



O comportamento do termo informação na Ciência da Informação

The behavior of the term information in "Information Science"

Marcos de Souza 

Doutor em Gestão e Organização do Conhecimento
Universidade Federal de Minas Gerais
marcosdesouza82@gmail.com

Fernanda Gomes Almeida 

Doutora em Gestão e Organização do Conhecimento
Universidade Federal de Minas Gerais
usernanda@gmail.com

Resumo

A informação enquanto conjunto de dados que os seres humanos dão sentido para interpretar fatos, por meio de um sistema de signos entre o transmissor e receptor, foi essencial para o desenvolvimento e sobrevivência da Ciência. Enquanto elemento que constitui a tríade dado, informação e conhecimento, também estudado na Ciência da Informação, o termo informação apresenta características generalistas quando analisado num contexto isolado. Dentre os objetivos, a pesquisa buscou: 1) identificar e discutir o comportamento do termo informação em pesquisas científicas da área de Ciência da Informação e; 2) realizar o mapeamento científico dos termos compostos formados por n-grama gerados a partir do termo raiz. A pesquisa empírica foi realizada por meio da construção de algoritmos utilizando a linguagem de programação Python que permitiu extrair a frequência de termos de 2.448 documentos. O termo informação, na sua forma raiz, tem apresentado comportamento contínuo e regular ao longo do período analisado, entretanto, os termos de composição como, por exemplo, "informação_tecnologia", "recuperação_informação" e "fonte_informação" somados, alcançam 41,98% da frequência do termo informação, apresentando maior especificidade de conceitos junto ao domínio da linguagem da Ciência da Informação. O real comportamento do termo informação se apresenta de maneira representativa e com maior especificidade através dos termos de composição, formados a partir do termo raiz. A interdisciplinaridade da Ciência da Informação contribuiu para o quantitativo e diversidade dos termos de composição extraídos do corpus.

Palavras-chave

Informação. Comportamento. Termo. Ciência da Informação.

Abstract

Information as a set of data that human beings give meaning to interpret facts, through a system of signs between transmitter and receiver, was essential for the development and survival of Science. As an element that constitutes the triad data, information, and knowledge, also studied in Information Science, the term information presents general characteristics when analyzed in an isolated context. Among the objectives, this research aims to: 1) identify and discuss the behavior of the term information in scientific research in the area of Information Science and; 2) perform the scientific mapping of compound terms formed by n-grams generated from the root term. The empirical research was carried out by constructing algorithms using the Python programming language that allowed extracting the frequency of terms from 2,448 documents. The term information, in its root form, has presented continuous and regular behavior throughout the analyzed period; however, the composition terms, such as "informação_tecnologia", "recuperação_informação" and "fonte_informação", to



gether, reach 41.98% of the frequency of the term information, presenting greater specificity of concepts within the domain of the Information Science language. The real behavior of the term information is presented representatively and with greater specificity through the composition terms, formed from the root term. The interdisciplinary nature of Information Science contributed to the quantity and diversity of the composition terms extracted from the corpus.

Keywords

Information. Behavior. Term. Information Science.

1 INTRODUÇÃO

Embora a Ciência da Informação seja uma área relativamente nova e que ainda busca um consenso sobre seu conceito entre as diferentes escolas, os termos dado, informação e conhecimento têm sido estudados ao longo dos anos nas áreas das ciências Exatas, Humanas e Sociais. Atrelada às tecnologias da informação, a tríade se tornou objeto de estudos da Ciência da Informação.

A diferenciação dos conceitos da tríade pode se tornar uma tarefa difícil, justamente por não possuir uma coesão entre autores sobre onde iniciam ou terminam cada um dos conceitos. Numa visão objetiva, Davenport (1998) refere-se a dado como a observação do estado do mundo; informação como dados com valores relevantes; e conhecimento como informação valiosa na mente humana. Le Coadic (1996) ressalta a importância da informação para desenvolvimento da ciência. Segundo esse autor, na ausência de informação, não existiria o conhecimento.

Esta pesquisa não pretende discutir aspectos epistemológicos, etimológicos ou filosóficos dos termos da tríade em questão, mas a forma que o termo informação se comporta ao longo do período estudado, em termos de surgimento, ápice e desuso. Partindo deste princípio, questiona-se: de que forma tem se apresentado o comportamento do termo informação na Ciência da Informação?

A pesquisa buscou: 1) identificar e discutir o comportamento do termo informação na área da Ciência da Informação e; 2) realizar o mapeamento científico dos termos de composição formados por n -gramas gerados a partir do termo raiz.

A composição é o processo de formação de palavras que consiste na concatenação de pelo menos duas bases (radical, tema ou palavra), formando uma nova unidade lexical a partir de duas unidades lexicais dotadas dos seus respectivos referenciais (RIO-TORTO, 1998). Por n -grama entende-se: um pedaço de n -caracteres extraídos de uma cadeia de caracteres, assumindo valores como 1 para unigrama, 2 para bigrama ou 3 para trigrama (SUK-KARIEH; PULMAN; RAIKES, 2003).

Pressupõe-se que o termo informação apresenta comportamento diferente de sua frequência extraída do *corpus*, mediante a composição de termos a partir do termo raiz, que apresenta significados específicos e de relevância ao domínio da linguagem explorada.

Justifica-se a importância da pesquisa uma vez que, realizado o mapeamento científico dos termos raiz e de composição, pode-se identificar os diferentes tipos de comportamentos de termos, que servem como indicador para pesquisadores em relação aos assuntos em ascensão ou descensão na área de Ciência da Informação. Além disso, o vocabulário de uma área está em constante evolução e a linguagem natural se apresenta variável, com pa-

lavras que caem de uso, modificam os significados ou surgem com o passar dos anos (PASCHOALIN; SPADOTO, 1996).

2 REFERENCIAL TEÓRICO

Na literatura científica é possível encontrar autores que enfatizam o marco inicial da Ciência da Informação a partir explosão informacional, destacando o grande volume de relatórios que documentavam a Segunda Guerra Mundial e para debater soluções tecnológicas (NHACUONGUE; FERNEDA, 2015).

