

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO (UFRJ)  
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS (CCJE)  
FACULDADE DE ADMINISTRAÇÃO E CIÊNCIAS CONTÁBEIS (FACC)  
CURSO DE BIBLIOTECONOMIA E GESTÃO DE UNIDADE DE INFORMAÇÃO (CBG)

**FRANCIELLY DOMINGUES QUINTAS PEREIRA**

A RECUPERAÇÃO DA INFORMAÇÃO EM SITES DE COMÉRCIO ELETRÔNICO

Rio de Janeiro

2018

FRANCIELLY DOMINGUES QUINTAS PEREIRA

**A RECUPERAÇÃO DA INFORMAÇÃO EM SITES DE COMÉRCIO ELETRÔNICO**

Trabalho de conclusão de curso apresentado ao Curso de Biblioteconomia e Gestão de Unidades de Informação da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de bacharel em Biblioteconomia.

Orientadora Prof<sup>a</sup>. Dr<sup>a</sup>.: Juliana de Assis

Rio de Janeiro

2018

## Ficha catalográfica

P436r

Pereira, Francielly Domingues Quintas

A recuperação da informação em sites de comércio eletrônico. / Francielly Domingues Quintas Pereira. -- Rio de Janeiro, 2018.

46 f.: il.

Orientadora: Juliana de Assis

Trabalho de Conclusão de Curso (Bacharel em Biblioteconomia e Gestão de Unidades de Informação) – Curso de Biblioteconomia e Gestão de Unidades de Informação, Universidade Federal do Rio de Janeiro.

1. Recuperação da informação. 2. Comércio eletrônico. 3. Representação da informação. 4. Organização da informação. I. Assis, Juliana de. II. Título.

CDD 005.74

**FRANCIELLY DOMINGUES QUINTAS PEREIRA**

**A RECUPERAÇÃO DA INFORMAÇÃO EM SITES DE COMÉRCIO ELETRÔNICO**

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia e Gestão de Unidades de Informação da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de bacharel em Biblioteconomia.

Rio de Janeiro, xx de xxxxxxxx de xxxx.

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Juliana de Assis – UFRJ  
Orientador (a)

---

Prof. Dr.<sup>a</sup> Patricia Mallman – UFRJ  
Membro interno

---

Prof.<sup>a</sup> Ma. Carla Beatriz – UFRJ  
Membro interno

Agradeço à Deus, familiares e amigos.



## AGRADECIMENTOS

Primeiramente gostaria de começar agradecendo à Deus, por ter me guiado com sabedoria durante toda a caminhada dentro da universidade, e me protegido ao longo desses últimos 4 anos de curso na Graduação. Agradeço por ter iluminado meus pensamentos com amor e mantido minha paz de espírito, e minha saúde, tanto mental e física fortes até aqui. Sem Ele hoje não teria forças suficientes para concluir este trabalho e o curso de Biblioteconomia.

Gostaria de agradecer também à minha família querida, minha mãe e minha avó. Minha mãe Gláucia Domingues que sempre esteve ao meu lado e sempre confiou no meu potencial, sempre acreditou que a educação é a base de tudo para que o ser humano seja rico eternamente em seu intelecto, pois todo aprendizado que carregamos nunca nos será roubado. Minha mãe fez questão de me colocar em boas escolas, de me fazer aprender um novo idioma, e graças a ela e sua luta diária hoje estou concluindo uma etapa muito importante de minha vida acadêmica, falando fluentemente inglês. Agradeço à minha avó, Maria Cirlene, por também confiar e apoiar meus ideais, e acreditar que a educação sempre viria em primeiro lugar para mim. Elas duas são a base de tudo na minha vida, meu porto seguro. Agradeço também ao meu namorado Eduardo por me trazer conforto, e me ajudar nas horas de estresse, por me ajudar a fazer alguns trabalhos da faculdade, e por me envolver de amor e carinho quando mais precisava.

Aos meus professores, agradeço por todo aprendizado que me foi passado durante esses 4 anos de Graduação, fico grata pelos ensinamentos e pelo convívio harmonioso.

Não poderia deixar de agradecer também a todos os amigos queridos que conheci graças à UFRJ. Bruna Teresa, Fernanda, Deborah, Eli, Priscila, Lorena, Gabriel, Mauricio, Aneli e Luna. Graças a vocês voltei a valorizar a importância que os amigos têm em nossas vidas. O tempo que passei na UFRJ ao lado de vocês foi intenso e divertido. Passamos por muitas aventuras juntos, rimos, choramos, brigamos, bebemos muito em chopadas e dançamos até cansar. Comemos bastante salgadinho e batata frita na Angela Lanches após as aulas, e algumas vezes durante elas também. Mesmo tendo personalidades totalmente diferentes, nós fomos a família um do outro, dando conselhos e tentando ajudar ao máximo nos problemas que surgiam ao longo desse período. Eu não sou de chorar, mas devo admitir que sinto um aperto imenso por saber que vocês não farão mais parte do meu cotidiano, a saudade já começou a me sufocar. Porém sou eternamente grata por vocês existirem, e por

terem feito a minha vida muito melhor no Fundão. Obrigada por tudo galera. Amo todos vocês!



## RESUMO

Aborda a recuperação da informação em sites de comércio eletrônico, nacionais e internacionais, a partir de uma observação feita de diferentes produtos ofertados nas lojas do Mercado Livre e Aliexpress, sob a perspectiva do profissional da ciência da informação, trabalhando os conceitos de comércio eletrônico, representação da informação, organização da informação e organização do conhecimento. Possui como objetivo geral analisar o modelo utilizado de representação da informação em sites de comércio eletrônico visando contribuir para a recuperação da informação. Define-se o conceito de comércio eletrônico e representação da informação para descobrir qual é o método utilizado nos sites escolhidos. Definiu-se como surgiram os sites Aliexpress e Mercado Livre, explicando seu histórico de criação e quem são seus criadores. Apresenta uma abordagem qualitativa. Como campo empírico analisou-se o modo como os termos são representados e as ferramentas utilizadas nos sites de comércio eletrônicos utilizados na pesquisa. No recorte empírico optou-se por selecionar 58 membros de dois grupos do Whatsapp. As técnicas de coleta de dados são a observação participante e o questionário. Constatou-se a preferência dos usuários por sites de comércio eletrônico nacionais; a utilização de uma grande quantidade de termos para descrição dos produtos; o uso de palavras como melhor forma de se recuperar um item do *site*. Concluiu-se que o melhor critério para uma recuperação da informação satisfatória seria a utilização de um número significativo de termos, por meio de palavras, que especifiquem o produto com detalhes, para que sua representação seja feita de modo correto, com resultados de busca satisfatórios.

**Palavras-chave:** Recuperação da Informação. Comércio Eletrônico. Representação da Informação. Organização da Informação. Organização do conhecimento. Aliexpress. Mercado Livre.

## ABSTRACT

It addresses the retrieval information on national and international electronic commerce sites, based on an observation made of different products offered at Mercado Livre and Aliexpress, based on the professional perspective of Science Information, working with concepts about e-commerce, Information Representation, Information Organization and Knowledge Organization. The objective of this study is to analyze the used model of information representation in e-commerce sites aiming to contribute to the information retrieval. Define the concept of e-commerce and representation of information to find out which method is used in the chosen sites. It was defined how the sites Aliexpress and Mercado Livre appeared, explaining their creation history and who their creators are. It presents a qualitative approach. As an empirical field we analyzed how the terms are represented and the tools used in the e-commerce sites used in the research. In the empirical cut, 58 members from two groups of the Whatsapp were chosen. The techniques of data collection are participant observation and the questionnaire. Users' preference for national e-commerce sites was verified; to use a large number of terms to describe the products; the use of words as the best way to recover an item from the site. It was concluded that the best criterion for a satisfactory retrieval of information would be the use of a significant number of terms, through words, that specify the product in detail, so that its representation is done correctly, with satisfactory search results.

**Key-words:** Retrieval Information. E-commerce. Information Representation. Information Organization. Knowledge Organization. Aliexpress. Mercado Livre.

## LISTA DE ILUSTRAÇÕES

<b>Quadro 1</b> - Sites de comércio eletrônico e suas ferramentas.....	20
<b>Figura 1</b> - Barra de navegação do Aliexpress.....	21
<b>Figura 2</b> - Barra de navegação do Mercado Livre.....	22
<b>Figura 3</b> - Exemplos de taxonomia no site Aliexpress.....	23
<b>Figura 4</b> - Exemplos de taxonomia no site Mercado Livre.....	23
<b>Figura 5</b> - Termos para definir o item bolsa no Aliexpress.....	24
<b>Figura 6</b> - Termos para definir o item bolsa no Mercado Livre.....	24
<b>Gráfico 1</b> - Sites de comércio eletrônico mais utilizados.....	27
<b>Gráfico 2</b> - Sites com termos mais precisos.....	28

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>04</b>
1.1	PROBLEMA.....	05
1.2	JUSTIFICATIVA.....	05
1.3	OBJETIVOS.....	06
1.3.1	OBJETIVO GERAL.....	06
1.3.2	OBJETIVOS ESPECÍFICOS.....	06
<b>2</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>07</b>
2.1	COMÉRCIO ELETRÔNICO .....	07
2.2	REPRESENTAÇÃO DA INFORMAÇÃO.....	10
2.3	ORGANIZAÇÃO DA INFORMAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO.....	12
2.4	RECUPERAÇÃO DA INFORMAÇÃO.....	16
<b>3</b>	<b>METODOLOGIA .....</b>	<b>19</b>
3.1	CAMPO EMPÍRICO.....	19
3.2	TÉCNICAS DE COLETA E ANÁLISE DE DADOS.....	25
3.3	POPULAÇÃO/AMOSTRA.....	26
<b>4</b>	<b>ANÁLISE DE DADOS.....</b>	<b>27</b>
<b>5</b>	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>31</b>
	<b>REFERÊNCIAS.....</b>	<b>33</b>
	<b>APÊNDICE A – QUESTIONÁRIO.....</b>	<b>37</b>

## 1 INTRODUÇÃO

A recuperação da informação em sites de comércio eletrônico ainda possui algumas dificuldades que estão apresentadas neste trabalho, e a principal delas é como é feita a representação da informação relacionada aos produtos que estão disponíveis em catálogos de lojas virtuais.

