

Painel de Informação sobre a COVID-19: consultas SPARQL na Wikidata

COVID-19 Information Panel: SPARQL query at Wikidata

Ana Carolina Simionato Arakaki¹, Fabiano Ferreira de Castro², Felipe Augusto Arakaki³

¹ Universidade Federal de São Carlos (UFSCar), São Paulo, Brasil. ORCID: <https://orcid.org/0000-0002-0140-9110>

² Universidade Federal de São Carlos (UFSCar), São Paulo, Brasil. ORCID: <https://orcid.org/0000-0002-8712-2654>

³ Universidade de Brasília (UnB), Distrito Federal, Brasil. ORCID: <https://orcid.org/0000-0002-3983-2563>

Autor para correspondência/Mail to: Ana Carolina Simionato Arakaki, acsimionato@ufscar.br

Recebido/Submitted: 21 de setembro de 2020; **Aceito/Approved:** 17 de novembro de 2020



Copyright © 2020 Arakaki, Castro & Arakaki. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição-NãoComercial-Compartilhado 3.0 Não Adaptada. Ao serem publicados por esta Revista, os artigos são de livre uso em ambientes educacionais, de pesquisa e não comerciais, com atribuição de autoria obrigatória. Mais informações em <http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice>.

Resumo

Introdução: Uma das formas de enfrentamento da COVID-19 diz respeito aos aspectos relacionados à produção e à disseminação de informações confiáveis, claras e de rápida compreensão. Muitas são as ações e as iniciativas comunicacionais e informacionais em prol da disseminação e dos meios que garantam a aceitabilidade, a adesão e o cumprimento das medidas de prevenção e controle da COVID-19. Objetiva o desenvolvimento de um ambiente digital, entendido aqui como um painel com tópicos relacionados à COVID-19, baseados em consultas SPARQL Protocol and RDF Query Language (SPARQL) e no dataset do Wikidata. **Método:** Utiliza-se uma metodologia de caráter teórico e aplicado, a partir da Revisão Sistemática da Literatura para subsidiar a construção do *corpus* conceitual subjacente às tecnologias computacionais advindas da Web Semântica e *Linked Data* e sua aplicação na estruturação e da modelagem do ambiente, para disponibilização e compartilhamento dos dados científicos. **Resultados:** Os dados coletados na Revisão Sistemática da Literatura revelam pouca produção científica disponível no âmbito internacional, entretanto, iniciativas interessantes que já se preocupam com a abertura e a disponibilização dos dados científicos na Web. Além disso, o painel de informações sobre a COVID-19 desenvolvido está categorizado em seis principais eixos, como Mapa COVID-19, Sintomas da COVID-19, Possíveis tratamentos, Taxonomia, Trabalhos relacionados e Imagens relacionadas. **Conclusão:** O painel de informações sobre a COVID-19 se apresenta como um ambiente digital que potencializa a visualização, o acesso e o compartilhamento dos dados e das informações para usuários heterogêneos, contribuindo com o repasse de informações consistentes, estruturadas e confiáveis, bem como o fomento de diretrizes públicas para o controle da disseminação da doença.

Palavras-chave: COVID-19; Web semântica; *Linked data*; Wikidata; SPARQL.

Abstract

Introduction: One of the ways of coping with COVID-19 concerns aspects related to the production and dissemination of reliable, clear and quickly understood information. There are many communicational and informational actions and initiatives in favor of dissemination and of the means that guarantee the acceptability, adherence and compliance with the prevention and control measures of COVID-19. This research aims to develop a digital environment, understood here as a panel with topics related to COVID-19, based on SPARQL Protocol and RDF Query Language (SPARQL) queries and on the Wikidata dataset. **Method:** To do so, a theoretical and applied methodology is used, based on the Systematic Literature Review to support the construction of the conceptual corpus underlying the computational technologies from the Semantic Web and *Linked Data* and its application in the structuring and modeling of the environment, for making scientific data available and sharing. **Results:** The data collected in the Systematic Literature Review reveal little scientific production available at the international level, however, interesting initiatives are already concerned with the openness and availability of scientific data on the Web. In addition, the information panel on COVID-19 developed is categorized into six main axes, such as Map COVID-19, Symptoms of COVID-19, Possible treatments, Taxonomy, Related works and Related images. **Conclusion:** Thus, the information panel about COVID-19 presents itself as a digital environment that enhances the visualization, access and sharing of data and information for heterogeneous users, contributing to the transfer of consistent, structured and reliable information, as well as the promotion of public guidelines for controlling the spread of the disease.

Keywords: COVID-19; Semantic web; *Linked data*; Wikidata; SPARQL.

INTRODUÇÃO

A COVID-19 é uma doença causada pelo coronavírus, denominado SARS-CoV-2, que pode causar diversos problemas respiratórios semelhantes à gripe e sintomas como tosse, febre e, em casos mais graves, dificuldade para respirar. O surto foi encontrado pela primeira vez em Wuhan, na China, em dezembro de 2019 e foi chamado de pandemia pela Organização Mundial da Saúde (OMS) em 11 de março de 2020.

No cenário contemporâneo, a disponibilização dos dados científicos é fundamental para encontrar soluções para evitar a propagação do vírus, ou encontrar soluções para combater outras doenças já existentes. Nesse contexto, os dados sobre os principais sintomas, imagens, resultados de vacinas, quantidade de pessoas infectadas e de mortos, entre outros, são essenciais para a criação de estratégias de prevenção e de combate à doença, além de conscientização da população e precisam estar disponibilizados em acesso aberto e estruturados.

Ressalta-se ainda a importância dos dados para a sociedade, visto que auxiliam no direito constitucional de acesso à informação, bem como o fomento de diretrizes públicas para o controle da disseminação da doença.

Em razão da pandemia e, principalmente, pela necessidade do acesso, do compartilhamento e do uso de dados e da informação disponível para os sujeitos que apropriar-se-ão dos conteúdos que circulam nos ambientes digitais, este artigo busca apresentar um painel, entendido nesta pesquisa como um ambiente informacional digital, com tópicos relacionados à COVID-19, baseados em consultas *SPARQL Protocol and RDF Query Language* (SPARQL) e no *dataset* do *Wikidata*.

Para tanto, utiliza-se uma metodologia de natureza teórica e aplicada, pautada na Revisão Sistemática da Literatura (RSL), que busca a identificação do embasamento conceitual e aplicado, referente às iniciativas do uso das ferramentas tecnológicas da *Web Semântica* e *Linked data* e no desenvolvimento de um ambiente informacional digital, denominado nesta pesquisa como painel de informação sobre a COVID-19.

Este texto está estruturado em cinco seções. Na primeira seção, é apresentado o referencial teórico sobre **Linked Data**, *Web Semântica*, *Resource Description Framework* (RDF) e SPARQL. Na terceira seção, são apresentados os procedimentos metodológicos, os resultados e as discussões. Por fim, são expostas as considerações finais.

REFERENCIAL TEÓRICO

A disponibilização de dados científicos em acesso aberto pode trazer benefícios tanto para comunidade acadêmica, quanto para sociedade. Por um lado, cientistas podem utilizar e reutilizar esses dados para comprovar ou refutar hipóteses e criar teses, promovendo avanços para Ciência. Já para a sociedade, esses dados precisam ser tratados e disponibilizados de forma simples e de fácil compreensão.