A Ciência da Informação baseia-se em teorias e experimentos, bem como em combinações inter-relacionadas que envolvem informação, necessidade e uso da informação, conhecimento e registro do conhecimento em diferentes contextos como sociais, individuais, institucionais, além das tecnologias da informação para tratamento dessas questões (SARACEVIC, 1996).

A Ciência da Informação apresenta características interdisciplinares e que se manifesta em campos como biblioteconomia, administração, economia, artes gráficas, linguística, comunicação, psicologia, filosofia, sociologia, matemática, lógica, computação, pesquisa operacional, telecomunicações, direito, política entre outras áreas (BORKO, 1968; LE COADIC, 1996).

Historicamente, na Ciência da Informação, os termos dado, informação e conhecimento, podem ser reconhecidos por meio das concepções de Borko (1968), destacadas no texto *Information science: what is it?*, onde o autor apresenta fundamentações disciplinares como requisitos científicos para solução de problemas de informação e conhecimento durante processo da elaboração identitária da Ciência da Informação (SEMIDÃO, 2013).

É importante destacar que dados não possuem significados próprios, mas apresentam uma série de observações, medidas ou fatos, podendo ser representados em forma de números, palavras, sons ou imagens ao se produzir a informação. A informação é definida como dados organizados de forma significativa e que devem estar associados a um contexto de significado. Já o conhecimento está relacionado a consciência do entendimento adquirido por meio da experiência do sujeito, possuindo, assim, um sentido mais amplo da informação (BOISOT, 1998).

A diferenciação entre os termos da tríade pode se tornar uma tarefa difícil, pois não existe um consenso de onde inicia e termina cada conceito, entretanto, ao elaborar um processo relacionando os três conceitos, podem-se alcançar resultados representativos (DAVENPORT, 1998). A relação entre conhecimento e informação perpassa por um processo de interação, já que a geração do conhecimento depende da informação e a coleta da informação requer conhecimento por parte do indivíduo (BOISOT, 1998).

A convergência dos núcleos epistemológicos da Ciência da Informação tem conceituado a tríade dado, informação e conhecimento, considerando informação como processo na mudança de um estado de coisas sob o aspecto cognitivo nas áreas: 1) gestão da informação e do conhecimento – aperfeiçoamento do potencial das organizações; 2) recuperação da informação – melhoria nos processos de seleção de informações com objetivo no processo científico, bem como o aprimoramento de recursos tecnológicos para uma recuperação da informação mais próxima da analogia da cognição humana; 3) organização do conhecimento – domínio sobre formas de categorização e; 4) mediação – referente a emancipação para a contribuição da construção do conhecimento social e quebra hegemonia (SEMIDÃO; ALMEIDA, 2013).

Ao conceituar informação, Le Coadic (1996) apresenta um sistema de signos e a necessidade da informação se tornar existente entre o meio transmissor e receptor:

Informação é um conhecimento inscrito (gravado) sob a forma escrita (impressa ou numérica), oral ou audiovisual. A informação comporta um elemento de sentido. É um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espacial-temporal: impresso, sinal, elétrico, onda sonora, etc. Essa inscrição é feita graças a um sistema de signos (a linguagem), signo este que é um elemento da linguagem que associa um significante a um significado: signo alfabético, palavra, sinal, pontuação (LE COADIC, 1996, p. 5).

Por meio da conceituação do autor, torna-se possível perceber uma proximidade, mas, também, um distanciamento entre os termos informação e conhecimento na Ciência da Informação, enfatizado através da afirmação “A informação é o sangue da ciência. Sem informação, a ciência não pode se desenvolver e viver. Sem informação a pesquisa seria inútil e não existiria o conhecimento” (LE COADIC, 1996, p. 27). Setzer (1999) considera que informação possui característica objetiva-subjetiva, sendo a primeira, no sentido descritivo de uma forma, e a segunda, dependente do usuário. Além disso o autor ressalta que:

Informação é uma abstração informal, que representa algo significativo para alguém através de textos, imagens, sons ou animação. [...] Esta não é uma definição – isto é uma caracterização, porque ‘algo’, ‘significativo’ e ‘alguém’ não estão bem definidos; assumimos aqui um entendimento intuitivo desses termos. [...] Não é possível processar informação diretamente em um computador. Para isso é necessário reduzi-la a dados. [...] Uma distinção entre dado e informação é que o primeiro é puramente sintático e o segundo contém necessariamente semântica (SETZER, 1999, p. 7).

Em um aspecto filosófico, O’Brien (2003, p. 224) conceitua informação como “um conjunto de dados aos quais seres humanos deram forma para torná-los significativos e úteis” remetendo à associação que o ser humano realiza para interpretar os fatos. Considera-se informação como todo dado trabalhado ou tratado de maneira que possa gerar um sentido natural e lógico. No contexto tecnológico, esses dados necessitam ser processados para posteriormente gerar informações (FELIX, 2003; REZENDE, 2013).

A informação possui duas dimensões, sendo a primeira pessoal – relacionada à experiência pessoal, bem como a prática de vida; e a segunda, coletiva – fragmentos do conhecimento produzido ao longo dos anos pela sociedade (CARDOSO, 1996). As informações estão intrinsecamente relacionadas a dois conceitos: dados brutos e conhecimento. A informação é considerada “[...] dados dotados de relevância e propósito” (DRUCKER, 1988, p.46, tradução nossa). Os dados são a “simples observação do estado do mundo”, enquanto o conhecimento é a “informação valiosa na mente humana” (DAVENPORT, 1998, p.18).

A evolução da sociedade e o fator tecnológico contribuíram para uma categorização do termo informação, sendo não estruturada e estruturada. A informação não estruturada foi marcada por uma época onde o processo da produção e organização da informação era realizado por bibliotecários, administradores desse tipo de informação. Tratava-se de um processo que não era tão ágil, pois fazia uso de fontes impressas. Essa atividade foi, ao longo dos séculos, suprimida pela informação estruturada que apresentou uma abordagem mais popular com a informatização e minimizou o quantitativo de papéis, além de direcionar logicamente a informação, quantificar e disseminar a informação de maneira que fosse possível gerar conhecimento e reduzir custos com relação a pessoas envolvidas no processo (DAVENPORT, 1998).