Muitas vezes quando o consumidor vai efetuar uma busca simples por um item, seja ele de qualquer categoria, ocorrem alguns problemas para se obter resultados satisfatórios, pois muitas das vezes um produto está representado pela sua imagem com uma descrição em um pequeno texto relacionado às suas características. Um exemplo seria a utilização excessiva de termos para nomear o produto, o que costuma ser feito pelo vendedor da loja em questão, ao exemplo do Mercado Livre, fazendo com que isso dificulte o usuário do site de encontrar o que ele deseja. O vendedor costuma utilizar essa técnica para que o comprador encontre o item de forma mais rápida, o que quase nunca acontece. Outros sites, como o Aliexpress, utilizam *bots*, ou seja, robôs que indexam automaticamente os produtos na loja. Como muitos sabem o Aliexpress é um site internacional que pertence ao grupo Alibaba Group<sup>1</sup>, e por ser uma rede conectada ao mundo todo, possui um sistema de tradução automática para o idioma do país em que o comprador está efetuando o acesso. A tradução dos termos também seria uma outra problemática para a recuperação da informação.

Observa-se que o consumidor costuma fazer sua pesquisa pelo produto por meio de termos que estejam de certa forma diretamente relacionados às suas características físicas, como por exemplo, suas cores, o material de que é feito, suas medidas, entre outros. No entanto, muitas vezes ocorrem problemas nos resultados de busca obtidos devido à falta de um critério para definição de tais termos, ocultando itens e interrompendo muitas das vezes a finalização de uma compra.

Para um resultado de busca satisfatório, o vendedor poderia utilizar uma quantidade menor de termos para tornar a pesquisa feita pelo usuário do site mais objetiva, e os responsáveis pela administração dos sites poderiam construir um manual com regras básicas de representação da informação, porém pode ser que alguns sites já utilizem essa técnica para facilitar as pesquisas, apesar de não haver uma divulgação ampla sobre isso.

O trabalho se propõe a fazer uma análise do modelo utilizado de representação da informação em sites de comércio eletrônico visando contribuir na recuperação da informação,

---

<sup>1</sup> Site do grupo Alibaba que promove interação comercial e social entre centenas e milhares de usuários ao redor do mundo <https://www.alibabagroup.com/en/global/home>

além de verificar quais os métodos utilizados para inclusão de produtos nas lojas virtuais que serão usadas de exemplo para o estudo. O projeto também irá analisar qual é a percepção que o usuário possui em relação aos resultados obtidos na pesquisa, e a partir disso conseguir propor quais são os melhores critérios que deverão ser utilizados para viabilizar melhorias na recuperação da informação.

Tanto a Recuperação da Informação quanto a Representação da Informação, estão relacionadas à utilização de sites de comércio eletrônicos, sejam eles nacionais ou internacionais, com a utilização de mecanismos de buscas e da maneira como é feita a relação semântica entre os termos pesquisados.

### 1.1 PROBLEMA

De acordo com a literatura baseada em Sistemas de Recuperação da Informação e também de Representação da Informação, quais seriam os critérios necessários para um resultado efetivo da busca de informação em sites de comércio eletrônico?

### 1.2 JUSTIFICATIVA

O uso de sites de comércio eletrônico cresceu consideravelmente nos últimos anos devido à facilidade na hora da compra do produto e da praticidade com que se dá sua forma de pagamento. Entretanto um grande problema vem ocorrendo em algumas lojas virtuais, principalmente as internacionais, que são as falhas de pesquisa de determinados itens. Geralmente são identificados termos com significados confusos, que muitas das vezes não estão relacionados ao objeto da compra. A grande demanda de produtos indexados nestes sites também fez com que os seus termos perdessem mais ainda sua relevância, visto que existem diversas categorias e subcategorias que podem ou não estar conectadas entre si conforme suas características.

A representação e a recuperação da informação estão diretamente relacionadas ao uso de sites de comércio eletrônico no contexto de mecanismos de busca utilizados para obter um resultado satisfatório de acordo com a relação semântica dos termos pesquisados. Baseado nessa estrutura compreende-se que o usuário que deseja realizar uma compra de um produto tenha facilidade no momento da navegação, além de um resultado de pesquisa satisfatório quando utiliza apenas a intuição. Sendo assim, saber qual o melhor critério para utilização de termos mais específicos e corretos facilitará o entendimento do usuário e potencial comprador, uma vez que será identificado o problema, e proporcionará uma visão diferenciada de recuperação da informação para lojas virtuais.

## 1.3 OBJETIVOS

Seguem abaixo os objetivos da pesquisa:

### 1.3.1 Objetivo geral

Analisar a representação da informação em sites de comércio eletrônico visando compreender a recuperação da informação e propor critérios.

### 1.3.2 Objetivos específicos

- a) verificar os métodos utilizados para representação da informação em lojas virtuais;
- b) observar a percepção do usuário em relação aos resultados obtidos na pesquisa;
- c) propor critérios para viabilizar melhorias na recuperação da informação;

## 2 REFERENCIAL TEÓRICO

O referencial teórico baseia-se em conceitos da Recuperação da Informação e de Comércio Eletrônico. Os conceitos que são utilizados neste trabalho: Comércio Eletrônico, Representação da Informação, Organização da Informação, Organização do Conhecimento, Recuperação da Informação.

### 2.1 Comércio eletrônico

O comércio eletrônico, termo abreviado da língua inglesa conhecido por e-commerce, (*eletronic commerce*), baseia-se em compras feitas em lojas virtuais através de meios eletrônicos. Com o crescimento da procura por diversos produtos em sites de compras devido a rapidez e facilidade no atendimento ao público, o comércio eletrônico tornou-se uma opção prática para o consumidor da era digital que muitas vezes não consegue encontrar tempo para frequentar lojas físicas.

O comércio eletrônico surgiu em meados dos anos 1990 com a popularização da internet e ao crescimento desse modelo de mercado quando começaram as primeiras demandas dentro dessa área. Mais precisamente, segundo Giovanini e Britto (2012 p.4), “[...] o comércio eletrônico surgiu em março de 1991, por meio da Acceptable Use Police, que são regras do governo americano”. Tais regras passaram a permitir o tráfego comercial na internet, marcando o início de uma nova era de consumo. CATALANI et al. (2014, p. 14) afirma que:

A disponibilidade de padrões abertos de domínio público e a proliferação de computadores pessoais, modems e redes locais vividas ao longo da década de 1980 compuseram o cenário adequado para que a internet tivesse, ao longo dos anos 1990, um crescimento meteórico, atingindo rapidamente milhões de pessoas, em especial depois da criação da World Wide Web, popularizada a partir de 1994. Naturalmente, tal fato despertou o interesse de empresas, e muitos começaram a se perguntar: “afinal, como ganhar dinheiro com a internet”.

E uma das suas características principais envolve a quebra de barreiras geográficas, sendo uma ferramenta facilitadora e que envolve a troca de informações entre diversos países.



Fuoco (2003, p. 15) apud Cavalcante e Brascher (2014, p.194) alega que “[...] o comércio eletrônico vem do termo em inglês e-commerce e está relacionado a qualquer tipo de tentativa para realizar negociações dos produtos de modo eletrônico”. A autora também apresenta outro termo em inglês chamado de e-business e trata de outra abreviação que significa electronic business e abrange esse determinado tipo de negócio no mundo da web, ou seja:

O e-business envolve o uso do sítio como um dos meios do relacionamento com clientes, o vasto uso de recursos de software que interliguem as operações dentro da empresa e fora dela, com seus fornecedores e clientes. Também envolve o uso do e-mail como instrumento de marketing direcionado, a integração dos sistemas internos da companhia com os canais de venda, estoques e fornecedores. Tudo gira em tempo real, dando agilidade e eficiência à empresa. É a economia corporativa se adaptando ao mundo eletrônico. (FUOCO, 2003, p. 15, apud CAVALCANTE e BRASCHER, 2014, p.194)

Assim, percebe-se que o relacionamento vendedor e cliente é muito importante dentro do comércio eletrônico, visto que a comunicação é um dos fatores principais para um bom entendimento dos negócios feitos. Para que isso seja mais organizado foram criados modelos que serão definidos abaixo.

A classificação do modelo de negócios via comércio eletrônico está dividida em seis conceitos diferentes que estão baseadas por Catalani et al (2004, p.36-37):

- a) **B2C** - Business to consumer: é o modelo mais comumente utilizado e consiste na forma de varejo online onde é possível empresas venderem seus produtos para os consumidores finais da internet. Exemplo: lojas virtuais como a Americanas.com;
- b) **B2B** – Business to business: negociação entre as empresas, de diversas maneiras, como troca de informações por e-mails para facilitar os negócios. Exemplo: e-marketplaces, Mercado Eletrônico, que é uma empresa brasileira;
- c) **B2E** – Business to employee: empresas que utilizam a intranet para oferecer produtos aos seus funcionários com menores preços. Exemplo: PREVI criado pelo Banco do Brasil;
- d) **B2G** – Business to government: quando a empresa vende para o governo. Exemplo: Mastermaq Software;

- e) **C2C** - Consumer to consumer: negociação direta entre consumidores. Exemplo: No Brasil temos o Mercado Livre, que é o site mais famoso do ramo, o Enjoei e o OLX;
- f) **C2B** – Consumer to business: pouco utilizado no Brasil, neste modelo o consumidor que oferta o produto para a empresa. Exemplo: We do Logos e Freela;

No Brasil o comércio eletrônico cresceu muito nos últimos anos pelo aumento do número de brasileiros com acesso à internet, que está diretamente relacionado ao grande número de pessoas compradoras de smartphones no país, que ultrapassa 35% do total. Mais da metade da população possui internet em suas residências com um percentual de 20% de consumidores virtuais ativos. De acordo com dados do site e-bit<sup>2</sup>, conhecido por ser uma das principais fontes de informação e inteligência competitiva que fornece ao consumidor qual a reputação e a melhor loja para realizar suas compras, em 2016 o comércio eletrônico apresentou um faturamento de R\$ 44,4 bilhões, com um crescimento de 7,4% em relação ao de 2015.