A partir desses princípios, um grupo de trabalho formado pelo *World Wide Web Consortium* (W3C) elaborou um documento para auxiliar a estruturação e a publicação de dados na Web, denominado de boas práticas de publicação de dados, conforme apresentado no Quadro 1).

De acordo com [Penteado, Bittencourt, e Isotani \(2019, p. 179\)](#), essas práticas objetivam fazer com que os dados sejam encontrados e compreendidos tanto por humanos quanto por computadores, facilitando a interação entre produtores e consumidores dos dados, oportunizando dados precisos, regularmente atualizados e disponíveis o tempo todo. Seu objetivo é fomentar a expansão continuada da *Web* como meio para a troca de dados e, com isso, tirar proveito máximo das suas capacidades, como a habilidade de ligar um fato a outro, descobrir recursos relacionados e criar visualizações interativas.

(continua)

| Boas Práticas para publicação de dados | | |
|----------------------------------------|------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| Categoria | Boa prática | Benefício |
| Metadados | Boa Prática 1: Fornecer metadados | Reúso, Compreensão, Facilidade de descoberta e Facilidade de processamento |
| | Boa Prática 2: Fornecer metadados descritivos | Reúso, Compreensão e Facilidade de descoberta |
| | Boa Prática 3: Fornecer metadados estruturais | Reúso, Compreensão e Facilidade de processamento |
| Licença de Dados | Boa Prática 4: Fornecer informações sobre a licença de dados | Reúso e Confiança |
| Procedência de Dados | Boa Prática 5: Fornecer informações de procedência dos dados | Reúso, Compreensão e Confiança |
| Qualidade de Dados | Boa Prática 6: Fornecer informações de qualidade de dados | Reúso e Confiança |
| Versionamento de Dados | Boa Prática 7: Fornecer indicador de versão | Reúso e Confiança |
| | Boa Prática 8: Fornecer o histórico de versão | Reúso e Confiança |
| Identificadores de Dados | Boa Prática 9: Usar URIs persistentes como identificadores de conjuntos de dados | Reúso, Capacidade de conexão e Interoperabilidade |
| | Boa Prática 10: Usar URIs persistentes como identificadores dentro de conjuntos de dados | Reúso, Capacidade de conexão, Facilidade de descoberta e Interoperabilidade |
| | Boa Prática 11: Atribuir URIs para as versões dos conjuntos de dados e séries | Reúso, Facilidade de descoberta e Confiança |

(continuação)

| Categoria | Boa prática | Benefício |
|-------------------------|------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|
| Formatos de Dados | Boa Prática 12: Usar formatos de dados padronizados legíveis por máquinas | Reúso e Facilidade de processamento |
| | Boa Prática 13: Usar representações de dados que sejam independentes de localidade (local e neutral) | Reúso e Compreensão |
| | Boa Prática 14: Fornecer dados em formatos múltiplos | Reúso e Facilidade de processamento |
| Vocabulários de Dados | Boa Prática 15: Reutilizar vocabulários, dando preferência aos padronizados. | Reúso, Facilidade de processamento, Compreensão, Confiança e Interoperabilidade |
| | Boa Prática 16: Escolher o nível de formalização adequado | Reúso, Compreensão e Interoperabilidade |
| Acesso a Dados | Boa Prática 17: Fornecer download em massa (<i>bulk download</i>) | Reúso e Facilidade de acesso |
| | Boa Prática 18: Fornecer subconjuntos para conjuntos de dados extensos | Reúso, Capacidade de conexão, Facilidade de acesso e Facilidade de processamento |
| | Boa Prática 19: Usar negociação de conteúdo para disponibilizar dados em formatos múltiplos | Reúso e Facilidade de acesso |
| | Boa Prática 20: Fornecer acesso em tempo real | Reúso e Facilidade de acesso |
| | Boa Prática 21: Fornecer dados atualizados | Reúso e Facilidade de acesso |
| | Boa Prática 22: Fornecer uma explicação para os dados que não estão disponíveis | Reúso e Confiança |
| | Boa Prática 23: Disponibilizar dados por meio de uma API | Reúso, Facilidade de processamento, Interoperabilidade e Facilidade de acesso |
| | Boa Prática 24: Usar padrões <i>Web</i> como base para construção de APIs | Reúso, Capacidade de conexão, Facilidade de descoberta, Facilidade de acesso, Facilidade de processamento |
| | Boa Prática 25: Fornecer documentação completa para as APIs | Reúso e Confiança |
| | Boa Prática 26: Evitar alterações que afetem o funcionamento de sua API | Confiança e Interoperabilidade |
| Preservação de Dados | Boa Prática 27: Preservar identificadores | Reúso e Confiança |
| | Boa Prática 28: Avaliar a cobertura do conjunto de dados | Reúso e Confiança |
| <i>Feedback</i> | Boa Prática 29: Coletar <i>feedback</i> de consumidores de dados | Reúso, Compreensão e Confiança |
| | Boa Prática 30: Compartilhar o <i>feedback</i> disponível | Reúso e Confiança |
| Enriquecimento de Dados | Boa Prática 31: Enriquecer dados por meio da geração de novos dados | Reúso, Compreensão, Confiança e Facilidade de processamento |
| | Boa Prática 32: Fornecer visualizações complementares | Reúso, Compreensão, Facilidade de acesso e Confiança |
| Republicação de Dados | Boa Prática 33: Fornecer <i>feedback</i> para o publicador original | Reúso, Interoperabilidade e Confiança |
| | Boa Prática 34: Obedecer aos termos de licença | Reúso e Confiança |
| | Boa Prática 35: Citar a publicação original do conjunto de dados | Reúso, Facilidade de descoberta e Confiança |

Quadro 1. Boas Práticas para publicação de dados.

Fonte: Baseado em W3C (2017).

A estruturação da disponibilização de dados apresenta diversos benefícios como: uso e reúso, facilidade de processamento, interoperabilidade, facilidade de acesso, facilidade de descoberta, confiança, facilidade de conexão e compreensão. É observado pelo Quadro 1), que a maioria das orientações para as boas práticas se apresentam como benefício para o reúso de dados. Esse fator nos leva a compreensão de que, no contexto da pandemia e da realidade instaurada pelo distanciamento social, o uso estratégico dos dados abertos e disponíveis na *Web* pode favorecer o compartilhamento e o acesso remoto aos dados, no atendimento às comunidades de interesse.

O uso e a aplicação das boas práticas atreladas às tecnologias da *Web Semântica*, em especial do *Linked Data*, possibilitaram a otimização dos processos de uso, reúso, compartilhamento, acesso, descoberta, confiança, e interoperabilidade dos dados. De acordo com Hyvönen (2012, p. 10, tradução nossa), “[...] em torno de 2005, as ideias sobre *Linked Data* e *Web de Dados* começaram a ganhar impulso como uma abordagem simples para *Web Semântica*, focada na publicação de grandes conjuntos de dados existentes e usando apenas ontologias RDF simples e leves.”

A concepção do *Linked Data* foi apresentada por Berners-Lee em 2006 e são quatro princípios que servem como sugestões para estruturação de dados. Esses princípios são:

- a) Use URIs como nomes para as coisas;
- b) Use HTTP URIs para que as pessoas possam procurar esses nomes;
- c) Quando alguém procurar um URI, fornecer informações úteis, usando os padrões (RDF, SPARQL);
- d) Incluir links para outros URIs, para que eles possam descobrir mais coisas.