Num viés administrativo, as informações, quando planejadas e disponibilizadas previamente aos interessados, apresentam um papel fundamental no processo de tomada de decisões nas organizações. Algumas características da informação são: a) possui conteúdo único; b) exige mais de duas palavras; c) deve apresentar detalhes em sua descrição sem apresentar generalidade; d) não pode ser formalizada por meio de um verbo; e) não apresentam características físicas ou virtuais e; f) podem ser organizados por assuntos, categorias, coisas, módulos, objetos, programas ou sistemas (REZENDE, 2013).

A informação é recurso fundamental para a sobrevivência de uma organização, uma vez que os dados brutos trabalhados e estruturados em sistemas de informação possibilitam aos gestores a realização da tomada de decisão de maneira estratégica, com base no conhecimento dos gestores em seus respectivos nichos de mercado (OLIVEIRA, 2012).

Abordar o tema informação sem contemplar o termo conhecimento é uma tarefa difícil. Boa parte dos autores da Ciência da Informação inter-relaciona o termo informação com comunicação, mensagem e conhecimento (CARDOSO, 1996). Ainda no contexto das inter-relações, Christovão e Braga (1997) destacam:

1) Informação pode ser definida como a interface, o evento entre um estímulo externo (mensagem) e um cognóscio que tal estímulo ou mensagem altera e 2) documentos contêm mensagens, as quais podem ou não produzir informação, dependendo do estado de conhecimento prévio/anterior do receptor humano (CHRISTOVÃO; BRAGA, 1997, p. 34).

Silva (2014) apresenta múltiplas inter-relações da informação no campo da Ciência da Informação, destacando os fundamentos técnico-pragmáticos, humanos e científicos.

Enquanto técnico-pragmático, o autor aborda a tríade gestão, processos e tecnologias inter-relacionando com a linguagem e a aplicação em centros de informação, destacando assim: informação na perspectiva dos processos, abordando o pragmatismo, físico, social e ontológico; informação na perspectiva da gestão, enfatizando o pragmatismo estratégico aos valores da informação; informação na perspectiva da gestão; informação na perspectiva da tecnologia; informação na perspectiva da linguagem e; informação em unidades de informação (SILVA, 2014).

Nos fundamentos humanos da informação, Silva (2014) realiza reflexões sobre a sociedade contemplando cultura, política/economia, educação e as inter-relações com ideologia e memória, destacando: informação e sociedade da informação social; informação e cultura sobre utilitarismo cultural e cultura digital da informação; informação e política/economia; informação e educação; informação e ideologia e; informação e memória.

Já nos fundamentos científicos da informação, são abordados: informação e ética; informação e valor; informação como fenômeno científico-natural; informação como fenômeno temporal-multitemporal; informação como fenômeno de objetividade, subjetividade e intersubjetividade e; informação como fenômeno pluri, inter e transdisciplinar (SILVA, 2014).

Souza (2020) apresenta uma classificação de comportamentos e características assumidas por termos extraídos de *corpus* e analisados durante um período temporal conforme apresentado no Quadro 1.

Quadro 1 – Comportamentos e características de termos.

<p>Comportamento contínuo – assiduidade, que não apresenta interrupções de frequência de termo ao longo do período analisado:</p> <p>a) regular – apresenta regularidade de frequência do termo com variações menores que 100%;</p> <p>b) irregular – apresenta irregularidade de frequência com variações maiores que 100% para cima e para baixo;</p> <p>c) ascensão – apresenta aumento de frequência do termo ao longo do período analisado superior a 100%;</p> <p>d) descensão – apresenta queda de frequência do termo ao longo ou período igual a 100% e se mantendo nulo nos seguintes.</p>	<p>Comportamento inconstante – apresenta ausência de frequência de termos entre o intervalo analisado:</p> <p>a) regular – apresenta regularidade de frequência em um determinado período analisado e que não ultrapassa 100%;</p> <p>b) irregular – apresenta irregularidades entre os anos analisados onde a frequência ultrapassa 100% para mais e para menos;</p> <p>c) ascensão – apresenta uma frequência crescente do termo nos últimos anos analisados ou período;</p> <p>d) descensão – apresenta queda de frequência dos últimos termos analisados ou período.</p>
---	---

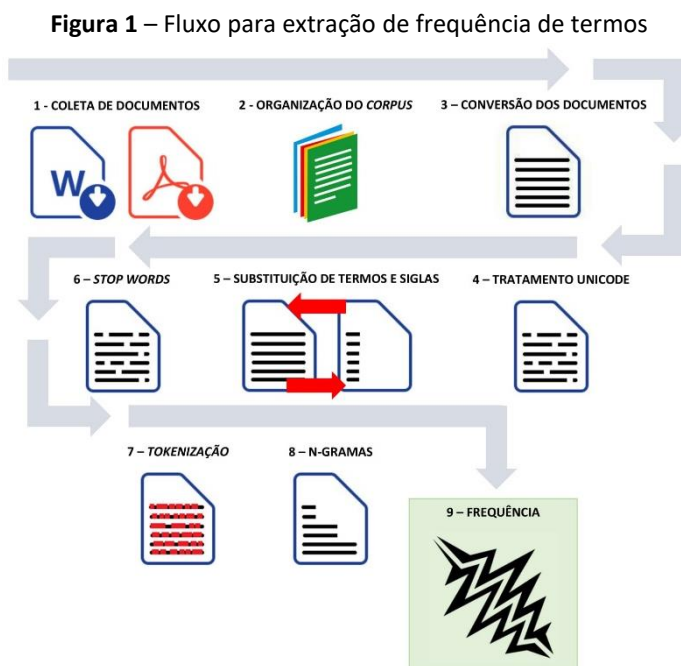
Fonte: (SOUZA, 2020).

Responder aos por quês das alterações dos comportamentos realizada ao longo de um determinado período pode ser tarefa difícil, uma vez que diferentes fatores podem contribuir para o aumento ou redução no quantitativo de frequências de termos extraídas a cada ano. Alguns desses fatores são: a interdisciplinaridade da área da Ciência da Informação; a independência de cada Grupo de Trabalho - GT do ENANCIB para aprovar diferentes quantitativos de pesquisas; o surgimento de novos programas de pós-graduação na área do domínio da linguagem e; a possibilidade de termos aparecerem em mais de um GT.