No âmbito internacional, segundo Cateora (2007, p. 425-426) apud Mendonça (2016 p. 242) “[...] o e-commerce é mais desenvolvido nos Estados Unidos por conta do vasto número de pessoas que possuem computadores pessoais e pelo baixo custo da Internet, mas é na Europa que se tem o maior consumo”. Enquanto isso na Ásia, a região Sudeste concentra o maior crescimento de mercado, com uma velocidade superior a outras regiões, somando mais de 600 milhões de consumidores que vai do Vietnã até a Indonésia, cujo principal interesse são aparelhos tecnológicos. Contudo, algumas empresas como a Unilever, que é de bens de consumo, e a japonesa Shiseido, de cosméticos, afirmam que o boom do comércio eletrônico está relacionado ao fato de que pode se aprofundar no mercado, não dependendo apenas de redes de varejo físicas de difícil acesso.

---

<sup>2</sup> Site da Nielsen Company que mede a reputação de sites de comércio eletrônico. <https://www.ebit.com.br/>

## 2.2 Representação da Informação

A Representação da Informação está presente no mundo desde as primeiras pinturas que representavam a caça aos animais entre povos de tribos antigas, até os dias atuais com imagens em aparelhos celulares ou computadores. Alvarenga (2003) [...] afirma que a pintura foi um elemento usado para representar a humanidade, e que representar significa colocar “algo em lugar de”. Lima e Alvarez (2012) [...] definem representação como o ato de utilizar elementos simbólicos (palavras, figuras, desenhos, etc.) para substituir uma ideia ou fato. Existem diferentes forma de representação da informação, como por exemplo, a de imagens, como citado, por termos dentro de um dicionário, entre outras.

A palavra representar é um verbo transitivo que significa revelar, mostrar, expor (por meio do papel), enquanto que representação é um substantivo feminino e significa exposição verbal ou escrita do que existe dentro da mente.

A representação da informação de acordo com Furgeri (2006) [...] é um conjunto de elementos básicos e de regras para a conjunção desses elementos, sendo esse um processo semelhante que ocorre na Ciência da Computação, onde os computadores digitais representam uma informação qualquer através de um conjunto de elementos (bits) e regras específicas (agrupamento de bits para formar caracteres).

Baseado em Galvão (2001, p.3) “[...] a representação da informação pode tanto explicitar os dados necessários para a identificação dos produtores de informações (autores) e do contexto de produção da informação (local e data), como também explicitar, de maneira concisa, os conteúdos informacionais presentes em um texto ou um conjunto de textos”. Portanto pode-se dizer que a explicitação dos conteúdos é realizada por dois tipos de representação: índices e resumos.

Galvão (2001, p.3) “[...] afirma que os índices associam termos, palavras e/ou descritores aos vários conteúdos informacionais, que podem estar dispersos em vários textos, permitindo sua identificação, seleção e recuperação das informações mais relevantes para os usuários”. Enquanto que os resumos condensam todo o conteúdo informacional de textos originais, mas preservam suas estruturas e vocabulários.

Na Ciência da Informação, existem dois tipos de resumos: o resumo informativo e o resumo indicativo. Macedo e Tálamo (1994) apud Galvão (2001, p.3) “[...] apontam que os resumos informativos são os mais convenientes pois apresentam, de forma concisa e objetiva, o conteúdo informacional relevante do texto, podendo substituí-lo. Enquanto que os resumos

indicativos, segundo as autoras, atêm-se apenas a alguns tópicos do texto, servindo apenas para tomada de decisão sobre a leitura e não o texto original”.

Muitos fenômenos podem interferir na busca por informações, como neologismos, sinonímias, polissemias, hiponímias e metáforas, além das dificuldades de tradução de linguagem específica para a natural e vice-versa, nas construções dos textos. Os exemplos de ferramentas desenvolvidas para representação da informação são as classificações e o que será abordado nesse texto, as taxonomias e como se aplicam.

A palavra Taxonomia, que é de origem grega, de acordo com Currás (2010, p. 58) apud Cavalcante (2016, p.31):

A taxonomia, etimologicamente, se deriva do grego: *taxis* = ordenação e *nomia* = lei, norma, regra. Encontram-se também na biografia as expressões “taxinomia” e “taxeonomia”, sem dúvidas derivadas das distintas formas de terminação que a expressão original grega apresentasse nos diferentes casos de sua declinação. (CURRÁS, 2010, p. 58 apud CAVALCANTE, 2016, p. 31).

Com o advento da Web, os avanços da tecnologia na década de 1990 fizeram com que as informações fossem organizadas baseado em uma estrutura, no caso com a utilização das taxonomias, que funcionam como uma estrutura que pode organizar todos os dados contidos em um site de comércio eletrônico, com produtos que possuem diversas características. A partir desse ponto, é importante destacar que a ideia de uma taxonomia é a classificação e a hierarquização.

Compreende-se que taxonomias são estruturas hierarquizadas, por meio de ferramentas que trabalham na classificação, e segundo Gomes, Motta e Campos (2006) apud Cavalcante (2012, p. 33):

Taxonomia é, por definição, classificação sistemática. Ali as classes se apresentam segundo uma ordem lógica, apoiada em princípios. Por outro lado, a taxonomia é mais restrita em suas possibilidades de exploração por conter apenas relações hierárquicas. (GOMES, MOTTA e CAMPOS, 2006 apud CAVALCANTE, 2012, p.33).

Após compreensão da definição da taxonomia, e de sua estruturação, é importante destacar que existem dois conceitos que denominam dois tipos diferentes de taxonomias, que atuam na recuperação da informação de maneiras diferentes: a taxonomia navegacional e a taxonomia descritiva. Conforme Cavalcante (2016, p. 40) “[...] a taxonomia navegacional está relacionada a recuperação da informação por meio da navegabilidade de seus descritores

(*browsing*), enquanto que a taxonomia descritiva serve de suporte aos mecanismos de busca. (*searching*).” Nesse caso, as taxonomias navegacionais não são visíveis aos usuários pois estão inseridas nos elementos gráficos dos sites.

### **2.3 Organização da Informação e Organização do Conhecimento**

Com o passar dos anos, a utilização da internet fez crescer consideravelmente o número de informações recebidas para seus usuários, que adquirem por meios de mecanismos de buscas todo o tipo de informação de que necessitam. Com esse volume informacional de fácil acesso, a organização de conteúdo disponível na Web tornou-se algo mais complexo. Um exemplo disso são os sites de comércio eletrônico, que muitas das vezes possuem um grande número de termos relacionados a diferentes tipos de produto, e devido a quantidade infinita de itens para serem colocados em catálogos, surge essa problemática de como deve ser feita a organização da informação.

Para uma maior compreensão da Organização da Informação e Organização do Conhecimento, serão definidos os conceitos de dado, informação e conhecimento, sendo estes conceitos utilizados para exemplificação.

O significado do termo dados está relacionado a informação que ainda não foi tratada, ou seja, a informação ainda no início de seu desenvolvimento. Na tecnologia da informação, de acordo com Lima, (2012, p. 24):

[...] dado é a menor partícula da informação que, quando isolado, não possibilita a decodificação clara e comum ao entendimento humano, pois está fora do contexto que favorece a sua compreensão. Informação é o conjunto de dados que permite extrair algum significado, podendo favorecer a obtenção de conhecimento. (LIMA, 2012, p. 24).

Baseado na citação acima acerca do que é um dado, compreende-se que um dado não pode definir por si só o significado de alguma coisa, ele precisa estar dentro de um grupo com vários dados diferentes, para que assim uma informação seja criada, e então ele poderá ter algum significado, podendo passar algum conhecimento. Entretanto, o conceito de conhecimento está atrelado aos aspectos cognitivos da mente humana, com processos mentais de assimilação, constituindo assim um saber. Wilson (2006, p.38) define conhecimento como:

[...] aquilo que sabemos. Envolve processos mentais de compreensão, entendimento e aprendizado que passam na mente, e apenas na mente, independentemente de interação com o mundo exterior à mente e a interação dos outros. (WILSON, 2006, p.38).

Segundo (CAVALCANTE, 2012) “[...] a organização da informação está diretamente relacionada ao processo de recuperação da informação, onde a mesma pode vir a ser registrada em documentos físicos ou digitais, utilizando-se de mecanismos de representação da informação contidos nestes tipos de suportes”.

Baseado na Ciência da Informação, de acordo com (LIMA, 2012) é uma área das ciências que está voltada para organização, representação e recuperação da informação. Ela atua como um modelo para responder a perguntas que se refere a representação da informação, representação do conhecimento, se as áreas do conhecimento estão sendo representadas do mesmo modo, e se tudo pode ser representado. Dahlberg define organização do conhecimento de outra forma:

A ciência que estrutura e organiza sistematicamente unidades do conhecimento (conceitos) segundo seus elementos de conhecimento (características) inerentes e a aplicação desses conceitos e classes de conceitos ordenados a objetos/assuntos. (DAHLBERG, 2006, p.12).

Segundo (Cavalcante, 2012), [...] o conhecimento é considerado como um conjunto estruturado de conceitos disponível na mente. Por outro lado, entende a informação como a expressão física do conhecimento para fins comunicativos, referindo-se a linguagem, em suas diversas formas, como principal vetor. Sendo assim, a linguagem verbal não pode ser negligenciada pela Ciência da Informação, visto que ela está voltada para informação escrita.