Considera-se que a base de estruturação dos dados em *Linked Data* é o *Resource Description Framework* (RDF). O RDF é o modelo de descrição de recursos, seguindo uma estrutura de associação de triplas, em que um recurso está ligado a um valor por meio de uma propriedade.

O modelo RDF oferece a possibilidade para as comunidades de descrição de recursos definirem a semântica de seus metadados de maneira formal, isto é, definindo o significado dos elementos de metadados, conforme as suas necessidades específicas de descrição, em um modelo processável por máquinas (Ferreira & Santos, 2013, p. 21)

Essa estruturação Recurso - Propriedade - Valor, relaciona um recurso com seu(s) valor(es), além de relacionar um Recurso com outro Recurso. Essa configuração possibilita que outras linguagens computacionais como o SPARQL possam recuperar essas informações. De acordo com Santarém Segundo (2017, p. 3870) “o SPARQL é um conjunto de especificações que fornecem linguagens e protocolos para consultar e manipular o conteúdo publicado em RDF na Web”. Exemplos de uso dessas tecnologias em ambientes digitais estão a base *DBpedia* e a *Wikidata*. “Um dos projetos que demarcam o surgimento da *Web 3.0* é o *DBpedia*, que permite extrair informações estruturadas na *Wikipédia*, tornando-as informações disponíveis na *Web* de modo semântico e estruturado” (Paletta & Mucheroni, 2014, p. 84). De acordo com Campos, Campos, e Barbosa (2019), a *Wikidata* é uma base de dados aberta que pode ser lida e editada tanto por pessoas como por máquina e pertence à iniciativa da *Wikimedia Foundation* (WMF). Ela fornece um ponto de ligação para a *Wikipédia*, *Wikimedia Commons*, outras *wikis* da comunidade *Wikimedia* e outras iniciativas ao redor do mundo.

O *software* na qual está hospedada a *Wikidata* permite a anotação semântica de dados em páginas *wiki* e sua exportação para RDF [...] Tendo em vista a *Wikipédia* como fonte de dados importados, em contraste com a *DBpedia*, elo central da nuvem LOD, a *Wikidata* possui uma qualidade maior e quantidade menor desses dados importados, uma vez que possui um processo de curadoria manual desses dados com informação de proveniência, o que toma tempo (Campos et al., 2019, p. 67-68).

Considerando o cenário contemporâneo, marcado principalmente pelo uso das Tecnologias da Informação e Comunicação e das ferramentas tecnológicas disponíveis, sobremaneira pela *Web Semântica* e *Linked Data*, o aproveitamento e a utilização dessas tecnologias podem contribuir na otimização dos processos de busca, de acesso, de compartilhamento e de uso e reúso das informações no atendimento às necessidades dos usuários no contexto da COVID-19.

PROCEDIMENTOS METODOLÓGICOS

A metodologia utilizada é caracterizada como um estudo exploratório e aplicado, com natureza qualitativa. A pesquisa dividiu-se em quatro momentos: 1) levantamento sistemático da literatura científica sobre o tema; 2) inicialmente foi identificada, a partir da literatura qual a base para extração dos dados – *Wikidata* que seria utilizada e ainda; 3) quais os temas mais abrangentes que seriam abordados para recuperação na *Wikidata* sobre o Coronavírus (COVID-19); na sequência, 4) focou-se na construção e na elaboração das estratégias de consulta SPARQL na *Wikidata*.

Para possibilitar a construção de um referencial teórico e metodológico crítico e sistemático sobre o tema abordado, o estudo adota a Revisão Sistemática da Literatura (RSL) como método de pesquisa. A RSL consiste em uma revisão bibliográfica acrescida de critérios e por etapas rigorosas, que garantem a qualidade e a representatividade dos documentos obtidos. “As revisões sistemáticas são desenhadas para serem metódicas, explícitas e passíveis de reprodução”. (Sampaio & Mancini, 2007, p. 83). Ser passível de reprodução é uma das principais características de uma RSL, principalmente por permitir que os resultados localizados possam ser facilmente auditados.

Além de contribuição original, as revisões sistemáticas promovem o entendimento do chamado ‘estado da arte’ em um campo do conhecimento, bem como ajudam a identificar a melhor evidência de pesquisa disponível, como apontam Briner e Denyer (2012) e Siddaway, Wood, e Hedges (2019).

A Revisão Sistemática da Literatura foi conduzida em três etapas principais, conforme apresentada na Figura 1.



Figura 1. Principais etapas da Revisão Sistemática da Literatura.

Fonte: Jesus e Castro (2019, p. 47).

O software *StArt* (*State of the Art through Systematic Review*), desenvolvido pelo Laboratório de Pesquisa em Engenharia de Software (LAPES), do Departamento de Ciência da Computação, da Universidade Federal de São Carlos (UFSCar) foi utilizado para auxiliar em todas as etapas da RSL. O *StArt* conta com um protocolo estruturado, no qual serão registradas informações importantes como palavras-chave, critérios de inclusão e de exclusão, bases de dados consultadas e categorias de análise, conforme observado no Quadro 2).

(continua)

| Protocolo da Revisão Sistemática da Literatura | |
|------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Título | Painel de informação sobre o COVID-19 |
| Pesquisadores | Ana Carolina Simionato Arakaki, Fabiano Ferreira de Castro e Felipe Augusto Arakaki |
| Descrição | Mapeamento da produção científica disponível no âmbito internacional e nacional que abordam e analisam tecnologias oriundas da <i>Web Semântica</i> e <i>Linked data</i> sobre a divulgação das informações referentes ao COVID-19. |
| Objetivos | Identificar um referencial teórico e metodológico a respeito da utilização de ferramentas tecnológicas aplicadas em ambientes, sistemas e plataformas digitais acerca do COVID-19. |
| Questão de pesquisa | Como as ferramentas tecnológicas advindas da <i>Web Semântica</i> e <i>Linked data</i> têm sido discutidas nas agendas científicas internacional e nacionalmente na aplicação e no compartilhamento de dados sobre o COVID-19? |
| Intervenção | Estudos conceituais e práticos sobre SPARQL, <i>Resource Description Framework</i> (RDF) e <i>Linked data</i> aplicados ao cenário COVID-19 |
| Controle | Análise exploratória do tema, considerando artigos e eventos publicados sobre a temática. |
| Resultados | Iniciativas de utilização das ferramentas tecnológicas da <i>Web Semântica</i> e <i>Linked data</i> aplicadas no compartilhamento de dados sobre COVID-19 em ambientes, sistemas e plataformas digitais. |
| População | Profissionais da Ciência da Informação, da área de Saúde e sociedade em geral interessados na temática. |
| Palavras-chave e estratégia de busca | <i>DBPedia</i> and COVID-19 (com e sem aspas); SPARQL and COVID-19 (com e sem aspas); <i>Linked data</i> and COVID-19 (com e sem aspas); RDF or <i>Resource Description Framework</i> and COVID-19 (com e sem aspas). |