3 METODOLOGIA

Esta pesquisa se classifica quanto à finalidade/natureza como aplicada, quanto à abordagem do problema como quali-quantitativa e quanto aos objetivos como exploratória (GIL, 2010).

O referencial teórico da pesquisa foi construído com base em artigos científicos e livros indexados no Google Scholar, Google Books e Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Os dados empíricos foram adaptados de McKinney (2018) sendo: a) interação com o mundo externo – refere-se à coleta de documentos do tipo artigos completos e resumos expandidos dos anais do Encontro Nacional de Pesquisa em Ciência da Informação - ENANCIB para constituição do *corpus*. A coleta de dados foi realizada entre os dias 27 de abril e 01 de maio de 2018 e, posteriormente, acrescidos um novo conjunto de documentos no ano de 2019. Optou-se, nesta etapa, por selecionar textos unicamente no idioma português; b) preparação e pré-processamento – organização, limpeza, manipulação, combinação, normalização, tratamento dos dados; c) transformação – operações matemáticas e estatísticas aplicadas em grupos de conjuntos de dados e; d) apresentação – apresentação dos resultados conforme representado na Figura 1.



Fonte: (SOUZA; ALMEIDA, 2021).

Utilizou-se, para o desenvolvimento do algoritmo de extração das frequências dos termos do *corpus*, a linguagem de programação Python¹, juntamente com as bibliotecas PDFMiner², NLTK³ e plotly⁴. Os algoritmos para extração de frequência dos termos, bem como o gráfico dinâmico⁵ com o comportamento de termos de composição estão disponibilizados através do GtiHub⁶. A metodologia utilizada foi a mesma da pesquisa de Souza e Almeida (2021) intitulada: *O comportamento do termo dado na Ciência da Informação* (grifo nosso).

4 RESULTADOS E DISCUSSÕES

O *corpus* foi constituído por artigos científicos e resumos expandidos publicados nos anais do Encontro Nacional de Pesquisa em Ciência da Informação – ENANCIB entre os anos de 2012 e 2018.

Foram coletados 2.475 documentos e descartados 27 arquivos que apresentaram problemas de Unicode⁷ durante a conversão dos documentos, resultando, assim, em 2.448 documentos utilizados para extração de frequências conforme apresentado no Gráfico 1.

¹ Python. Disponível em: <https://www.python.org/>. Acesso em: 29/06/2021.

² PDFMiner. Disponível em: <https://pypi.org/project/pdfminer/>. Acesso em: 29/06/2021.

³ NLTK - *Natural Language Toolkit*. Disponível em: <https://www.nltk.org/>. Acesso em: 29/06/2021.

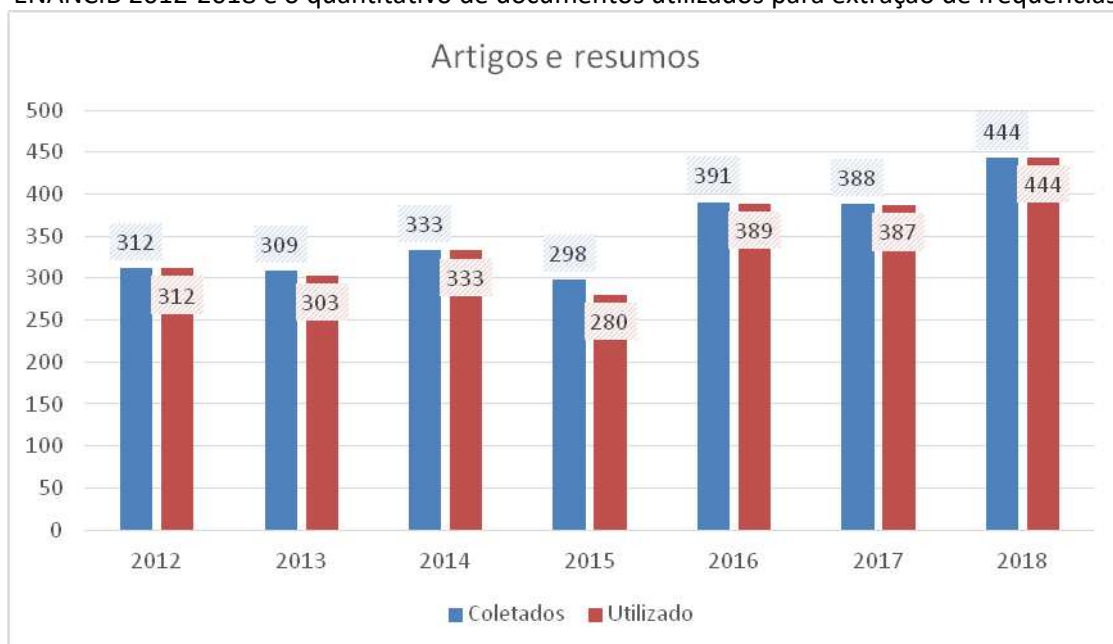
⁴ plotly. Disponível em: <https://plotly.com/>. Acesso em: 29/06/2021.

⁵ Algoritmo – Gráfico dinâmico contendo os termos de composição a partir do termo raiz. Disponível em: <https://bit.ly/3gCnzKR/>.

⁶ Github. Plataforma de desenvolvimento colaborativo para hospedar, revisar códigos, gerenciar projetos e criar *software* de maneira colaborativa. Algoritmo - Extração de termos. Disponível em: <https://bit.ly/2YTrJls/>.

⁷ Unicode - Padrão que permite aos computadores representar e manipular, de forma consistente, texto de qualquer sistema de escrita existente.

Gráfico 1 – Quantitativo anual de artigos completos e resumos expandidos publicados por anais do ENANCIB 2012-2018 e o quantitativo de documentos utilizados para extração de frequências



Fonte: Elaborado pelos autores.

Nas etapas de preparação e pré-processamento, os documentos que constituem o *corpus* foram convertidos⁸ para o formato TXT⁹, legível pelo computador. A conversão possibilitou a padronização e unificação dos documentos, além de uma redução do tamanho de 2GB para 94MB. Posteriormente, foi realizada a etapa de conversão de termos em siglas e vice-versa, por meio de Expressões Regulares. Esta conversão permite unificar termos com o mesmo significado de forma que não exista uma divisão de frequências, como, por exemplo, todas as siglas CI encontradas no *corpus* passam a ser convertidas para Ciência da Informação.

