



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
MESTRADO EM CIÊNCIA DA INFORMAÇÃO**

ANA KAROLINE FRANÇA DE MENEZES

**ADERÊNCIA DOS PRINCÍPIOS FAIR NO CONTEXTO DA LITERATURA
ENVOLVENDO AS VACINAS CONTRA A COVID-19**

**JOÃO PESSOA -PB
2022**



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
MESTRADO EM CIÊNCIA DA INFORMAÇÃO**

ANA KAROLINE FRANÇA DE MENEZES

**ADERÊNCIA DOS PRINCÍPIOS FAIR NO CONTEXTO DA LITERATURA
ENVOLVENDO AS VACINAS CONTRA A COVID-19**

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Informação do Programa (PPGCI) da Universidade Federal da Paraíba como requisito parcial para obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Informação, conhecimento e sociedade
Linha de pesquisa: Organização, acesso e uso da informação.

Orientadora: Prof.^a. Dra. Virgínia Bentes Pinto.
Coordenador: Prof. Dr. Wagner Junqueira de Araújo.

**JOÃO PESSOA –PB
2022**

Catálogo na publicação Seção de Catalogação e Classificação

M543a Menezes, Ana Karoline França de.
Aderência dos princípios FAIR no contexto da literatura envolvendo as vacinas contra a COVID-19 /Ana Karoline França de Menezes. - João Pessoa, 2022.
139 f. : il.

Orientação: Virgínia Bentes Pinto. Coorientação: Wagner Junqueira de Araújo.
Dissertação (Mestrado)- UFPB/CCSA.

1. Ciência da informação. 2. Gestão de dados. 3. Princípios FAIR. 4. COVID-19. I. Pinto, Virgínia Bentes. II. Araújo, Wagner Junqueira de. III. Título.

UFPB/BC

CDU 02(043)

ANA KAROLINE FRANÇA DE MENEZES

**ADERÊNCIA DOS PRINCÍPIOS FAIR NO CONTEXTO DA LITERATURA
ENVOLVENDO AS VACINAS CONTRA A COVID-19**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Universidade Federal da Paraíba, como requisito parcial para obtenção do título de Mestra em Ciência da informação.

Área de concentração: Informação, conhecimento e sociedade.
Linha de pesquisa 3: Organização, acesso e uso da informação.

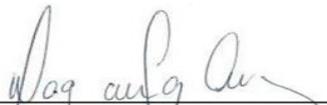
Aprovada em: 18/03/2022

BANCA EXAMINADORA



Prof. Dra. Virgínia Bentes Pinto (orientadora)

Universidade Federal do Ceará



Prof. Dr. Wagner Junqueira de Araújo (Coorientador)

Universidade Federal da Paraíba



Prof. Dr. Denyson Axel Ribeiro Mota (Membro externo)

Universidade Federal do Cariri



Prof. Dr. Guilherme Ataíde Dias (Membro Interno)

Universidade Federal da Paraíba

Prof. Dra. Andréa Soares Rocha da Silva (Suplente Externo)

Universidade Federal do Ceará

AGRADECIMENTOS

Aos amigos e familiares, por todo o apoio e pela ajuda, que muitos contribuíram para a realização deste trabalho. Aos meus pais: Francisco e Eliene, irmãs: Hanna e Junielle, assim como a meu sobrinho Guilherme, meu cunhado Marcelo, e a meus avós: Alaíde e João Duda que me incentivaram nos momentos difíceis e compreenderam a minha ausência enquanto eu me dedicava à realização deste trabalho.

Aos amigos: Arthur, Patrícia, Adriano, Maurício, Suênia, Lucas, Bruno, Carlos entre outros, que sempre estiveram ao meu lado, pela amizade incondicional e pelo apoio demonstrado ao longo de todo o período em que me dediquei a este trabalho.

À Professora Virgínia Bentes e ao Professor Wagner Junqueira por terem sido minha orientadora e meu coordenador e terem desempenhado tal função com dedicação e amizade. Aos professores, Prof. Dr. Denysson Ribeiro Mota, Prof. Dr. Guilherme Ataíde Dias, Profa. Dr^a. Andréa Silva e ao Prof. Dr. Marckson Roberto Ferreira de Sousa pelas correções e ensinamentos que me permitiram apresentar um melhor desempenho no meu processo de formação profissional.