(continuação)

| Protocolo da Revisão Sistemática da Literatura | |
|-------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Critério de seleção das fontes de busca | Bases de dados de abrangência multidisciplinar, bases de dados voltadas para a sistematização de literatura cinzenta e bases temáticas da área da Saúde e Comunicação e Informação. |
| Idioma dos estudos | Inglês e Português. |
| Métodos para as fontes de buscas | Leitura do título e resumo dos documentos, aplicação dos critérios de inclusão e exclusão, leitura da introdução e conclusão ou ainda do documento completo nos casos em que as análises anteriores se mostrarem inconclusivas para a seleção. |
| Fontes de busca | MedLine Ovid; Web of science; Scielo; PubMed; Literatura Latino-Americana em Ciências da Saúde (LILACS); Information Science & Technology Abstracts - ISTA (EBSCO); Library and Information Science Abstracts - LISA (ProQuest); Library, Information Science & Technology Abstracts with Full Text - LISTA; Base de Dados Referencial de Artigos de Periódico em Ciência da Informação (BRAPCI); OpenGrey (www.opengrey.eu); Grey Literature Report (New York Academy of Medicine); www.greylit.org;; Portal de Periódicos CAPES. |
| Critérios de inclusão (I) e exclusão (E) | (I) Incluir discussão conceitual das ferramentas tecnológicas da <i>Web Semântica</i> e <i>Linked data</i> no contexto do COVID-19; (I) Discutir o papel das ferramentas tecnológicas da <i>Web Semântica</i> e <i>Linked data</i> no contexto do COVID-19; (I) Discutir a aplicação e casos de uso das ferramentas tecnológicas da <i>Web Semântica</i> e <i>Linked data</i> no contexto do COVID-19; (E) Não estar nos idiomas estabelecidos (PT, ENG); (E) Não estar no formato estabelecido (artigo); (E) Não abordar ou apenas mencionar a temática de interesse; (E) O documento completo não está disponível via portal de periódicos, vinculação institucional com a UFSCar ou não pode ser localizado com acesso gratuito no ambiente <i>Web</i> na íntegra. |
| Definição da tipologia dos estudos | Artigos de periódicos, sem restrição do período de publicação. |
| Avaliação da qualidade dos estudos | A avaliação da qualidade dos documentos será feita a partir da análise da metodologia dos estudos aceitos. |
| Campos de avaliação de qualidade | <ul style="list-style-type: none"> • Metodologia do Estudo; • Método de avaliação do periódico |
| Campos para extração dos estudos | <ul style="list-style-type: none"> • Conceitos apresentados • Ambiente/Sistema/ plataforma digital; • Casos de uso • Tipo de estudo (estudo empírico, estudo de revisão etc.) |
| Sumarização dos resultados | Após extração de dados, realizada seguindo o roteiro proposto no campo " <i>Information Extraction Fields</i> ", os dados foram agrupados em quadros de acordo com a categoria de análise, os resultados semelhantes serão agrupados permitindo a realização de uma análise quantitativa e também a elaboração de inferências que permitirão responder a pergunta de pesquisa. |

Quadro 2. Preenchimento do protocolo de busca no StArt.

Fonte: Elaborado pelos autores (2020).

Destaca-se que os resultados da RSL foram utilizados para respaldar a construção teórica dessa pesquisa, além de propiciar a compreensão da aplicação e da construção do cenário de informações relacionadas à COVID-19, baseado em consultas *SPARQL Protocol and RDF Query Language* (SPARQL) e no *dataset* do *Wikidata*.

Em relação ao estudo aplicado, todas as consultas apresentadas neste texto tiveram a sua última verificação no dia 17 de setembro de 2020, e os assuntos escolhidos foram: número de casos da doença no mundo, sintomas da doença, possíveis tratamentos, terminologia relacionada à COVID19, imagens de pesquisas da doença e, por fim, artigos publicados e no prelo sobre o Coronavírus.

Além disso, foram realizadas consultas no *dataset* da *Wikidata* pelo acesso aos dados, a partir do SPARQL *Endpoint*, um *web service* em que os usuários podem interagir utilizando o protocolo SPARQL para recuperar os dados. Este ambiente fornece uma interface em que o usuário insere uma consulta SPARQL, retornando na sequência os dados que atenderam às regras definidas.

Por fim, os dados extraídos foram organizados e disponibilizados para consulta em um painel de visualizações e está sendo disponibilizado pelo *site* institucional do Departamento de Ciência da Informação (DCI) e do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos (UFSCar), no endereço <https://www.dci.ufscar.br/covid-19/cenario-de-informacao>, para a comunidade acadêmica e interessados no assunto.

RESULTADOS E DISCUSSÃO

Em relação ao estudo teórico, pautado na Revisão Sistemática da Literatura, foram recuperados 271 documentos, dentre os quais 260 foram descartados por não contemplarem os critérios de inclusão do protocolo de busca e, dessa maneira, foram aceitos oito documentos que compõem o corpus desta pesquisa (Quadro 3).

(continua)

| Autores | Título | Abordagem |
|-----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Papadopoulos, Papadakis, e Litke (2020) | A Methodology for Open Information Extraction and Representation from Large Scientific Corpora: The COVID-19 Data Exploration Use Case | Os autores destacam uma proposta de uma metodologia para o processamento e transformação de representações estruturadas e não estruturadas de informações de texto em linguagem natural e de maneira automatizada pelas ferramentas computacionais, com base nas ontologias, na geração de sistemas de extração de informações abertos pautados no conjunto de dados médicos sobre o COVID-19. Além disso, uma série de tarefas de pós-processamento e enriquecimento semântico de informações foram adotados incluindo vinculação de entidades com o Unified Medical Language System e sua visualização em um gráfico de base de dados. Como resultado foi desenvolvido um <i>dataset</i> (CORD-19) baseado em uma ontologia abrangente de conceitos biomédicos. Artigo que contempla aspectos teóricos e aplicados. |
| Fischer e Jobst (2020) | Capturing the Spatial Relatedness of Long-Distance Caregiving: A Mixed-Methods Approach | Trabalho teórico e aplicado em que os autores analisam o cuidado à distância (LDC), no contexto da avaliação do futuro do cuidado ao idoso e da manutenção da saúde e do bem-estar das pessoas cuidadas e dos cuidadores à distância. Os autores analisam e operacionam de forma abrangente, a relação espacial da assistência de longa distância, combinando as ciências espaciais e as considerações da geoinformática e desenvolvendo uma abordagem de método misto interdisciplinar combinando dados primários de um único caso com a teoria dos gráficos. Também é desenvolvida uma ontologia, chamada ontologia LDC, que consiste num primeiro passo para a integração de informações de um único estudo de caso em infraestruturas de conhecimento espacial. A correspondência de expressões individuais para definições semânticas comuns e até mesmo para gráficos de conhecimento externos abre fontes de dados acessíveis adicionais e densifica informações, por exemplo, pontos de interesse, que são necessários para uma análise espacial LDC mais objetiva. Os autores concluem que não existem fontes de dados oficiais com interfaces de informação apropriadas, como GeoSPARQL. |