Utilizando a biblioteca NLTK foram retiradas do *corpus* as *stop words* – palavras de parada como, por exemplo, “e”, “ou”, “para”, “com”, que possuem frequências elevadas, entretanto, baixa relevância junto ao domínio da linguagem. A biblioteca, em questão, possui uma lista padrão ao qual foi adicionada uma outra lista de palavras de parada aderentes ao domínio da linguagem.

Na etapa de transformação dos dados, também utilizando a biblioteca NLTK, foi realizada a *tokenização*, que divide os textos em frases, palavras, símbolos e outros elementos que são convertidos em uma sequência de palavras separadas por tratamento de pontuação e espaçamento.

Posteriormente, foi executada a função *n*-gramas que resultou em um total de 6.640.564 unigramas, 6.638.116 bigramas e 6.635.668 trigramas. A partir destes resultados, foram geradas listas contendo os mil primeiros termos de cada tipo de *n*-grama que gerou a apresentação dos resultados através de gráfico dinâmico. A quantidade de termos por lista de *n*-grama está relacionada ao fator de qualidade e representatividade do termo. Como exemplo de baixa representatividade, constam o bigrama “informação_autor” e o trigramas “gestão_informação_apoia”.

⁸ Algoritmo - Conversão de documentos. Disponível em: <https://bit.ly/2D19vfr/>.

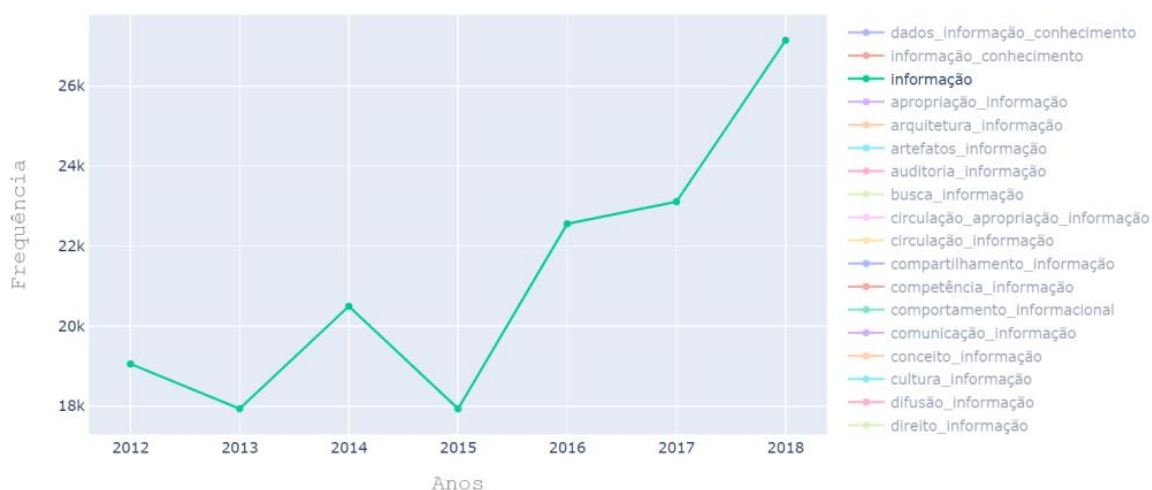
⁹ TXT - Ficheiros de texto que podem conter texto simples, sem formatação.

O comportamento diacrônico dos termos pode ser identificado mediante a frequência dos *n*-gramas extraídos dos *corpora* em intervalos diferentes, permitindo assim realizar uma classificação de comportamento dos termos que contribuem para um melhor entendimento da área estudada (SOUZA, 2020).

O termo informação alcançou frequência de 148.237, extraída de 2.448 artigos completos e resumos expandidos que constituem o *corpus*. No ano de 2012, o termo apareceu 19.056 vezes, enquanto no ano de 2013, o termo apresentou frequência de 17.936, o que significa queda de -6% referente ao ano anterior. No ano de 2014, o termo apresentou frequência de 20.495, equivalente a um aumento de 14%. Já no ano de 2015, houve uma queda no quantitativo de termos em -12%, resultando numa frequência de 17.934. No ano de 2016, o termo apresentou frequência de 22.559, representando um aumento de 26%. No ano de 2017, foi extraída a frequência de 23.111, equivalente a um aumento de 2%. Por fim, no último ano analisado, o termo alcançou frequência de 27.146 e um aumento de 17% conforme apresentado no Gráfico 2.

Gráfico 2 – Comportamento do termo Informação

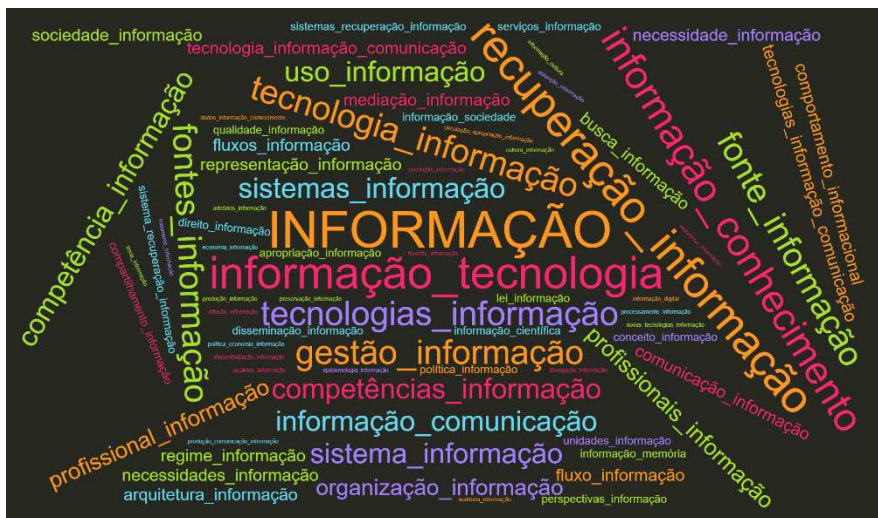
Frequência dos termos de composição a partir do termo "Informação"



Fonte: Elaborado pelos autores.