Em sites de comércio eletrônico, o modo como estão disponibilizados os conceitos podem influenciar diretamente na hora da recuperação da informação, sendo que a organização do conhecimento está atrelada ao fato de saber reconhecer um domínio, identificar as facetas, agrupar os conceitos e fazer uma relação entre eles. Assim SOWA (1984, p.344) destaca que “[...] conceitos e percepções não podem formar modelos perfeitos de mundo, são abstrações que selecionam características importantes para determinada finalidade e que ignoram detalhes ou complexidades que são importantes apenas para determinados objetivos”.

Desse modo, CAFÉ e BRASCHER (2011) “[...] explicam que uma mesma entidade pode ser caracterizada de diversas maneiras diferentes, de acordo com o domínio em questão e o tipo de sistema de organização do conhecimento que deseja construir, visto que arbitrariedade da seleção de agrupamentos não deve ser feita de modo aleatório”. É preciso

destacar quais os aspectos que se deseja utilizar em uma determinada representação do conhecimento.

E ainda dentro do raciocínio das autoras em questão, compreende-se que a organização do conhecimento toma por base a utilização de um ou mais domínios e que seus componentes (termos), deverão refletir modelo de comunicação de um tipo de comunidade, podendo acontecer de ao descrever uma realidade de maneira parcial, acabar resultando em baixos índices na recuperação da informação. Isso ocorre devido ao fato de a informação, que está articulada entre o termo e o conceito, pode ter seu significado variado de acordo com o contexto que ele está inserido.

No contexto da Organização do Conhecimento, existem ferramentas utilizadas que apresentam a interpretação estruturada e organizada do objeto, chamadas de Sistema de Organização do Conhecimento (SOC). Os SOC são definidos, de acordo com Carlan e Medeiros (2011), como [...] esquemas de representação do conhecimento, ou seja, são instrumentos que fazem a tradução dos conteúdos e dos documentos originais e completos, dentro de um esquema estruturado sistematicamente, representando tal conteúdo, com a finalidade de organizar a informação e o conhecimento, facilitando a sua recuperação da informação contida no documento. É necessário também fazer uma análise do que o usuário necessita, qual tipo de SOC apropriada, qual o tipo de hardware ou software que será mais adequado àquela rede, além de integração e manutenção do mesmo. Portanto, os SOC são novas linguagens documentárias que agregam elementos que estão incorporados nas inovações tecnológicas da era digital.

Dentre os sistemas de organização do conhecimento existentes, os mais tradicionais são segundo Sales e Café (2009, p. 102):

- a) **Tesouro:** é uma lista estruturada de termos, e caracteriza-se pela especificidade e pela complexidade existente no relacionamento entre os termos que comunicam o conhecimento especializado;
- b) **Taxonomia:** esquemas de classificação de assunto, de grupos ou seres individuais, associando-os a uma classe, que pode associar-se à uma subclasse, economizando tempo de pesquisa. Pode ser descritiva, navegacional ou de gerenciamento de dados;
- c) **Vocabulário controlado:** instrumento de controle terminológico que estabelece a forma de representar os termos que compõem um conjunto de área do conhecimento, tornando possível maior coerência entre os termos indexados;

- d) **Sistemas de classificação:** criados com o objetivo de organizar acervos de biblioteca para facilitar a pesquisa feita pelo usuário. A CDD (Classificação Decimal de Dewey) foi um dos primeiros sistemas de classificação criado em 1876 pelo bibliotecário Melvill Dewey e que deu origem a outros sistemas.

Ainda dentro do contexto dos SOC, as ontologias vêm ganhando um papel muito importante em diversas comunidades científicas, como a Ciência da Computação, por exemplo. Na Organização da Informação, a ontologia possui sentido especial, menos tradicional, isto é, ela é “um catálogo de tipos de coisas” onde se supõe que exista um domínio, baseado na perspectiva de um indivíduo que usa uma determinada linguagem. GRUBER (1993) define ontologia como:

Uma especificação explícita de uma conceituação. (...) em tal ontologia, definições associam os nomes de entidades no universo do discurso (ex., classes, relações, funções, ou outros objetos) com texto legível para humanos, descrevendo o que os nomes significam e axiomas formais que restringem a interpretação e o uso bem formado destes termos. (GRUBER 1993, p. 199)

As ontologias são mais que linguagens documentárias, permitindo que máquinas possam processar o “raciocínio” automatizado por meio de regras e inferências (SALES, CAMPOS e GOMES, 2008, p. 63). Portanto, CARLAN e MEDEIROS (2011) afirmam que [...] ontologias são o elemento da web semântica que possibilita o nível de representação semântico, e por meio delas, os softwares utilizados na web semântica podem utilizar o conhecimento codificado para compreender e interpretar semanticamente, os documentos e os objetos. Assim, compreende-se que as ontologias são uma SOC de extrema importância em sites de comércio eletrônico, devido ao fato de trabalharem a compreensão e a interpretação dos significados que um item possui na descrição de uma loja, tornando sua busca mais rápida e fácil.



## 2.4 Recuperação da Informação

Os sistemas de recuperação da informação, segundo CESARINO (1985), podem ser compreendidos como “[...] um conjunto de operações consecutivas executadas para localizar, dentro da totalidade das informações disponíveis, aquelas que são realmente relevantes. Executam as funções de seleção, análise, indexação, e busca das informações”. Em todas estas etapas a interação entre usuário e sistema é fundamental, mesmo ocorrendo falhas algumas vezes.

Araújo (1985) afirma que o “[...] marco moderno da recuperação de informação e da consolidação dos sistemas de recuperação da informação como entidade, é em geral, datado das décadas de 40 e 50, embora o termo recuperação da informação, derivado do inglês “information retrieval”, apenas tenha sido criado em 1951 por Calvin Mooers”. Este marco é caracterizado pela necessidade de armazenar e dar acesso rápido e preciso ao grande número de documentos que vinha tendo um crescimento exponencial desde o século XVII, além do surgimento do computador, que era visto como a grande solução para os problemas de armazenamento e recuperação da informação.

Nesse sentido, Saracevic (2006) aponta que “[...] a influência da recuperação da informação na primeira década de 1960 teria surgido da emergente necessidade de armazenamento e da recuperação em uma sociedade que cada vez mais se especializava”. Conforme Araújo (2012) aborda, [...] os Sistemas de Recuperação da Informação devem representar, armazenar, organizar e localizar itens de informação. O autor ainda destaca que a indexação é a principal função de SRI, e que seus componentes incluem documentos, necessidades dos usuários, consulta formulada, e o próprio processo de recuperação em si.

Existem dois modelos de SRI, que são os clássicos (booleano, vetorial e probabilístico) e os mais avançados (lógica fuzzy, redes neurais e algoritmos genéticos). De acordo com Monteiro et. al (2017): “[...] nos modelos clássicos de RI o texto é modelado puramente como uma sequência de palavras, sendo assim, outros modelos foram desenvolvidos visando a ampliação da performance e da operacionabilidade: baseados na teoria dos conjuntos, a lógica difusa ou nebulosa (fuzzy), o booleano estendido e o baseado em conjuntos; baseado nos modelos algébricos, vetorial generalizado, indexação semântica latente e redes neurais; e enfim, baseado nos modelos probabilísticos, BM25, modelos de linguagem, divergência de aleatoriedade e redes bayesianas”.

Aqui serão explicados detalhadamente os modelos de Recuperação da Informação, que se dividem em modelos clássicos e modelos estruturados, onde os modelos clássicos são, de

acordo com Souza (2006, p. 166) “[...] documentos descritos por um conjunto de palavras-chave representativas, também chamadas de termos de indexação, que buscam representar o assunto do documento, e sumarizar seu conteúdo de forma significativa”.

Os três modelos clássicos são o Booleano, Vetorial e Probabilístico, abaixo exemplificados por Souza (2006, p. 166):

- a) **Booleano:** baseia-se na teoria dos conjuntos, e para cada questão, são recuperados todos os documentos que o usuário buscou por meio da combinação de termos. Opera por um modelo binário, tratando de relevância e não relevância, não gera ordenação dos resultados de busca, e não apresenta ranqueamento. Exemplo: bases de dados científicas;
- b) **Vetorial:** o modelo mais utilizado em sites de comércio eletrônico por utilizar o método de ranking, no qual os documentos são organizados em um “saco de palavras” (bag of words), sendo representados como vetores de n-dimensional, onde n são a quantidade de termos índices, ou seja, palavras dentro de todos os documentos no sistema. Atribui peso ao termo, qual é o mais importante, e consegue ranquear e dizer sua importância. Não possui um modelo binário, calcula-se o grau de similaridade satisfeitos por um documento para serem considerados relevantes.
- c) **Probabilístico:** nesse modelo a interação com o usuário é a principal técnica utilizada, por meio do feedback de tudo aquilo que é pesquisado por ele, ou seja, o conjunto de documentos que forem considerados pertinentes ao que o usuário deseja recuperar, serão analisados e encaminhados para refinar os resultados de busca.

Seguindo o contexto do projeto de pesquisa, que trabalha a recuperação em sites de comércio eletrônico, Monteiro et al (2017) afirma que, “[...] quando se fala sobre objetos de multimídia, ocorrem algumas particularidades em suas formas de representação, que possibilitam a recuperação de imagens apenas pela comparação de imagens relacionadas, sem necessitar escrever uma consulta, já que o ranqueamento é feito de forma diferente ou mesmo é inexistente”. Um exemplo disso é o site de compras internacionais do grupo Alibaba, chamado Aliexpress, onde podem ser feitas buscas pelos itens da loja por meio da utilização de imagens. O usuário pode fazer a pesquisa por meio do aplicativo para smartphones, onde o mesmo autoriza o site a se conectar com sua câmera ou seu álbum de fotos do aparelho, e a partir disso, o sistema de recuperação do aplicativo faz um apanhado de todos os produtos que

podem vir a ser o produto desejado, ou que venham a ter alguma semelhança física com ele. Para Baeza-Yates e Ribeiro Neto (2013), “[...] dentre as estratégias de recuperação de multimídia, a recuperação feita por imagens é a mais simples por ser estática”.