(continua)

| Autores | Título | Abordagem |
|---------------------------------------------------|-----------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Dutta e DeBellis (2020)) | CODO: An Ontology for Collection and Analysis of Covid-19 Data | Estudo teórico e aplicado, onde os autores desenvolveram a Ontologia COviD-19 para casos e informações do paciente (CODO) fornece um modelo para a coleta e análise de dados sobre a pandemia COVID-19. A ontologia fornece um modelo de código aberto baseado em padrões que facilita a integração de dados de fontes de dados heterogêneas. A ontologia foi projetada analisando fontes de dados COVID-19 díspares, como conjuntos de dados, literatura, serviços etc. A ontologia segue as melhores práticas para vocabulários, reutilizando conceitos de outros vocabulários importantes e usando os padrões W3C RDF, OWL, SWRL e SPARQL. A ontologia já tem um usuário independente e incorporou dados do mundo real do governo da Índia. Esse trabalho aparece como uma proposta semelhante aos objetivos dessa pesquisa. |
| Wang et al. (2020) | CORD-19: The COVID-19 Open Research Dataset | Os autores discorrem sobre o conjunto de dados de pesquisa aberto denominado COVID-19 Open Research Dataset (CORD-19), o qual se constitui de um recurso crescente de artigos científicos sobre o COVID-19 e a pesquisa histórica de coronavírus relacionada. O CORD-19 é projetado para facilitar o desenvolvimento de mineração de texto e sistemas de recuperação de informação sobre sua coleção de metadados e documentos estruturados de texto completo. Desde o seu lançamento, o CORD-19 foi baixado mais de 75 mil vezes e serviu como base para muitos sistemas de mineração e descoberta de texto COVID-19. Além disso, os autores descrevem a mecânica da construção do conjunto de dados, destacando os desafios e as principais decisões de projeto, bem como uma visão geral de como o CORD-19 foi, as ferramentas e as próximas tarefas compartilhadas construídas em torno do conjunto de dados. |
| Ostaszewski et al. (2020) | COVID-19 Disease Map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms | Trabalho que aborda a iniciativa de desenvolvimento do mapa de doenças COVID-19, um repositório/ambiente informacional que congrega dados de pesquisa científica sobre coronavírus, constituídos por profissionais da Ciência da Computação preocupados com o compartilhamento dos dados e das informações referentes ao ciclo de replicação do SARS-CoV-2, suas interações com o hospedeiro, reação do sistema imunológico e mecanismos de reparo. Os idealizadores do projeto têm dentre seus objetivos o aspecto curatorial do repositório cruzando com dados e bases de conhecimento, para apoiar a exploração visual e computacional, bem como os esforços de modelagem de doenças. Trata-se de uma plataforma para exploração visual e análises computacionais de processos moleculares envolvidos na entrada, na replicação e nas interações patógeno-hospedeiro do SARS-CoV-2, bem como resposta imunológica, recuperação de células hospedeiras e mecanismos de reparo. O mapa irá apoiar a comunidade de pesquisa e melhorar a compreensão desta doença, para facilitar o desenvolvimento de diagnósticos e terapias eficientes. Esse trabalho aparece como uma proposta semelhante aos objetivos dessa pesquisa. |
| Nikas, Kadilierakis, Fafalios, e Tzitzikas (2020) | Keyword Search over RDF: Is a Single Perspective Enough? | Estudo baseado em uma pesquisa empírica que analisa a interação de vários métodos de apresentação dos resultados da pesquisa em conjuntos de dados RDF, por meio de linguagens de consulta estruturadas como SPARQL. Os autores comparam essa abordagem com sistemas existentes relacionados e relatam os resultados de uma avaliação baseada em tarefas com os usuários. Esse trabalho aparece como uma proposta semelhante aos objetivos dessa pesquisa. |

(continuação)

| Autores | Título | Abordagem |
|---------------------------------------------|------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Santipantakis, Vouros, e Doulkeridis (2020) | Towards Integrated and Open COVID-19 Data | Pesquisa que apresenta um protótipo para geração e atualização de um conjunto de dados abertos sobre casos confirmados COVID-19, que permite análises cross-country em diferentes níveis de granularidade. O conjunto de dados compilado é transformado em triplas RDF para preencher uma ontologia construída sobre ontologias conhecidas e os recursos são vinculados a repositórios externos de dados abertos. A base de conhecimento está sendo atualizada automaticamente e está disponível ao público por meio de um terminal SPARQL e um link de download direto. Os autores mostram como a integração de dados permite a análise espaço-temporal de dados e a descoberta de conhecimento, por meio de consultas significativas que não seriam viáveis de serem processadas de outra forma, além de permitir a interligação com outras fontes de dados disponíveis publicamente. Esse trabalho aparece como uma proposta semelhante aos objetivos dessa pesquisa. |
| Satti et al. (2020) | Ubiquitous Health Profile (UHPr): a big data curation platform for supporting health data interoperability | Destaca o problema da falta de dados de saúde interoperáveis, o que se apresenta com um grande desafio para alcançar cuidados de saúde ubíquos. Nesse sentido, a proposta Ubiquitous Health Profile (UHPr) é apresentada como uma solução de armazenamento de dados multidimensional em um mecanismo de curadoria de dados semiestruturado, que fornece suporte básico para o arquivamento de dados médicos heterogêneos e obtenção de interoperabilidade parcial de dados no domínio da saúde. A plataforma cobre várias dimensões da curadoria de dados de saúde com ferramentas de gerenciamento de big data. Os resultados oriundos dos testes e das aplicações mostraram que a plataforma alcançou interoperabilidade em nível de dados, avaliando os aspectos de tempestividade, precisão e escalabilidade. A plataforma UHPr foi avaliada juntamente com suas ferramentas de gestão associadas, como uma prova de conceito para aplicação de mediação semântica, em dados semiestruturados de saúde. |

Quadro 3. Artigos aceitos na Revisão Sistemática da Literatura.

Fonte: Elaborado pelos autores (2020)

Os dados coletados na Revisão Sistemática da Literatura revelam pouca produção científica disponível no âmbito internacional, entretanto, iniciativas interessantes já se preocupam com a abertura e a disponibilização dos dados científicos na *Web* e a adoção das tecnologias computacionais advindas da *Web Semântica* e *Linked Data*, como RDF e SPARQL, por exemplo, no auxílio à disseminação das informações sobre a COVID-19. Além disso, percebe-se que, dentre os documentos analisados, a natureza dos estudos refere-se à adoção de metodologias de caráter teórico-aplicadas, envolvendo aspectos do *Linked Data*, do RDF, da SPARQL e da utilização de ontologias para a modelagem de domínios do conhecimento.

Em relação ao estudo aplicado, a partir dos dados da *Wikidata* e das consultas realizadas no *web service Wikidata Query Service*, disponível no link <https://query.wikidata.org/>, esta seção apresenta as consultas e resultados que conjuntamente criam um painel de dados e de informação para os tópicos relacionados à COVID-19. Os resultados das consultas são atualizados simultaneamente, além de permitir outras formas de visualização, como também, este artigo apresenta gráficos com os resultados mais significativos.

Assim, na Figura 2 é apresentado o mapa do Coronavírus no mundo, apontando os números de casos e mortes de COVID-19. Como estratégia para consulta, foram extraídos os seguintes dados: quantidade de mortos (linha 6), quantidade de casos (linha 7), coordenadas geográficas (linha 8). Na linha 11, apresenta-se uma regra para apresentação dos dados a partir das coordenadas geográficas. A consulta pode ser visualizada e editada pelo link <https://w.wiki/cYB>. Além disso, na Figura 2 foi realizado um recorte para a visualização que, no momento da consulta, o Brasil apresentava 108.536 mortes e 3.359.570 casos confirmados.