As alterações de comportamento do termo no que se refere ao crescimento ou queda de percentual relaciona-se ao ano anterior. Entre o período mínimo e máximo analisado, o termo informação apresentou uma diferença acumulada de frequência equivalente a 42%, o que remete a um aumento significativo do uso do termo nas produções científicas da área da Ciência da Informação. Cabe ressaltar que o termo informação, contido no *corpus*, apresenta composições de termos dos tipos bigramas e trigramas com significados diferentes ao do termo raiz. A Figura 2 apresenta 64 termos de composição extraídos a partir de frequências.

Figura 2 – Termos de composição a partir do termo Informação



Fonte: Elaborado pelos autores.

Os termos de composição apresentam frequências menores quando comparado ao termo raiz, seja a cada ano ou no acumulado entre o período analisado. São exemplos de termos de composição os bigramas “informação_tecnologia”, com frequência de 4.959, “recuperação_informação”, com frequência de 4.001, “uso_informação”, com frequência de 2.391 e o trigrama “tecnologias_informação_comunicação”, com frequência de 1.127 numa lista de 64 termos que totalizam frequência de 62.244, equivalente a 41,98% da frequência do termo informação.

Trata-se de um percentual representativo, uma vez que o volume de termos de composição extraídos do *corpus* apresenta valores significativos. Além dos termos apresentarem características de relevância ao domínio da linguagem, torna-se possível identificar o mapeamento científico do termo conforme apresentado no Quadro 2.

Quadro 2 – Termos de composição e frequências.

informação_tecnologia,4959; recupera-	conceito_informação,746;
ção_informação,4001; informa-	sistemas_recuperação_informação siste-
ção_conhecimento,3381;	ma_recuperação_informação,
tecnologia_informação tecnologi-	714;
as_informação,2720; fonte_informação fon-	informação_memória,710;
tes_informação,2713; gestão_informação,2579;	compartilhamento_informação,697;
sistema_informação sistemas_informação,2473;	lei_informação,694; direito_informação,671;
uso_informação, 2391; informa-	unidades_informação,664; políti-
ção_comunicação,2389; competên-	ca_informação,592; qualidade_informação,522;
cia_informação competên-	tratamento_informação,475; usuá-
cias_informação,2122; profissional_informação	rios_informação,467; produção_informação,452;
profissionais_informação,1730; organiza-	informação_digital,387; produ-
ção_informação,1656; arquitetu-	ção_comunicação_informação,299;
ra_informação,1445; socieda-	informação_cultura,262; circula-
de_informação,1320; representa-	ção_apropriação_informação,246;
ção_informação,1272; necessida-	
des_informação necessidade_informação,1229;	

fluxo_informação fluxos_informação,1188; tecnologias_informação_comunicação tecnolo- gias_comunicação_informação,1127;	economia_informação,246; seguran- ça_informação,241; políti- ca_economia_informação,217; epistemolo- gia_informação,211; processamen- to_informação,207; troca_informação,205; filoso- fia_informação,175; no-
regime_informação,1105; comunica- ção_informação,1081; busca_informação,1075;	vas_tecnologias_informação,168; divulga- ção_informação,143; obtenção_informação,120; disponibilização_informação,101; circula- ção_informação,97; preservação_informação,67; auditoria_informação,53; da-
comportamento_informacional,1056; media- ção_informação,1054; informa- ção_sociedade,981; perspecti- vas_informação,911;	dos_informação_conhecimento,47; cultu- ra_informação,32; difusão_informação,31; artefa- tos_informação,25.
informação_científica,892;	
disseminação_informação,862; apropria- ção_informação,796;	
serviços_informação,752;	

Fonte: Elaborado pelos autores.

Dentre os comportamentos dos termos de composição, o termo “fonte_informação”, representado pela cor azul, apresentou frequência de 328 no ano de 2012; 361 em 2013, o que significa um crescimento de 10%; 452 em 2014, apresentando um crescimento de 25%; 368 em 2015, equivalente a uma queda de -19%; 339 em 2016, apresentando novamente queda de -8%; 412 em 2017, representando um aumento de 22% e; 453 em 2018, com crescimento de 10%.

O termo “recuperação_informação”, representado pela cor amarela, apresentou frequência de 497 no ano de 2012; 431 em 2013, equivalente a queda de -13%; 602 em 2014, o que significa um crescimento de 40%; 716 em 2015, apresentando crescimento de 19%; 564 em 2016, equivalente a queda de -21%; 632 em 2017, representando um crescimento de 12% e; 559 em 2018, equivalente a queda de -12%.

Já o termo “sistema_informação”, representado pela cor laranja, apresentou frequência de 428 no ano de 2012; 310 em 2013, equivalente a queda de -28%; 359 em 2014, resultando num crescimento de 16%; 282 em 2015, apresentando queda de -21%; 278 em 2016, apresentando queda de -1%; 342 em 2017, o que significa um crescimento de 23% e; 474 em 2018, apresentando crescimento de 39%. Todas as alterações comportamentais dos termos, seja crescimento ou queda de números e percentuais, são referentes ao ano anterior conforme, apresentado no Gráfico 3.

Todos os termos apresentados no gráfico apresentam comportamento contínuo com frequência em todos os períodos analisados e característica regular, apresentando variações de frequências menores que 100% entre o intervalo mínimo e máximo, ao longo do período analisado, sendo: 38% para “fluxo_informação”; 12% para “recuperação_informação” e; 11% para “sistema_informação”.

Termos como “fonte_informação”, “gestão_informação”, “uso_informação”, “competência_informação”, “organização_informação”, “arquitetura_informação”, “sociedade_informação”, “representação_informação” e “necessidades_informação” apresentam o mesmo comportamento e a mesma característica que os termos apresentados no Gráfico 3.

Gráfico 3 – Comportamento de termos de composição

Frequência dos termos de composição a partir do termo “Informação”



Fonte: Elaborado pelos autores.