Os mecanismos de buscas de acordo com Miranda (2005) “[...] são complexas ferramentas da internet, que são utilizadas pelo público em geral, e oferecem ampla gama de serviços online, como por exemplo: buscar informações, armazená-las e apresenta-las aos usuários, e-mails e outras formas de comunicação, além de envolver entretenimento e propaganda”. Para Monteiro (2015) “[...] os mecanismos de busca geram os índices contemporâneos, refletem agenciamentos híbridos e podem ser definidos como uma enorme base de dados de informações importantes a respeito de sites na web”.

Existem três tipos de mecanismos de buscas que são os mais importantes, apontados por Caldeira (2015):

- a) ***crawling***: um programa denominado crawler varre a web coletando páginas novas e atualizadas para serem incluídas no índice;
- b) ***indexing***: as informações recolhidas são armazenadas e indexadas na base de dados;
- c) ***searching***: uma interface de busca é exibida para o usuário realizar a pesquisa, a interface e o software relacionado que conecta a busca do usuário com o índice executa um algoritmo para encontrar e exibir as páginas relevantes;

É interessante destacar que esses três tipos de mecanismos de buscas podem ocorrer em sites de comércio eletrônico, pois o *searching* por exemplo é um mecanismo que trabalha a relação do usuário com a loja do qual ele faz suas pesquisas, utilizando a relevância a partir de termos mais buscados pelo mesmo.

O conceito de relevância segundo Moreira et al. (2017, p.168):

Tem como correlato o significado, em especial no Google, mas também na Web Semântica, tanto que surge uma inclinação para chamar esses Sistemas Digitais de Recuperação da Informação de “Sistemas de Significação” em razão das camadas Semântica e Semiótica em seus processos de indexação, geração de índices e interpretação das *queries* e dos resultados. (MOREIRA et al., 2017, p.168).

Portanto a relevância pode ser um conceito a espera de novas atualizações conceituais, cognitivas e técnicas, estando fortemente relacionada às tecnologias semânticas de ciberespaço, à personalização e a contextualização da pesquisa.

### 3 METODOLOGIA

Para execução deste trabalho, foi utilizada a pesquisa de caráter exploratório, tendo por objetivo proporcionar maior familiaridade com o problema, visto que se pretende torná-lo mais explícito. De acordo com Gil (2008, p.27) “[...] as pesquisas exploratórias têm como principal finalidade desenvolver, esclarecer e modificar conceitos e ideias, tendo em vista a formulação de problemas mais precisos ou hipóteses pesquisáveis para estudos posteriores”. Também foi feita uma pesquisa descritiva, que segundo Gil (2008, p.28) “[...] têm como objetivo primordial a descrição das características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis”. Sua principal característica seria a utilização de técnicas padronizadas de coletas de dados.

A pesquisa fez uma abordagem qualitativa, com o objetivo de fazer uma coleta de dados, por meio do estudo de particularidades e experiências individuais. Conforme Gil (1999, p.43) “[...] o objetivo fundamental da pesquisa é descobrir respostas para problemas mediante o emprego de procedimentos metodológicos”. Dito isto, optou-se por utilizar como abordagem metodológica o estudo de caso, que segundo Cervo e Bervian (2002) apud Machado (2005) “[...] como uma pesquisa sobre um determinado indivíduo, família ou grupo ou comunidade que seja representativo do seu universo, para examinar aspectos variados de sua vida”.

#### 3.1 CAMPO EMPÍRICO

Esse trabalho aborda como campo empírico duas plataformas: Mercado Livre e Aliexpress, por serem dois sites de comércio eletrônico com características semelhantes de lojas de departamento. O Mercado Livre foi fundado em 1999 pelo empresário argentino Marcos Galperín e está presente em 12 países da América Latina, e também em Portugal na Europa. O plano de negócios foi escrito por Galperín ainda quando ele trabalhava para obter seu diploma de MBA na Escola de Negócios de Stanford, e montou um grupo de pessoas para formar a equipe de profissionais com o intuito de colocar o site em prática. O Aliexpress é um site de comércio eletrônico internacional pertencente ao grupo Alibaba, que foi fundado pelo chinês Jack Ma e um grupo de 17 amigos em 1999, com o intuito de conectar vendedores e potenciais clientes em uma plataforma global que pudesse atender pequenas e médias

empresas. No entanto, o Aliexpress foi fundado alguns anos depois, mais precisamente em 2010. Ele fornece produtos para compradores online internacionais e é composto por pequenas empresas.

A seguir, foi criado um quadro 1 com dois sites de comércio eletrônico, nacional e internacional, que possuem características em comum relacionados a representação da informação e recuperação da informação. No quadro estão dispostas as ferramentas que estão ou não relacionadas com a estrutura de um desses dois sites, e caso seja confirmada essa relação, serão feitas marcações no quadro para correlacionar as ferramentas aos sites.

QUADRO 1: Sites de comércio eletrônico e suas ferramentas

<b>SITES</b>	<b>ALIEXPRESS</b>	<b>MERCADO LIVRE</b>
Ranqueamento	x	x
Tagueamento	x	x
Indexação Automática	x	
Barras de Navegação	x	x
Menus Pull-Down / Drop Down	x	x
Tabela de conteúdos	x	x

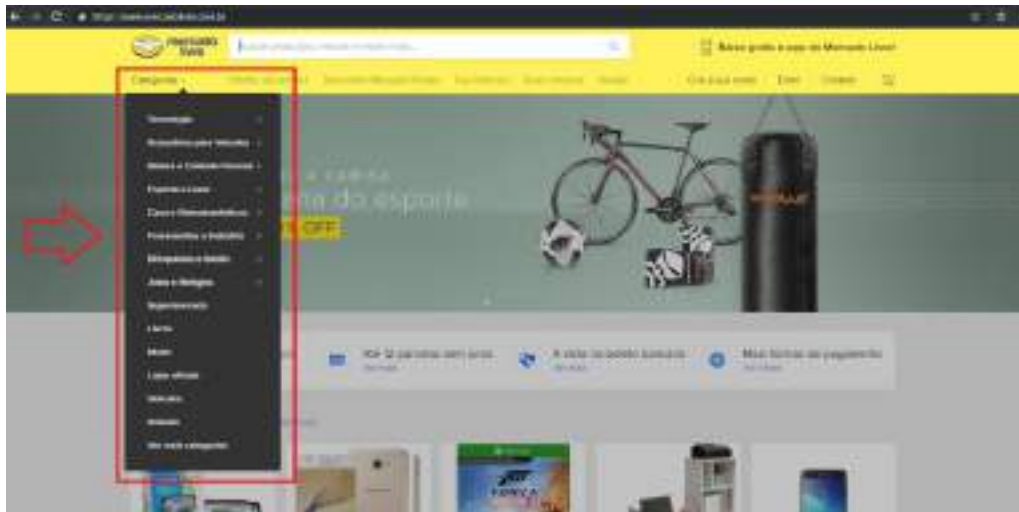
Fonte: a autora.

Existem alguns mecanismos de busca que podem ser utilizados na construção de um site comercial, facilitando as buscas do usuário e possível comprador, e agilizando a recuperação da informação. Foi feita uma análise diante dos exemplos que estão representados no quadro, onde os dois sites utilizados na pesquisa foram acessados para o reconhecimento inicial de cada ferramenta.

O primeiro site a ser observado foi o Aliexpress, que possui a ferramenta de barra de navegação, muito comumente utilizada em sites de comércio eletrônico. A barra de navegação geralmente encontra-se no topo do site, em destaque para que o usuário possa percebê-la de imediato e consiga fazer sua pesquisa de modo mais rápido. Funciona como uma Taxonomia, que tem como objetivo o estudo científico responsável por determinar e classificar sistematicamente diferentes coisas em categorias. A Taxonomia, explicada pela Ciência da Informação, é segundo Cavalcante e Brascher (2014, p.3) apud Gomes et al (2006, online) “[...] uma classificação sistemática, sendo mais restrita em suas possibilidades de exploração



FIGURA 2: BARRA DE NAVEGAÇÃO MERCADO LIVRE



Fonte: Mercado Livre.

Para compreender o uso de tags é importante destacar o significado de folksonomia, pois são conceitos relacionados a representação do conteúdo. A Folksonomia foi um termo cunhado em 2004 pelo arquiteto Thomas Vander Wall, sendo considerado um modo de indexar a informação, e faz analogia a taxonomia tendo como sua principal característica a criação de tags (palavras-chave). Ainda Vander Wall (2005) definiu que [...] “folksonomia é um modo de as pessoas etiquetarem objetos, uma classificação social, fazendo com que as pessoas que usem o mesmo vocabulário sejam capazes de encontrar o mesmo objeto”. Segundo Rapetti (2007, p.18) “As tags (etiquetas) são metadados, as palavras-chaves que são utilizadas para classificar o conteúdo da informação, porém com o diferencial de serem incluídas pelos usuários dessa informação”. Ainda, de acordo com Maculan et al. (2009) [...] “o ranqueamento é uma técnica utilizada para saber quais as tags mais utilizadas pelos usuários, para melhor compreensão da regularidade semântica nas descrições criadas por eles”.

FIGURA 3: EXEMPLOS DE TAXONOMIA NO SITE ALIEXPRESS



Fonte: Aliexpress.

FIGURA 4: EXEMPLOS DE TAXONOMIA NO SITE MERCADO LIVRE



Fonte: Mercado Livre.

Os menus drop-down ou menus pul-down são um mecanismo de busca que de certa forma está diretamente relacionado às barras de navegação na estrutura dos sites de comércio eletrônico, pois são um método prático de mostrar uma vasta lista com diversas opções de escolhas de termos desde que uma única opção seja exibida até o usuário clicar em uma categoria e ativar as subcategorias relacionadas na lista.