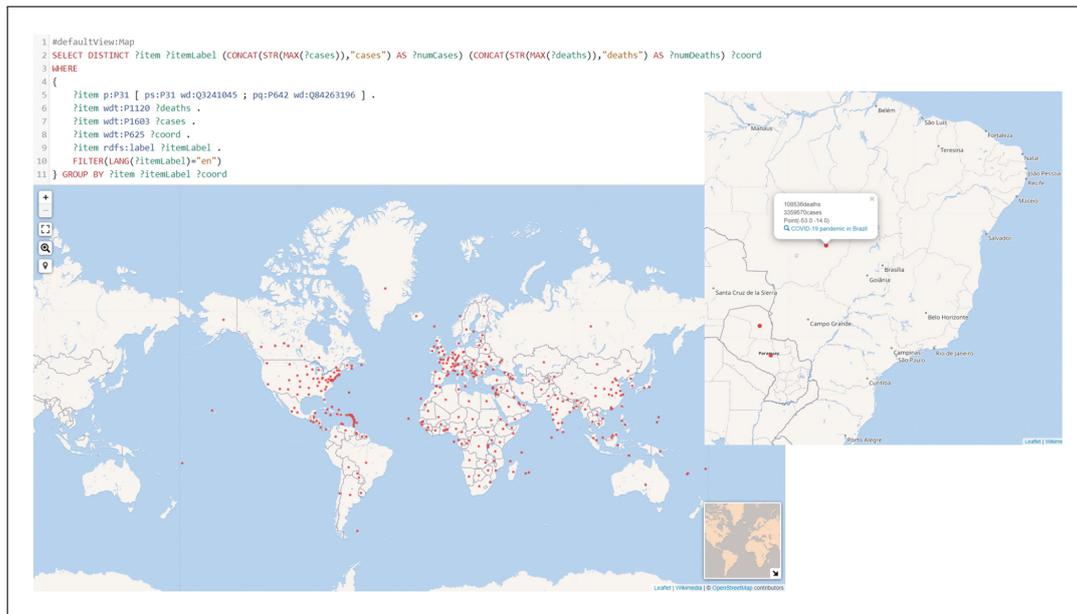


Figura 2. Número de casos e mortes de COVID-19 no mundo.
Fonte: Wikidata (2020b).

A segunda consulta realizada foi sobre os sintomas aparentes e as relações com a doença, representado pela Figura 3.

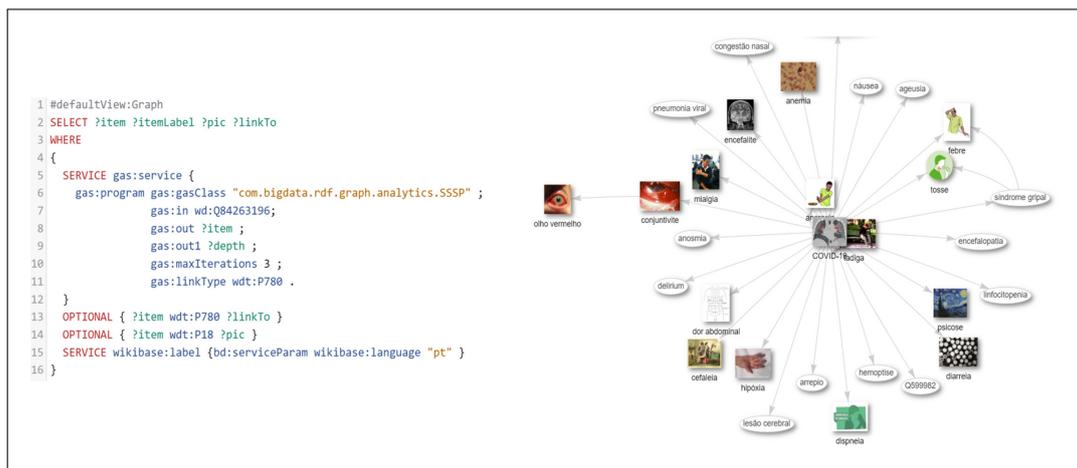


Figura 3. Sintomas relacionados à COVID-19.
Fonte: Wikidata (2020d).

A consulta pode ser visualizada pelo link <https://w.wiki/ca9>. Nas linhas 9 a 12 é visto o sequenciamento das condições para a visualização em gráfico. O recorte é feito pelos sintomas (P780) e a apresentação pode ser em texto e também em imagem (P18), nas linhas 11, 13 e 14. Outra consulta sobre possíveis tratamentos relacionados à COVID-19 também foi realizada, mas com uma estrutura de consulta semelhante a esta última consulta, e pode ser acessada pelo link <https://w.wiki/XZG>.

A terminologia do Coronavírus também foi utilizada para consulta e pode ser visualizada pela Figura 4 e pelo link <https://w.wiki/ND5>.

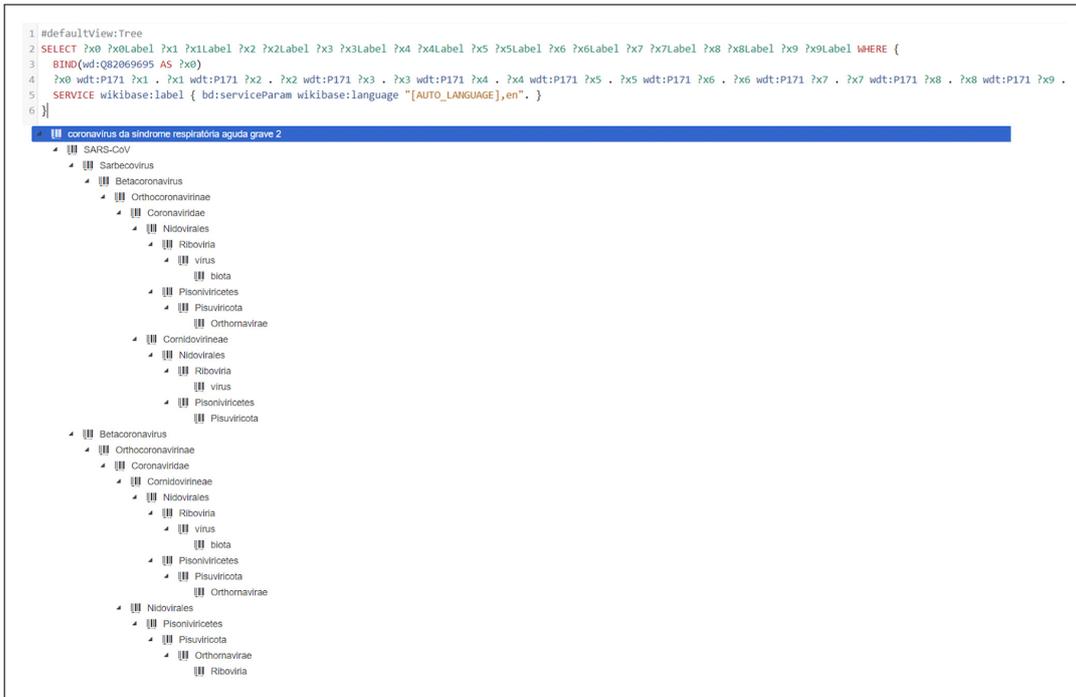


Figura 4. Terminologia do Coronavírus.
Fonte: Wikidata (2020e).