Aos professores do PPGCI, por todos os conselhos, pela ajuda e pela paciência com a qual guiaram o meu aprendizado. A todos que participaram, direta ou indiretamente do desenvolvimento deste trabalho de pesquisa, enriquecendo minha formação acadêmica.

RESUMO

A gestão de dados de pesquisa é um processo essencial para todo projeto de investigação científica. Nesse âmbito é que foi desenvolvido o padrão Findable, Accessible, Interoperable e Reusable (FAIR), que tem como propósito obter melhoria quanto à localização, acesso, interoperabilidade e reuso dos dados em uma perspectiva de compartilhamento. Esta pesquisa definiu como **objetivo geral**, analisar o modo como está sendo efetivada a aplicabilidade da ferramenta **FAIR Data BR** nos conjuntos de dados voltados para as áreas da Saúde e da Ciência da Informação que abordam as vacinas contra COVID-19. A **metodologia** é pautada na pesquisa bibliográfico, exploratória e com uma abordagem mista tendo sido realizada em quatro etapas: (i) Análise das abordagens sobre a gestão de dados científicos e a compreensão dos princípios FAIR fazendo uma ligação com a área da Ciência da Informação; (ii) análise de um conjunto de dados perante à aderência princípios Fair Data BR, experimentando então a ferramenta de verificação, desenvolvida no âmbito do grupo de pesquisa *Web, Representação do Conhecimento e Ontologias (WRCO)* da Universidade Federal da Paraíba, extração dos dados utilizando-se do mecanismo Dataset Search e por meio das palavras-chaves: VACCINE AND COVID; (iii) tratamento e interpretação dos dados; e (iv) proposição de recomendações para uma melhor adequação dos dados aos princípios FAIR. **Resultados:** Comparando-se as quatro categorias, observou-se que houve maior aderência quanto ao reuso e ao acesso do que a interoperabilidade e a localização. Tal diagnóstico nos leva a defender que os dados são mais reusáveis, entretanto, não são tão interoperáveis, e possuem um bom acesso, embora que o acesso não seja tão fácil de localizá-lo. **Conclui-se** que a ferramenta **Fair Data BR**, apresenta certa aderência nos repositórios investigados que abordam as vacinas contra a COVID-19 evidenciando usos satisfatórios, embora ainda se tenha identificado algumas deficiências na realização da gestão dos dados.

Palavras-chave: gestão de dados; FAIR; Ciência da Informação; COVID-19; vacinas.

ABSTRACT

The management of research data is an essential process for every scientific research project. In this context, the Findable, Accessible, Interoperable and Reusable (FAIR) standard was developed, which aims to improve the location, access, interoperability and reuse of data from a sharing perspective. This research is defined as a **general objective**, to analyze how the applicability of the **FAIR Data BR** tool is being implemented in the data sets aimed at the areas of Health and Information Science that address vaccines against COVID-19. The **methodology** is based on bibliographic, exploratory research and a mixed approach, having been carried out in four stages: (i) Analysis of approaches on scientific data management and understanding of FAIR principles making a link with the area of Information Science; (ii) analysis of a data set before the adherence principles Fair Data BR, then experimenting the verification tool, developed within the scope of the research group Web, Representation of Knowledge and Ontologies (WRCO) of the Federal University of Paraíba, data extraction using the Dataset Search mechanism and by means of the keywords: VACCINE AND COVID; (iii) processing and interpretation of the data; and (iv) proposing recommendations for a better adequacy of the data to the FAIR principles. **Results:** Comparing the four categories, it was observed that there was greater adherence regarding reuse and access than interoperability and localization. Such diagnosis leads us to argue that the data are more reusable, however, they are not as interoperable, and have good access, but are not so easy to locate. It is **concluded** that the **Fair Data BR** tool has some adherence in the repositories investigated that address the vaccines against COVID-19, showing satisfactory uses, although some deficiencies in the data management have been identified.

Keywords: data management; FAIR; Information Science; COVID-19; vaccines.