O termo “informação_tecnologia”, representado pela cor lilás, apresentou frequência de 458 no ano de 2012; 496 no ano de 2013, o que significa um crescimento de 8%; 529 em 2014, o que representa um crescimento de 7%; 560 em 2015, apresentado crescimento de 6%; 756 em 2016, novamente com crescimento de 35%; 1.007 no ano de 2017, equivalente a um crescimento de 33% e; 1.153 em 2018, com crescimento de 14%.

O termo “competência_informação”, representado pela cor vermelha, apresentou frequência de 106 no ano de 2012; 132 no ano de 2013, o que significa um crescimento de 25%; 126 em 2014, equivalente a uma queda de -5%; 239 no ano de 2015, apresentando um crescimento de 90%; 367 em 2016, referente a um crescimento de 54%; 449 em 2017, ou crescimento de 22% e; 703 em 2018, ou crescimento de 22%.

O termo “regime_informação”, representado pela cor azul, apresentou frequência de 92 no ano de 2012; 150 em 2013, o que significa crescimento de 63%; 97 no ano de 2014, equivalente a uma queda de -35%; 131 em 2015, apresentando crescimento de 35%; 124 em 2016, equivalente a uma queda de -5%; 201 no ano de 2017, representando um crescimento de 62% e; 310 em 2018, representando crescimento de 54%, conforme apresentado no Gráfico 4. Faz-se necessário ressaltar que todas as alterações de comportamentos dos termos referem-se ao ano anterior.

Os termos do Gráfico 4 apresentam comportamentos e características similares, sendo contínuo por apresentar frequência durante todo o intervalo analisado e, ascensão, por apresentar aumento de frequência ao longo do período analisado superior a 100%, sendo 153% para “informação_tecnologia”, 563% para “competência_informação” e 237% para “regime_informação”. Ressalta-se que, a partir de 2015, ano com menor número de publicações nos anais do ENANCIB, os termos “informação_tecnologia” e “competência_informação” começaram a ganhar notoriedade, mantendo curva de crescimento até o final do período estudado.

Gráfico 4 – Comportamento de termos de composição

Frequência dos termos de composição a partir do termo "Informação"



Fonte: Elaborado pelos autores.

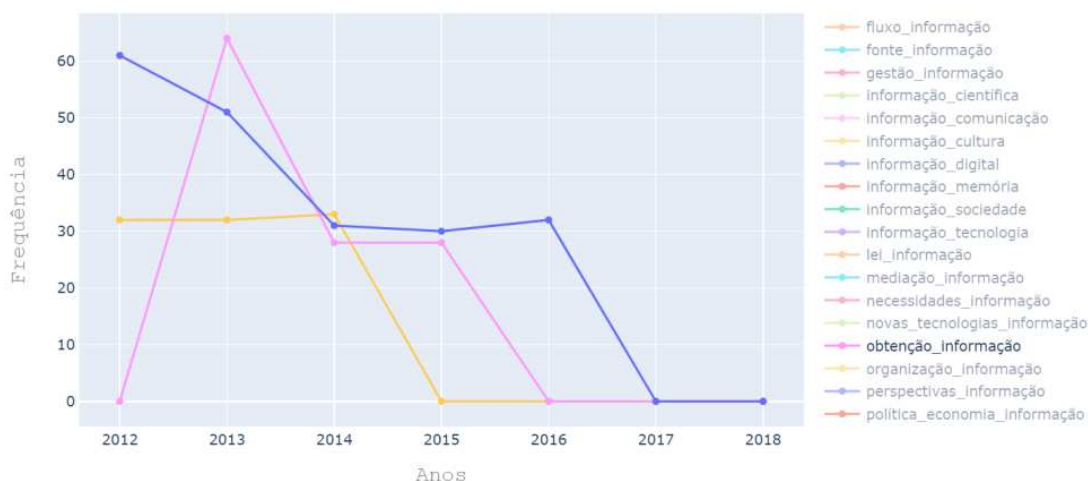
O termo "troca_informação", representado pela cor azul, apresentou frequência de 61 no ano de 2012; 51 em 2013, equivalente a queda de -16%; 31 em 2014, novamente queda de -39%; 30 em 2015, o que significa queda de -3%; 32 em 2016, o que significa um crescimento de 7%. Entre os anos de 2017 e 2018 o termo não apresentou frequência ou apresentou menções somente após o milésimo termo do tipo bigrama extraído do *corpus*.

O termo "obtenção_informação", representado pela cor rosa, não apresentou frequência nos anos 2012 e de 2016 a 2018. No ano de 2013 o termo apresentou frequência de 64; 28 em 2014, equivalente a uma queda de 56%, mantendo o mesmo valor de frequência para o ano de 2015.

Já o termo "circulação_informação" apresentou frequência nos três primeiros anos analisados, sendo em 2012 e 2013 igual a 32 e em 2014, frequência de 33, o que significa um crescimento de 3%, conforme apresentado no Gráfico 5.

Gráfico 5 – Comportamento de termos de composição

Frequência dos termos de composição a partir do termo "Informação"



Fonte: Elaborado pelos autores.

Os termos do Gráfico 5 apresentam comportamento inconstante, com ausência de frequência em determinados períodos e característica de descensão, apresentando queda de frequência nos últimos anos analisados ou em períodos, de maneira que o termo não apresenta mais frequências, caracterizando, assim, um desuso dos termos nas publicações científicas do domínio da linguagem.

5 CONSIDERAÇÕES FINAIS

O termo informação analisado no *corpus* de documentos científicos da área Ciência da Informação apresentou comportamento contínuo com característica regular ao longo do período analisado, representando, assim, um acúmulo de frequências igual a 148.237 e média anual de 21.177, sendo extraídas de 2.448 documentos. Embora o termo apresente característica regular, entre o intervalo mínimo e máximo analisado, o termo resultou em crescimento de frequência de utilização em pesquisas científicas na área de Ciência da Informação equivalente a 42%.

O termo informação, quando analisado de maneira isolada, possui características generalistas e se torna dependente de termos como dado para que se possa gerar conhecimento através de processos que envolve a tríade. Entretanto, o real comportamento do termo informação se apresenta de maneira representativa e com maior especificidade através dos termos de composição, formados a partir do termo raiz. São exemplos os termos “informação_tecnologia” com frequência de 4.959, “recuperação_informação” com 4.001, “fonte_informação” com 2.713, “sistema_informação” com 2.473, de um grupo de 64 termos do tipo bigrama e trigrama, extraídos do *corpus*, representando 41,98% de toda a frequência do termo informação.