Na Figura 4 foi utilizado o identificador persistente Q82069695 - Coronavírus da síndrome respiratória aguda grave 2 na linha 3, e o P171 para a hierarquia dos termos, sendo o identificador utilizado para as taxonomias. O P171 significa: táxon imediatamente superior (P171), isto é, a categoria taxonômica imediatamente acima do táxon em questão.

A próxima consulta refere-se a imagens de pesquisas sobre a COVID-19, sendo a consulta das imagens disponível pelo link <https://w.wiki/ND4> e na Figura 5.

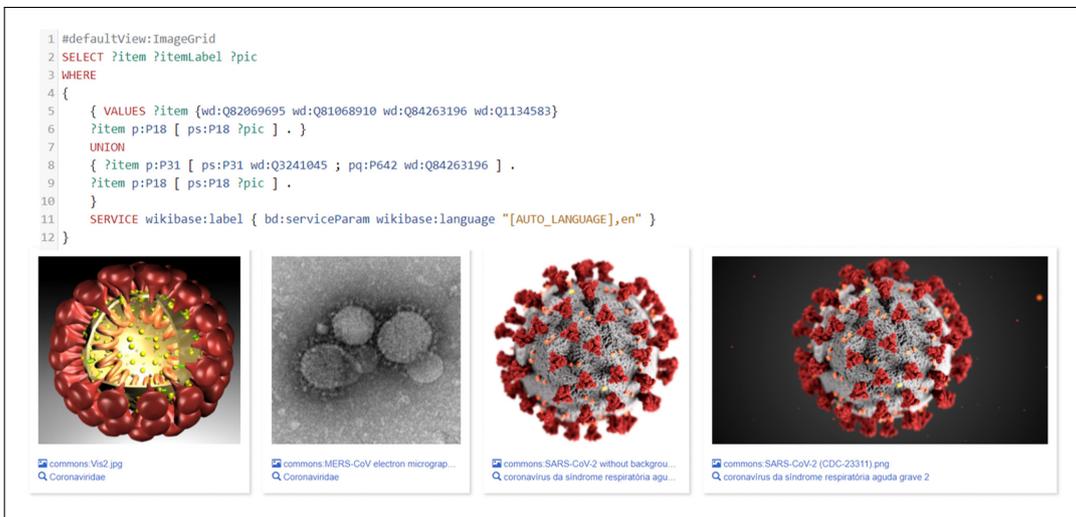


Figura 5. Imagens de pesquisas relacionadas ao COVID-19.
Fonte: Wikidata (2020a).

A imagem como dado científico respalda muitos condicionantes e resultados que a própria captura revela os trabalhos de pesquisa. Por essa razão, a imagem foi escolhida como uma fonte de informação e, ainda, como um dos tópicos de consulta para o Wikidata. No caso, foi realizada a consulta com os identificadores persistentes dos termos: coronavírus da síndrome respiratória aguda grave 2 (Q82069695), Pandemia de COVID-19 (Q81068910), (Q84263196), e Coronaviridae (Q113458). O Coronaviridae refere-se à família de vírus do COVID-19 e foi identificada pela busca sobre taxonomias.

A busca por trabalhos científicos sobre o Coronavírus foi representada na Figura 6. Foram extraídos: título do artigo na linha 9, DOI na linha 12, data na linha 14 e local de publicação na linha 13. As linhas 10 e 16 da consulta apresentam um filtro de idioma em inglês (en). As linhas 18 e 19 apresentam regras para apresentação

da consulta. A linha 18 apresenta um comando para apresentar os resultados de data decrescente. Já a linha 19 apresenta uma regra de limite dos 300 últimos resultados.

| date | venue | title | uri |
|-----------------------|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|
| 1 de janeiro de 2021 | Annals of laboratory medicine | Erratum: COVID-19 Molecular Testing in Korea: Practical Essentials and Answers From Experts Based on Experiences of Emergency Use Authorization Assays | <https://doi.org/10.3343/ALM.2021.41.1.126> |
| 1 de dezembro de 2020 | Emerging microbes & Infections | Hypothesis for potential pathogenesis of SARS-CoV-2 infection-a review of immune changes in patients with viral pneumonia | <https://doi.org/10.1093/emis/2221751.2020.1746198> |
| 1 de dezembro de 2020 | Emerging microbes & Infections | Establishment and validation of a pseudovirus neutralization assay for SARS-CoV-2 | <https://doi.org/10.1093/emis/2221751.2020.1743767> |
| 1 de dezembro de 2020 | Emerging microbes & Infections | Clinical observation and management of COVID-19 patients | <https://doi.org/10.1093/emis/2221751.2020.1741327> |
| 1 de dezembro de 2020 | Emerging microbes & Infections | The different clinical characteristics of corona virus disease cases between children and their families in China - the character of children with COVID-19 | <https://doi.org/10.1093/emis/2221751.2020.1744483> |
| 1 de dezembro de 2020 | Emerging microbes & Infections | Renin-angiotensin system inhibitors improve the clinical outcomes of COVID-19 patients with hypertension | <https://doi.org/10.1093/emis/2221751.2020.1746200> |
| 1 de dezembro de 2020 | Emerging microbes & Infections | Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients | <https://doi.org/10.1093/emis/2221751.2020.1747363> |
| 1 de dezembro de 2020 | Veterinary Quarterly | Emerging novel coronavirus (2019-nCoV)-current scenario, evolutionary perspective based on genome analysis and recent developments | <https://doi.org/10.1080/01652176.2020.1727963> |
| 1 de dezembro de 2020 | Medical Education Online | COVID-19: time for a bold new strategy for medical education | <https://doi.org/10.1080/10872981.2020.1764741> |
| 1 de dezembro de 2020 | Medical Education Online | The impact of COVID-19 on the undergraduate medical curriculum | <https://doi.org/10.1080/10872981.2020.1764740> |
| 1 de dezembro de 2020 | Medical Education Online | COVID-19: an FY1 on the frontline | <https://doi.org/10.1080/10872981.2020.1768688> |

Figura 6. Publicações em inglês sobre o COVID-19.
Fonte: Wikidata (2020c).

A consulta da Figura 6 está disponível pelo link <https://w.wiki/cYj>. A mesma consulta também foi realizada no idioma português, retornando apenas um resultado, a publicação intitulada “Da Tuberculose ao COVID-19: Legitimidade Jurídico-Constitucional do Isolamento/Tratamento Compulsivo por Doenças Contagiosas em Portugal”, que pode ser consultada pelo link <https://w.wiki/ckL>. No idioma português do Brasil, não foi encontrada nenhuma publicação.

CONSIDERAÇÕES FINAIS

A pesquisa teve como ponto central o desenvolvimento de um ambiente informacional digital a fim de apresentar um cenário de informação para os tópicos relacionados à COVID-19, por meio de consultas SPARQL e utilizando a base de dados da *Wikidata*.

No estudo teórico foram apresentadas as ferramentas tecnológicas oriundas da *Web Semântica* e *Linked Data*, tais como o modelo de dados RDF, a linguagem computacional SPARQL e exploradas na *Wikidata*. A Revisão Sistemática da Literatura oportunizou a identificação de algumas ações no âmbito internacional que têm se mostrado significativas no desenvolvimento e no compartilhamento de dados científicos que subsidiem iniciativas relacionadas ao assessoramento às informações acerca da COVID-19.