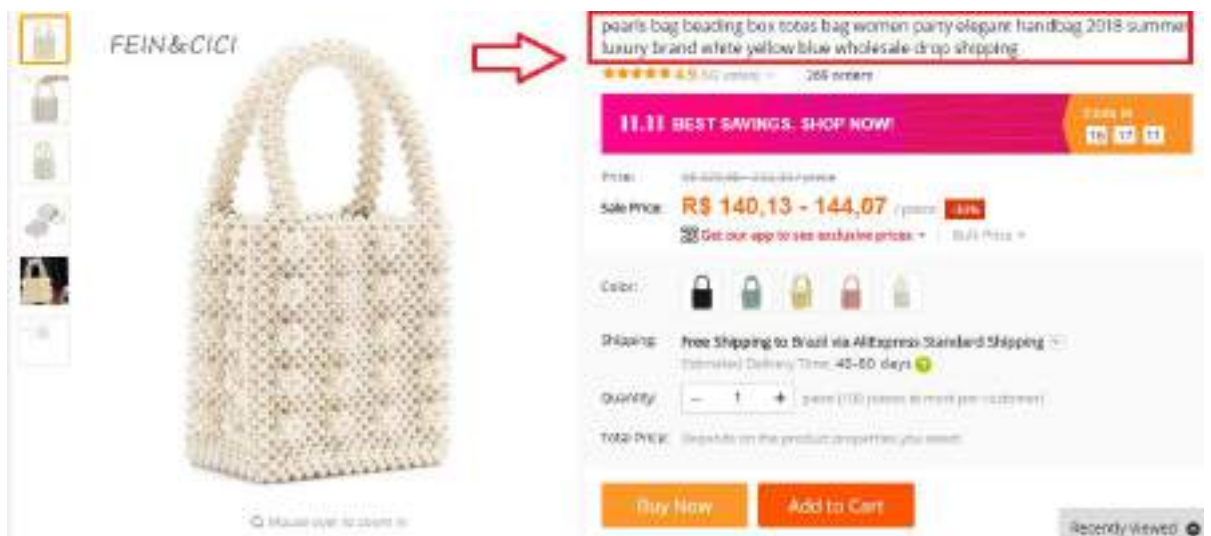
Compreende-se que alguns sites são indexados pelos seus buscadores, sendo estes, sites que fazem uma pesquisa de conteúdo especializada, como por exemplo, o Google, o Bing, etc. Isso facilita o usuário a encontrar o site que estava procurando e até outros que não



estava. Assim, ocorre uma organização dos resultados de busca nos sites de acordo com sua relevância, criando-se uma lista.

Dentro de sites de comércio eletrônico, a utilização de termos variados por meio do usuário pode acarretar em falhas na hora de recuperar um determinado, pois uma palavra pode estar relacionada a uma característica específica do produto, ou como pode vir a ser um termo muito abrangente. Abaixo um exemplo da variedade de termos utilizados para definir uma bolsa:

FIGURA 5: TERMOS PARA DEFINIR O ITEM BOLSA NO ALIEXPRESS



Fonte: Aliexpress.

FIGURA 6: TERMOS PARA DEFINIR O ITEM BOLSA NO MERCADO LIVRE



Fonte: Mercado Livre.

No caso de sites como o Aliexpress e Mercado Livre, por se tratar de uma rede colaborativa de troca de informação, o controle desse vocabulário é responsabilidade do vendedor, e conforme a sua concepção ele vai estabelecer quais palavras-chave estarão mais de acordo com o item que será pesquisado. Entretanto, não é possível afirmar se realmente um controle por meio desses vendedores pois muitas vezes são indexados termos em excesso para que isso facilite a busca do usuário/comprador da loja, podendo causar alguns erros, como excesso de resultados não satisfatórios.

### 3.2 TÉCNICAS DE COLETA E ANÁLISE DE DADOS

Como instrumento de coleta de dados foi utilizado o Questionário. De acordo com Gil (2008, p.128) o questionário é “a técnica de investigação composta por um número mais ou menos elevados de questões apresentadas por escrito às pessoas, tendo por objetivo o conhecimento de opiniões, crenças, interesses, expectativas, situações vivenciadas, etc”.

Além disso, foi utilizada a Observação participante, que segundo o Infopedia “é uma técnica de investigação social, em observador partilha, na medida em que as circunstâncias o permitam, as atividades, as ocasiões, os interesses e os afetos de um grupo de pessoas ou de uma comunidade”. A observação participante foi feita pela autora por meio de pesquisas em ambos os sites para ter uma análise mais clara de como a recuperação da informação e da representação da informação estão aplicadas.

### 3.3 POPULAÇÃO

Foi utilizado o Questionário com um determinado número de pessoas de dois grupos do Whatsapp. Um grupo é composto por alunos de Biblioteconomia da UFRJ, contabilizando 13 pessoas, e o outro grupo também é formado por alunos da UFRJ de Biblioteconomia, totalizando 45 membros. Os dois grupos do Whatsapp contabilizam ao todo 58 pessoas, sendo estes amigos em comum que serão ouvidos para pesquisa e obtenção de respostas sobre a utilização dos dois sites.

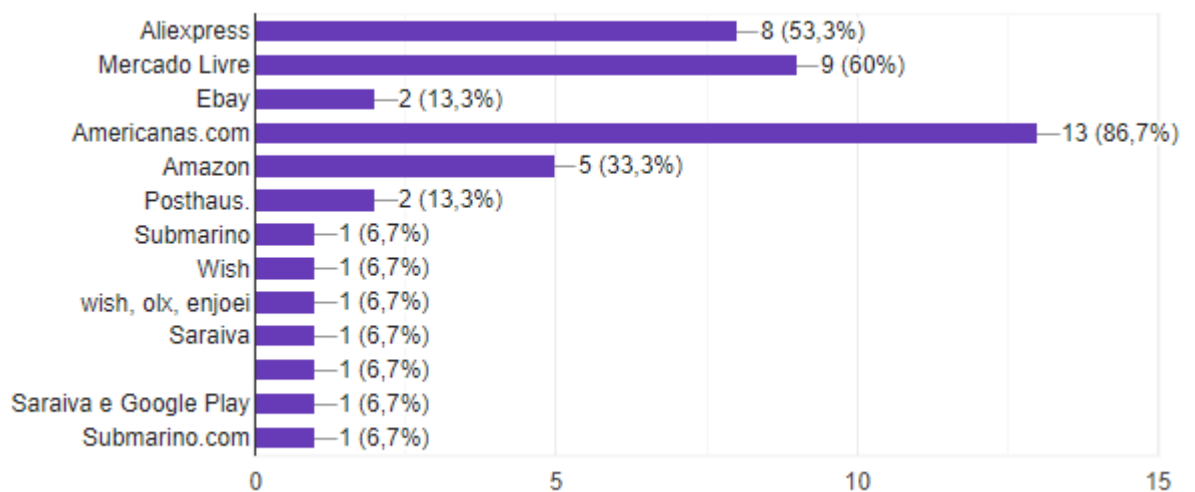
Na escolha dos perfis foram utilizados critérios enfatizando que os participantes possuem afinidade com o uso de sites de comércio eletrônico, para que as questões fossem melhor compreendidas e assim obtivesse resultados satisfatórios. Foram escolhas aleatórias, sem a necessidade de restrição de idade, gênero ou local.

#### 4 ANÁLISE DE DADOS

Nesta seção foram analisados os dados obtidos nessa pesquisa, mediante a observação participante e o uso de questionário. Com base nos dados demonstrados abaixo podemos compreender a representação da informação e a recuperação da informação em sites de comércio eletrônico, tanto nacionais quanto internacionais.

O questionário foi enviado para 58 membros de dois grupos do Whatsapp distintos que costumam utilizar com frequência sites de comércio eletrônico, e foram obtidas respostas oriundas de 26% dos indivíduos da pesquisa. Abaixo temos o gráfico resultante da primeira questão abordada, sobre quais sites de comércio eletrônico são utilizados com mais frequência pelos usuários.

**Gráfico 1 – Sites de comércio eletrônico mais utilizados**



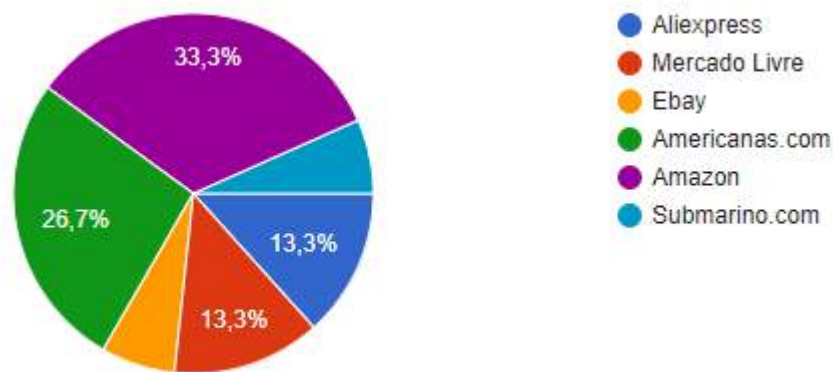
Fonte: a autora.

Conforme podemos observar com o gráfico acima, o site da loja Americanas.com é o mais utilizado pelos usuários, seguidos do Mercado Livre e do Aliexpress. Os três sites possuem algumas características em comum como grande variedade de produtos vendidos, que estão divididos em diversos setores como casa e banho, roupas, acessórios, eletrodomésticos e aparelhos eletrônicos.

A partir disso, na segunda pergunta os usuários foram questionados em quais sites eles encontram com mais facilidade um produto, e novamente os sites que tiveram mais destaque

foram o Aliexpress, Mercado Livre e Americanas.com, sendo que o site da Amazon também apareceu nas preferências dos respondentes. É importante destacar que a estrutura destes sites são muito semelhantes e geralmente os resultados de busca podem estar de acordo com o perfil do usuário e também referente a quantidade de termos utilizados na pesquisa.

**Gráfico 2** – Sites com termos mais precisos



Fonte: a autora.