A interdisciplinaridade da Ciência da Informação contribuiu para o quantitativo e diversidade dos termos de composição extraídos do *corpus*. Dentre os termos de composições extraídos, os comportamentos e características identificadas, constatou-se que o maior volume de termos possui característica contínua e regular. Tal característica entre os resultados representa uma estabilidade dos termos no domínio da linguagem estudada. Faz-se necessário enfatizar que o comportamento dos termos pode ser flutuante de acordo com o intervalo analisado, bem como a inserção ou exclusão de períodos / documentos.

Confirmou-se o pressuposto da pesquisa ao constatar que o comportamento do termo informação, identificado através de frequências, não representa sua realidade, já que os termos de composição que compõem o mapeamento científico do termo apresentam comportamentos diferentes, além de apresentarem significados de maior relevância ao domínio de linguagem quando comparado ao termo raiz. Sugere-se, para pesquisas futuras, o estudo do comportamento do termo conhecimento na Ciência da Informação, bem como a comparação entre os termos que compõem a tríade dado, informação e conhecimento.

REFERÊNCIAS

BOISOT, M. **Competitive advantage in the information economy**. Oxford; New York: Oxford University Press, 1988.

BORKO, H. Information science: what is it? **American Documentation**, v.19, n.1, p.3-5, 1968.

CARDOSO, A.M.P. Pós-modernidade e informação: conceitos complementares? **Perspectivas em Ciência da Informação**. Belo Horizonte, v. 1, n. 1, p. 63-79, 1996.

CHRISTOVÃO, H. T.; BRAGA, G. M. Ciência da informação e sociologia do conhecimento científico: a intertematicidade plural. **Transinformação**, Campinas, v. 9, n. 3, p. 15–32, 1997.

DAVENPORT, T. E. **Ecologia da informação**: por que só a tecnologia não basta para o sucesso na era da informação. São Paulo: Futura, 1998.

DRUCKER, P.F. The Coming of the New Organization. **Harvard Business Review**, Boulder, v. 66, n.1, p. 45-53, 1998.

FELIX, W. **Introdução à Gestão da Informação**. Campinas: Alínea, 2003.

GIL, A. C. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

LE COADIC, Yves-François. **A ciência da informação**. Brasília: Briquet de Lemos, 1996.

MCKINNEY, W. **Python para análise de dados**: tratamento de dados com pandas, numpy e ipython. São Paulo: Novatec, 2018.

NHACUONGUE, J. A.; FERNEDA, E. O campo da ciência da informação: contribuições, desafios e perspectivas. **Perspectivas em Ciência da Informação**. Belo Horizonte, v. 20, n. 2, p. 3-18, 2015.

O'BRIEN, J. A. **Sistemas de Informação e as decisões gerenciais na era da internet**. 9. ed. São Paulo: Saraiva, 2003.

OLIVEIRA, D. P. R. **Sistemas de Informações Gerenciais: Estratégica - Táticas - Operacionais**. 15. ed. São Paulo: Atlas, 2012.

PASCHOALIN, M. A.; SPADOTO, N. T. **Gramática**: teoria e exercícios. São Paulo: FTD, 1996.

REZENDE, D. A. **Sistemas de Informações Organizacionais**: guia prático para projetos em cursos de Administração Contabilidade Informática. 5. ed. São Paulo: Atlas, 2013.

RIO-TORTO, G. M. Mecanismos de produção lexical no português europeu. **Alfa**, v.42, n.esp, p. 15-32, 1998. Disponível em: <https://periodicos.fclar.unesp.br/alfa/issue/view/298>. Acesso em: 01 jul. 2021.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**. Belo Horizonte, v. 1, n. 1, p. 41-62, 1996.

SEMIDÃO, R. A. M. Dados, Informação e Conhecimento: elementos de análise conceitual. **DataGramZero**, Rio de Janeiro, v. 14, n. 4, p. 10, 2013. Disponível em: <https://brapci.inf.br/index.php/article/download/52967>. Acesso em: 01 jul. 2021.

SEMIDÃO, R. A. M.; ALMEIDA, C. C. Tríade dados, informação e conhecimento: elementos de compreensão epistemológica da ciência da informação. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 14., 2013, Florianópolis. **Anais...** Florianópolis: UFSC, 2013.

SETZER, V. Dado, informação, conhecimento e competência. **DataGramZero**, Rio de Janeiro, n. 0, dez. p. 1-14, 1999. Disponível em: <https://www.ime.usp.br/~vwsetzer/datagrama.html>. Acesso em: 01 jul. 2021.

SILVA, J. L. C. **Múltiplas interpolações da informação no campo da Ciência da Informação no âmbito dos fundamentos técnico-pragmáticos, humanos e científicos**. 2014. 490 f. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto de Ciência da Informação, Universidade Federal da Bahia, Salvador, 2014. Disponível em: <https://repositorio.ufba.br/ri/handle/ri/17065>. Acesso em: 01 nov. 2020.

SOUZA, M. **O comportamento de termos da Ciência da Informação por meio da modelagem de tópicos**. 2020. 404 f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Programa de Pós-Graduação em Gestão e Organização do Conhecimento, Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2020. Disponível em: <https://repositorio.ufmg.br/handle/1843/34292>. Acesso em: 01 nov. 2020.

SOUZA, M.; ALMEIDA, F. G. O comportamento do termo dado na ciência da informação. **Ciência da Informação em Revista**. Maceió, v. 8, n. 2, p. 39-54, 2021. Disponível em: <https://www.seer.ufal.br/index.php/cir/article/view/11764/>. Acesso em: 13 nov. 2021.

SUKKARIEH, J. Z.; PULMAN, S. G.; RAIKES, N. Auto-marking: using computational linguistics to score short, free text responses. *In: ANNUAL CONFERENCE OF THE INTERNATIONAL ASSOCIATION FOR EDUCATIONAL ASSESSMENT*, 29, 2003, Manchester. Proceedings... [S.l.]: IAEA, 2003. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.108.7417&rep=rep1&type=pdf>. Acesso em: 01 jul. 2020.