Percebe-se que a Web, como um ambiente digital, poderá favorecer o compartilhamento de dados científicos, mostrando-se um cenário informacional promissor no repasse de informações confiáveis e significativas aos usuários, a partir da adoção de modelos, arquiteturas de metadados e linguagens computacionais que possibilitarão a construção efetiva e estruturada dos dados nos domínios de conhecimento específicos.

Nesse sentido, o painel de informação para os tópicos relacionados à COVID-19 desenvolvido nesta pesquisa foi estruturado em seis categorias de conteúdo, a saber: Mapa COVID-19, Sintomas do COVID-19, Possíveis tratamentos, Taxonomia, Trabalhos relacionados e Imagens relacionadas. Cada categoria trabalha de forma sinérgica e integrada, em que as consultas às informações são realizadas na linguagem SPARQL, possibilitando o compartilhamento e o acesso aos dados na *Web* e garantindo a disseminação das informações para usuários especializados e não especializados.

Assim, esta pesquisa se apresenta como um ponto de partida naquilo que envolve a construção e a modelagem de ambientes digitais que se utilizam das tecnologias advindas da *Web Semântica* e *Linked Data*, no que diz respeito ao compartilhamento de dados e informações melhor estruturados no atendimento às solicitações dos usuários no cenário informacional COVID-19.

REFERÊNCIAS

- Berners-Lee, T. (2006). *Linked data-design issues*. Recuperado de <http://www.w3.org/DesignIssues/LinkedData.html>
- Briner, R. B., & Denyer, D. (2012). Systematic review and evidence synthesis as a practice and scholarship tool. In D. M. Rousseau (Ed.), *Handbook of evidence-based management: companies, classrooms and research* (p. 112-129). Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199763986.013.0007
- Campos, L. M., Campos, M. L. d. A., & Barbosa, N. T. (2019). A wikidata e os desafios da interoperabilidade na era dos dados abertos ligados na web: uma breve reflexão. *Prisma. com*(40), 64–77. Recuperado de <https://ojs.letras.up.pt/index.php/prisma.com/article/view/6528>
- Dutta, B., & DeBellis, M. (2020). Codo: an ontology for collection and analysis of covid-19 data. In *In 12th International Conference on Knowledge Engineering and Ontology Development (KEOD) 2-4 November*. Recuperado de <https://arxiv.org/abs/2009.01210>
- Ferreira, J. A., & Santos, P. L. V. A. d. C. (2013). O modelo de dados resource description framework (rdf) e o seu papel na descrição de recursos. *Informação & Sociedade*, 23(2), 13–23. Recuperado de <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/15436/9681>
- Fischer, T., & Jobst, M. (2020). Capturing the spatial relatedness of long-distance caregiving: A mixed-methods approach. *International Journal of Environmental Research and Public Health*, 17(17), 6406. doi: 10.3390/ijerph17176406
- Hyvönen, E. (2012). Publishing and using cultural heritage linked data on the semantic web. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 2(1), 1–159. doi: 10.2200/S00452ED1V01Y201210WBE003
- Jesus, A. F. d., & Castro, F. F. d. (2019). Dados bibliográficos para o linked data: uma revisão sistemática de literatura. *Brazilian Journal of Information Studies: Research Trends*, 13(1), 45–55. doi: 10.36311/1981-1640.2019.v13n1.08.p45
- Nikas, C., Kadilierakis, G., Fafalios, P., & Tzitzikas, Y. (2020). Keyword Search over RDF: Is a Single Perspective Enough? *Big Data and Cognitive Computing*, 4(3), 22. doi: 10.3390/bdcc4030022
- Ostaszewski, M., Mazein, A., Gillespie, M. E., Kuperstein, I., Niarakis, A., Hermjakob, H., ... others (2020). Covid-19 disease map, building a computational repository of sars-cov-2 virus-host interaction mechanisms. *Scientific data*, 7(1), 1–4.
- Paletta, F. C., & Mucheroni, M. L. (2014). O desenvolvimento da web 3.0: Linked data e dbpedia. *Prisma. com*(25), 73–90.
- Papadopoulos, D., Papadakis, N., & Litke, A. (2020). A methodology for open information extraction and representation from large scientific corpora: The cord-19 data exploration use case. *Applied Sciences*, 10(16), 5630. doi: 10.3390/app10165630
- Penteado, B. E., Bittencourt, I. I., & Isotani, S. (2019). Análise exploratória sobre a abertura de dados educacionais no brasil: como torná-los prontos para o ecossistema da web? *Revista Brasileira de Informática na Educação*, 27(1), 175-195. doi: 10.5753/rbie.2019.27.01.175
- Sampaio, R. F., & Mancini, M. C. (2007). Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. *Revista Brasileira de Fisioterapia*, 11(1), 83-89. doi: 10.1590/S1413-35552007000100013
- Santarém Segundo, J. E. (2017). Web semântica: introdução a recuperação de dados usando SPARQL. In *In Encontro Nacional de Pesquisa em Ciência da Informação, MG, Brasil, 15*. Recuperado de <http://enancib2014.eci.ufmg.br/documentos/anais/anais-gt8>
- Santipantakis, G. M., Vouros, G. A., & Doukeridis, C. (2020). Towards Integrated and Open COVID-19 Data. *arXiv*. Recuperado de <https://arxiv.org/abs/2008.04045>
- Satti, F. A., Ali, T., Hussain, J., Khan, W. A., Khattak, A. M., & Lee, S. (2020). Ubiquitous Health Profile (UHPr): a big data curation platform for supporting health data interoperability. *Computing*, 102(11), 2409–2444. Recuperado de <https://www.springerprofessional.de/en/ubiquitous-health-profile-uhpr-a-big-data-curation-platform-for-/18298800>
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70, 747–770. doi: 10.1146/annurev-psych-010418-102803
- Wang, L. L., Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., Eide, D., ... others (2020). Cord-19: The covid-19 open research dataset. *ArXiv*. Recuperado de <https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge> (22 abr. 2020)
- Wikidata. (2020a). *[Imagens relacionadas ao Covid-19]*. Recuperado de <https://w.wiki/ND4>
- Wikidata. (2020b). *[Número de casos e mortes de COVID-19 no mundo]*. Recuperado de <https://w.wiki/cYB>
- Wikidata. (2020c). *[Publicações em inglês sobre o Covid-19]*. Recuperado de <https://w.wiki/cYj>
- Wikidata. (2020d). *[Sintomas relacionadas ao Covid-19]*. Recuperado de <https://w.wiki/ca9>
- Wikidata. (2020e). *[Terminologia]*. Recuperado de <https://w.wiki/ND5>
- World Wide Web - W3C. (2017). *Boas práticas para dados na web*. Recuperado de <https://w3c.br/traducoes/DWBP-pt-br/>
- Arakaki, A. C. S., Castro, F. F. & Arakaki, F. A. (2020). Painel de Informação sobre a COVID-19: consultas SPARQL na Wikidata. *AtoZ: novas práticas em informação e conhecimento*, 9(2), 160 – 172. Recuperado de: <http://dx.doi.org/10.5380/atoz.v9i2.76684>