No gráfico 2 podemos observar que o site da Amazon ganha mais destaque em relação a precisão, com 33,3% dos votos, seguido do site Americanas.com com 26,7%. Na pergunta posterior a essa foi questionado aos usuários quais sites descrevem melhor os produtos que são vendidos e o porquê, e a grande maioria destacou estes mesmos sites descritos, explicando que a grande variedade de informações oferecidas sobre o produto, com uma quantidade maior de termos em sua descrição, além da facilidade com o idioma, que mesmo em sites internacionais, pode vir a ser traduzido pela loja de acordo com a localização de endereço de seu usuário.

No entanto, ao serem questionados sobre a quantidade de termos utilizados pelos sites para representação dos itens, ocorreu uma unanimidade nas respostas já que todos os usuários escolheram a opção com mais características, porque assim eles podem ter certeza que o que estão comprando é exatamente aquilo que os termos em comum descrevem. Quando questionados sobre quais são as preferências na busca por um produto, com o uso de imagens ou palavras, 11 usuários deram preferência as palavras por aumentarem a precisão da busca, pela praticidade, além de descrever proporções, o que não é possível no uso de imagens. Assim, apenas 2 usuários escolheram imagem como modo mais prático de busca, e um

usuário destacou que em ambos os casos gosta de ter as duas opções disponíveis, dependendo do que se está pesquisando. É fato que utilizar palavras colabora para que se tenha resultados precisos e satisfatórios em um site de comércio eletrônico, que muitas vezes contêm uma sobrecarga de informações de seus produtos. A utilização de imagens em sites como o Aliexpress por exemplo, pode ser feita via aplicativo através do uso de sua câmera ou álbum de fotos do celular, mas devido ao grande número de termos relacionados as imagens, muitas vezes características gerais como cores do produto podem resultar itens que não pertencem a categoria desejada. Sendo assim, confirma-se por este questionário que as palavras são mais eficazes que as imagens na recuperação da informação.

Como foram escolhidos sites diversos para abordagem no questionário, os usuários também foram indagados sobre existir alguma dificuldade na utilização de sites internacionais. Seis usuários afirmaram que isso dificultava sim devido a barreira linguística, que é o maior impedimento, além de muitas vezes os termos recuperados não possuem nenhuma relação com o produto, e por fim um destacou os preços mais altos, que não influencia nesta pesquisa em específico. Os 9 restantes que disseram que não atrapalhava o site ser internacional destacaram a facilidade de manipular a plataforma, e que pelo idioma inglês ser padrão estavam acostumados a fazer buscas.

Os usuários também foram questionados sobre qual site era mais fácil e rápido encontrar um item pesquisado, nacional ou internacional. A maioria deu preferência aos sites nacionais, contabilizando 11 respondentes, devido ao costume e a praticidade de manuseio da plataforma, além do idioma também trazer uma familiaridade e rapidez nos resultados de busca. Os 4 que escolheram sites internacionais destacaram que a ampla gama de palavras-chaves e as opções de pesquisas variadas os tornam mais interessantes, e também promovem uma busca de modo intuitivo. Nesse caso, o maior empecilho na recuperação da informação seria o idioma, que muitas vezes pode atrapalhar nas buscas para quem não o domina.

Por fim, no questionário é perguntado quais os problemas que o usuário encontra ao tentar recuperar um produto em site de comércio eletrônico. O uso de palavras-chaves nem sempre encontram um objeto, a dificuldade de especificar, ou até restringir a pesquisa ao produto de interesse, a quantidade de resultados que não condizem com o termo original e a representação de produtos similares foram as respostas mais significativas.

Constata-se através desses dados que a maioria dos usuários de sites de comércio eletrônico deram maior preferência aos produtos que possuem mais termos descrevendo as características dos produtos pesquisados, refutando a ideia inicial da pesquisa de que isso seria um empecilho na recuperação da informação desejada. Além disso, o fato de o site ser

nacional foi destacado devido ao idioma português ser um ponto positivo nas buscas. Contudo, é interessante perceber a partir da observação participante, que uma grande variedade de termos não é o suficiente para que ocorra uma boa representação da informação e uma recuperação satisfatória, pois de acordo com o questionário e nas pesquisas feitas a partir dele, a precisão dos termos é crucial para que ocorra um resultado de busca satisfatório.

## 5 CONSIDERAÇÕES FINAIS

No desenvolvimento dessa pesquisa, buscou-se analisar a representação da informação em sites de comércio eletrônico visando compreender a recuperação da informação. Os objetivos da pesquisa foram verificar os métodos utilizados para representação da informação em lojas virtuais, analisar a percepção do usuário em relação aos resultados obtidos na pesquisa e propor um conjunto de critérios para viabilizar melhorias na recuperação da informação.

Na pesquisa foram examinados como o usuário de sites de comércio eletrônico faz suas buscas por itens nas lojas de departamento mais conhecidas, quais as maiores dificuldades encontradas na hora da recuperação da informação, e como é feita a representação dos produtos em sites nacionais e internacionais, por meio da análise de alguns tipos ferramentas que funcionam como mecanismos de buscas, dando foco especialmente para o Aliexpress e o Mercado Livre, estudados na metodologia.

É de grande importância abordar este tema devido ao fato que hoje em dia é muito comum o acesso as lojas de departamentos por meio de sites de comércio eletrônico, graças a rapidez com que é feita uma compra, além da grande variedade de produtos encontrados em um só lugar. Os sites de comércio eletrônico apresentam uma enorme diversificação de itens em seus catálogos, fazendo com que o consumidor, como a aluna que escreveu esta pesquisa por exemplo, tenha dificuldades em encontrar com rapidez um produto.

No âmbito acadêmico, os conceitos de Recuperação da Informação e de Representação da Informação ganham estão diretamente relacionados ao tema proposto, visto que é importante observar se os mecanismos de buscas estão dando resultados satisfatórios em concordância com a relação semântica estabelecida entre os termos pesquisados. Contudo, os usuários desses sites desejam que suas buscas resultem em uma recuperação das informações dos itens positiva, assim como os profissionais que trabalham com as vendas nesses sites desejam que seus produtos sejam encontrados por meio da representação da informação correta.

Os resultados obtidos na análise de dados apontam que a maioria dos usuários de sites de comércio dão preferência aos itens que possuem mais termos para definição das características principais dos itens representados, porém frisando que tais termos devem ser específicos para não tornar a busca superficial. O idioma em língua portuguesa também foi colocado como um dos pontos positivos devido à familiaridade com o idioma de seu país de



origem, pois muitas vezes em sites internacionais alguns termos traduzidos ficam confusos e não resultam em uma pesquisa que esteja de acordo com a informação que o usuário deseja recuperar. Por fim, é apontado que a pesquisa feita por meio do uso de palavras é preferível ao uso de imagens, já que muitas vezes uma imagem é uma representação da informação pouco específica do item, enquanto que as palavras promovem uma descrição muito mais específica das características.

Baseado nos objetivos específicos, concluiu-se que foram feitas as devidas análises da representação da informação visando uma maior compreensão da recuperação da informação, os métodos utilizados para representação da informação foram devidamente verificados por meio da análise das ferramentas utilizadas nos sites de comércio eletrônico Mercado Livre e Aliexpress, por meio do uso de um quadro fazendo um *checklist* e apresentando imagens das lojas. A percepção do usuário foi analisada por meio do uso de questionário, com observação participante, e a partir de suas respostas foram levados em consideração alguns pontos feitos durante a pesquisa. Assim, são propostos alguns critérios para viabilizar melhorias na recuperação da informação como a utilização de termos corretos representados por palavras em sites de comércio eletrônico, o uso de uma grande quantidade de termos mais específicos, e que se dê preferência ao idioma de domínio de seu usuário.

## REFERÊNCIAS

ALIEXPRESS. Disponível em: < <https://www.aliexpress.com/>>. Acesso em: 11 dez. 2018.

ALMEIDA, Mauricio Barcellos; BAX, Marcello Peixoto. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, v.32, n.3, p. 7-20, set./dez. 2003. Disponível em: <<http://revista.ibict.br/ciinf/article/view/984/1024>>. Acesso em: 1 jul. 2018

ARAVINDAN, A.; SETBOONSARNG, C. **Empresas usam dados para aproveitar boom de e-commerce na Ásia**. Exame, 23 de outubro de 2017. Disponível em: <<https://exame.abril.com.br/negocios/empresas-usam-dados-para-aproveitar-boom-de-e-commerce-na-asia/>>. Acesso em: 6 de jun. 2018.

ARLAN, Eliana; MEDEIROS, Marisa Brascher Basílio. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, [S.l.], v. 4, n. 2, p. 55-73, fev. 2012. ISSN 1983-5213. Disponível em: <<http://periodicos.unb.br/index.php/RICI/article/view/6209/5102>>. Acesso em: 30 jun. 2018.

B2W DIGITAL. Visão geral do setor. **Comércio eletrônico no Brasil**. 2018. Disponível em: <<https://ri.b2w.digital/institucional/comercio-eletronico-no-brasil>>. Acesso em: 6 jun. 2018.

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Recuperação de informação: conceitos e tecnologia das máquinas de busca**. 2. ed. Porto Alegre: Bookman, 2013.

BARBALHO, A. B. C. **Contratos internacionais de comércio eletrônico: características e regulamentação**. Conteúdo Jurídico, Brasília: 16 mar. 2017. Disponível em: <<http://www.conteudojuridico.com.br/artigo,contratos-internacionais-de-comercio-eletronico-caracteristicas-e-regulamentacao,588687.html>>. Acesso em: 6 jun. 2018.

CAFÉ, Lígia; BRÄSCHER, Marisa. Organização do Conhecimento: teorias semânticas como base para estudo e representação de conceitos. **Informação & Informação**, [S.l.], v. 16, n. 2, p. 25-51, dez. 2011. ISSN 1981-8920. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/10388/9282>> Acesso em 29 jun. 2018.