LISTA DE ILUSTRAÇÕES

Figura 01-	Procedimento de análise FAIR.....	23
Figura 02-	Ciclo de vida dos dados	29
Figura 03-	Ilustração esquemática dos dados da cauda longa.....	31
Figura 04-	Fair Principles.....	35
Figura 05-	Representação esquemática da estrutura do Coronavírus.....	52
Figura 06-	Relações filogenéticas de coronavírus.....	54
Figura 07-	Transmissão, tratamento e vacinas contra a covid-19.....	56
Figura 08-	Curva de crescimento da produção científica acumulada sobre COVID-19 no Pubmed em 18 dias, de 15 de março a 1º de abril de 2020.....	57
Figura 09-	<i>Dataset Search</i>	59
Figura 10-	Ferramenta Fair Data BR.....	60
Figura 11-	Nuvens de tag dos dados da área da Saúde.....	63
Figura 12-	Nuvens de tag dos dados da área	64
Figura 13-	Tipologia de dados de pesquisa.....	66
Figura 14-	Infográfico FAIR	74
Figura 15-	Uso e compartilhamento dos dados pela ODI.....	79
Figura 16-	Interoperabilidade nos repositórios.....	83
Figura 17-	Formatos dos metadados nos repositórios.....	84

LISTA DE GRÁFICOS

Gráfico 01-	Pesquisa sobre banco de dados na BRAPCI	42
Gráfico 02-	Pesquisas sobre repositório de dados na BRAPCI.....	43
Gráfico 03-	Média temporal	61
Gráfico 04-	Média por área	63
Gráfico 05-	Tipologia de dados (FASE).....	67
Gráfico 06-	Tipologia dos dados (Natureza).....	68
Gráfico 07-	Tipo de Dado (Origem).....	69
Gráfico 08-	Média FAIR.....	75
Gráfico 09-	Média FAIR (<i>Findable</i>).....	76
Gráfico 10-	Média FAIR (<i>Accessible</i>)	77
Gráfico 11-	Média FAIR(<i>Interoperable</i>).....	80
Gráfico 12-	Média FAIR(<i>Reusable</i>).....	84

LISTA DE QUADROS

Quadro 01-	Tipologia de dados	28
Quadro 02-	Princípios FAIR.....	34
Quadro 03-	Pesquisa sobre Covid-19 na Ciência da Informação	46
Quadro 04-	Requisitos para as métricas	72

LISTA DE SIGLAS

ACS	<i>American Chemical Society</i>
ABRC	<i>Association des bibliothèques de recherche du Canada</i>
BOAI	<i>Budapest Open Access Initiative</i>
CRSH	<i>Conseil de Recherches en Sciences Humaines</i>
CNPq	Conselho Nacional de Pesquisa
DTL	<i>Dutch Techcentre for Life Sciences</i>
ECDC	<i>European Center for Disease Prevention and Control</i>
FAIR	<i>Findable, Accessible, Interoperable, Reusable</i>
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
FIOCRUZ	Fundação Oswaldo Cruz
NIH	<i>National Institutes of Health</i>
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
OMS	Organização Mundial da Saúde
ODI	<i>Open Data Institute</i>
PGD	Plano do gerenciamento de dados
RDC-DRC	<i>Research data Canadá-Données de Recherche Canadá</i>
TDICS	Tecnologias Digitais da Informação e Comunicação
WRCO	Representação do Conhecimento e Ontologias

SUMÁRIO

1 INTRODUÇÃO.....	14
2 PROCEDIMENTOS METODOLÓGICOS	20
3 GESTÃO DE DADOS CIENTÍFICOS.....	24
3.1 GESTÃO DE DADOS E REPOSITÓRIOS.....	24
4 PRINCÍPIOS FAIR	33
4.1 FINDABLE	36
4.2 ACCESSIBLE.....	38
4.3 INTEROPERABLE	39
4.4 REUSABLE	39
5 ABORDAGENS DA CIÊNCIA DA INFORMAÇÃO E A PANDEMIA DA COVID	
19.....	41
5.1 DADOS E A CIÊNCIA DA INFORMAÇÃO.....	41
5.2 CIÊNCIA DA INFORMAÇÃO NO PANORAMA DA PANDEMIA	45
6 COVID 19: UM OLHAR SOBRE O IMPACTO PANDÊMICO	51
6.1 ENTENDENDO A COVID-19	52
6.2 VACINAÇÃO E DADOS CIENTÍFICOS	55
7 PESQUISA DE CAMPO	59
7.2 RESULTADOS E DISCUSSÕES SOBRE AS MÉDIAS FAIR	69
8 CONCLUSÃO	87
REFERÊNCIAS.....	91
APÊNDICE A - TÍTULOS DOS CONJUNTOS DE DADOS.....	107