CALDAS, W. F.; MOREIRA, M. P. Folksonomia e classificação de etiquetas: estudo de caso Flickr. In: **ENANCIB, IBICT**, 10.ed., João Pessoa, PB. Out. 2009. Disponível em: <<http://enancib.ibict.br/index.php/enancib/xenancib/paper/viewFile/3164/2290>>. Acesso em: 30 nov. 2018.

CALDEIRA, Fátima Hassan. O mecanismo de busca do Google e a relevância na relação sistema-usuário. **Revista Digital do Programa de Pós-Graduação em Letras da PUCRS**, v.8, n.1, p.91-106, janeiro-junho, 2015. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/letronica/article/view/19616/13418>>. Acesso em 1 jul. 2018.

CAVALCANTE, R. S. **Crítérios para avaliação de taxonomias navegacionais em sítios de comércio eletrônico**. 2012. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Ciência da Informação, Universidade de Brasília, 2012. Disponível em: <[http://repositorio.unb.br/bitstream/10482/10917/1/2012\\_RaphaeldaSilvaCavalcante.pdf](http://repositorio.unb.br/bitstream/10482/10917/1/2012_RaphaeldaSilvaCavalcante.pdf)> Acesso em 18 abr. 2018.

CATALANI, L. et al. **E-commerce**. – 2.ed. – Rio de Janeiro: FGV, 2006. 168p.

CATIVO, J. Definição: folksonomia. **Biblioteconomia Digital**. Informação, tecnologia e conhecimento, jan. 2011. Disponível em: <<https://www.biblioteconomiadigital.com.br/2011/01/definicao-folksonomia.html>>. Acesso em: 30 nov. 2018.

CHINA LINK TRADING. **Jack Ma**: a história do fundador da Alibaba. 2018. Disponível em: <<http://www.chinalinktrading.com/blog/jack-ma-historia/>>. Acesso em 4 jul. 2018

CHIZHIKOVA, M. M. 3 Principais fatores de ranqueamento do Google. **E-commerce Brasil**, dez. 2016. Disponível em: <<https://www.ecommercebrasil.com.br/artigos/3-principais-fatores-de-ranqueamento-do-google/>>. Acesso em: 30 nov. 2018.

DESTINO NEGÓCIO. **Mercado Livre lidera operações de comércio eletrônico na América Latina. De Stanford à Nasdaq**: a história do Mercado Livre. Disponível em: <<https://destinonegocio.com/br/casos-de-sucesso/mercado-livre-lidera-operacoes-de-comercio-eletronico-na-america-latina/>>. Acesso em 4 de jul. 2018.

EBIT. **Reputação e certificação**. 2018. Disponível em: <<https://www.ebit.com.br/entenda-a-ebit>>. Acesso em: 6 de jun. 2018.

EUGÊNIO, Márcio. Brasil, 2017. **Crescimento do comércio eletrônico no Brasil**. Disponível em: <<https://ecommercenews.com.br/artigos/tendencias-artigos/crescimento-do-comercio-eletronico-no-brasil/>>. Acesso em: 6 de jun. 2018.

FINEP. Biblioteca FINEP. **Empresas e serviços – Vocabulário controlado**. 2018. Disponível em: <<http://www.finep.gov.br/biblioteca/biblioteca-da-finep/produtos-e-servicos/biblioteca-vocabulario?view=bibliotecavocabulario&layout=bibliotecavocabulario>>. Acesso em 1 jul. 2018.

FUOCO, T. **Guia do Valor econômico de comércio eletrônico**. São Paulo: O Globo, 2003.

GERHARDT, Tatiana. SILVEIRA, Denise. **Métodos de pesquisa**. Porto Alegre: Editora da UFRGS, 2009. Estudo de Caso, p. 39.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. – 5ed. – São Paulo: Atlas, 1999.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. – 6ed. – São Paulo: Atlas, 2008.

GIOVANINI, T.; BRITO, R. R. E-commerce como ferramenta de venda através de sites de compra coletiva. In: Simpósio de Excelência em Gestão e Tecnologia, 9.ed., Resende, RJ.

AEDB: 2012. Disponível em: <<https://www.aedb.br/seget/arquivos/artigos12/43816448.pdf>>. Acesso em: 30 nov. 2018.

GRUBER, T. R. A translation approach to portable ontologies. **Knowledge Acquisition**, v.5, n.2, p.199-220, 1993.

HOLANDA, C.; BRAZ, M. I. Indexação automática de conteúdos na web: análise de sites de museus. **Biblionline**, v.8, n.1, p. 42-59, 2012. Disponível em: <<http://www.periodicos.ufpb.br/ojs/index.php/biblio/article/view/11766/7509>>. Acesso em: 30 nov. 2018.

MACEDO, N. D.; TÁLAMO, M. F. M. **Resumos:** subsídios para sua elaboração. *Cadernos de Análise Documentária*, v.1, n.1, p.1-12, maio 1994.

MACHADO, W. T. O estudo de caso como método de pesquisa científica. **Portal da Classe Contábil**, 2005. Disponível em: <<https://www.classecontabil.com.br/o-estudo-de-caso-como-metodo-de-pesquisa-cientifica/>>. Acesso em: 29 de nov. 2018.

MACULAN, B. C. M. de S.; ASSIS, J. de; ALVES, A. V.; PEREIRA, F. Taxonomia, folksonomia, acessibilidade e usabilidade: proposta de interseção na área de organização do conhecimento, com foco na recuperação de informação. 2009. In: **Seminário em Ciência da Informação**, Londrina, 27 a 29 de setembro de 2009. Disponível em: <<http://eprints.rclis.org/23854/>>. Acesso em: 30 nov. 2018.

MENDONÇA, H. G. de. E-commerce. **Revista Inovação, Projetos e Pesquisa**, v.4, n.2, p. 240-251, jul./dez. 2016.

MERCADO LIVRE. Disponível em: <<https://www.mercadolivre.com.br/>>. Acesso em: 11 de dez. 2018.

MONTEIRO, Silvana Drumond et al. Sistemas de recuperação da informação e o conceito de relevância nos mecanismos de busca: semântica e significação. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 22, n. 50, p. 161-175, set. 2017. ISSN 1518-2924. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n50p161/34700>  
Acesso em 29 jun. 2018.

MONTEIRO, S. D. Knowledge Graph e a significação: novos agenciamentos semióticos dos índices contemporâneos. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 8, n. 2, p. 203-218, jul/dez. 2015.

NOVELLINO, Maria Salet Ferreira. Instrumentos e metodologias de representação da informação. **Informação & Informação**, [S.l.], v. 1, n. 2, p. 37-45, dez. 1996. ISSN 1981-8920. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/1603/1358>>. Acesso em 18 jun. 2018.

Observação participante in Artigos de apoio Infopédia [em linha]. Porto: Porto Editora, 2003-2018. Disponível em: < [https://www.infopedia.pt/\\$observacao-participante](https://www.infopedia.pt/$observacao-participante)>. Acesso em 4 de jul. 2018.

**Organização e representação do conhecimento na perspectiva da Ciência da Informação** / Georgete Medleg Rodrigues e Ilza Leite Lopes (org.). Brasília: Thesaurus, 2003. 271 p.

PORTAL DO BIBLIOTECÁRIO. **Tesauro e ontologia: uma introdução**. 2015. Disponível em: <<http://portaldobibliotecario.com/biblioteconomia/tesauro-e-ontologia-uma-introducao/>>. Acesso em 29 de jun. 2018.

PROFESSOR DIGITAL. **Qual a diferença entre dados, informação e conhecimento?** Disponível em: <<https://www.luis.blog.br/qual-a-diferenca-entre-dados-informacao-e-conhecimento/>>. Acesso em: 29 jun. 2018.

RODRIGUES, B. Os 4 Is do conteúdo para e-commerce. **E-commerce Brasil**, fev. 2017. Disponível em: <<https://www.ecommercebrasil.com.br/artigos/os-4-is-do-conteudo-para-e-commerce/>>. Acesso em: 30 nov. 2018.

SALES, R. de; CAFÉ, L. Diferenças entre tesauros e ontologias. **Perspectivas em Ciência da Informação**, v.14, n.1, p. 99-116, jan./abr. 2009. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-99362009000100008](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362009000100008)>. Acesso em: 29 out. 2018.

SIGNIFICADOS. Significado de taxonomia. O que é taxonomia? Jul. 2016. Disponível em: <<https://www.significados.com.br/taxonomia/>>. Acesso em: 30 nov. 2018.

SOUZA, R. R. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. **Perspectivas em Ciência da Informação**, v.11, n.2, p.161-173, mai./ago. 2006. Disponível em: < <http://www.scielo.br/pdf/pci/v11n2/v11n2a02>>. Acesso em: 30 nov. 2018.

## APÊNDICE A – QUESTIONÁRIO

### PESQUISA SOBRE SITES DE COMÉRCIO ELETRÔNICO

A pesquisa analisa a representação da informação em sites de comércio eletrônico para a compreensão da recuperação da informação

1. Quais os sites de comércio eletrônico você costuma utilizar para fazer compras?
  - ( ) Aliexpress
  - ( ) Mercado Livre
  - ( ) Ebay
  - ( ) Americanas.com
  - ( ) Amazon
  - ( ) Outros
2. Em qual site de comércio eletrônico você encontra um item com mais rapidez?
3. Qual é o site que te possibilita fazer a busca com termos mais precisos?
  - ( ) Aliexpress
  - ( ) Mercado Livre
  - ( ) Ebay
  - ( ) Americanas.com
  - ( ) Amazon
  - ( ) Outros
4. Quais sites de comércio eletrônico descrevem melhor o produto? Por que?
5. Você prefere um item que possua a descrição dos produtos com mais ou menos características? Justifique sua resposta.
6. O que você prefere utilizar na busca por um item: imagem ou palavra? Por que?
7. Sente dificuldades em encontrar algum item da loja quando o site é internacional? Por que?
8. Quais problemas você encontra ao tentar recuperar um produto em site de comércio eletrônico?
9. Qual site você acha mais fácil e rápido para encontrar um item: nacional ou internacional? Por que?