1 INTRODUÇÃO

As mudanças de paradigmas acerca das reflexões no que concerne a dados abertos de pesquisa ganharam proporções maiores nesse século XXI, todavia desde a década de cinquenta já se percebia na literatura alguns trabalhos contemplando essa temática. A pesquisa empreendida por García-García e López Borrull Peset (2015), aponta que o primeiro documento trazendo a questão de dados foi o *Journal of chemical and Engineering Data*, publicado em 1956 pela *American Chemical Society* (ACS).

Contudo, esses primeiros movimentos a respeito dos dados de pesquisa parecem não ter surtido efeito e, somente a partir do século XXI, retomam com certa força particularmente com o movimento da *Ciência Aberta* que se consolidou em 2003 com a publicação da “Declaração de Berlim sobre o Acesso Aberto ao Conhecimento em Ciências e Humanidades”. Tal declaração trouxe no seu bojo a estruturação dessa cultura de modo a favorecer o acesso nos mais diversos campos de saberes e, particularmente na temática das grandes enfermidades, a exemplo da pandemia da COVID-19, percebe-se a necessidade perante da abertura dos dados de pesquisa, para avançar a possibilidade da cura dessa doença. Tal percepção já está bem amadurecida no contexto da saúde conforme defende o *National Institutes of Health* (NIH) ao afirmar que “Os dados devem estar disponíveis livremente o mais rápido possível, salvaguardando a privacidade dos participantes e protegendo dados confidenciais e proprietários” (NIH, 2003, Online).

Outros campos também voltam seus interesses para a gestão de dados, como, por exemplo, a Ciência da Computação, a Matemática, a Administração e a Ciência da Informação. Essa última, conforme aponta Saracevic (1996), surgiu devido às evoluções tecnológicas sendo percebida como um campo interdisciplinar e que teve como foco suprir as crises informacionais surgidas na época. Em virtude disso, a área se tornou bastante abrangente tendo seu foco nas necessidades informacionais, trazendo também um olhar para a gestão de dados.

Isto posto, na perspectiva da ciência contemporânea, tem-se destacado a *E-Science*, percebida como a ciência online, em que clama por uma forma de colaboração entre os pesquisadores e na qual em larga escala produz grande volume de dados científicos. Nesse sentido

destacam-se os repositórios de dados, que conforme Sayão e Sales (2015, p. 96):

Os repositórios de dados de pesquisa são infraestruturas de base de dados desenvolvidas para apoiar todo o ciclo da gestão de dados de pesquisa, incluindo as ações mais dinâmicas e contundentes sobre os dados, que coletivamente são chamadas de curadoria de dados de pesquisa, que visam adicionar valor aos dados, avaliando, formatando, agregando e derivando novos dados.

Nesse cenário se observa que a gestão desses dados é de grande importância para que haja padronização de modo a contribuir para o reaproveitamento, interoperabilidade, acesso e localização desses dados. Corrêa (2019, p.153) retrata que “a gestão de dados é a disciplina responsável por planejar, implantar e executar estratégias, procedimentos e práticas necessários para o gerenciamento, de forma efetiva, dos recursos de dados e informações [...]”. Partindo do panorama científico tem-se hoje um volume de dados voltados não somente para os pesquisadores que a geraram, como também para comunidade científica em geral, à vista disso Madeiro (2018, p. 283) reflete que

A relevância da temática em tela é de interesse dos pesquisadores, curadores de dados, comunidades científicas, órgãos de fomento à pesquisa, editores de periódicos científicos, gestores e executores de programas e projetos de dados de pesquisa, profissionais da computação e da Ciência da Informação em geral. Devido à crescente expansão e ao compartilhamento do volume de dados de pesquisa, são necessários o estabelecimento e a consolidação de uma política de gestão de dados científicos, visando à garantia dos mesmos por meio de processos de curadoria, preservação, arquivamento e compartilhamento de coleções de dados de pesquisa, em sustentabilidade contínua e permanente.

Diante do exposto é pertinente observar a importância de profissionais aptos para trabalhar na gestão de dados, particularmente nesses novos ambientes de acesso rápido e fácil uma vez que o compartilhamento de dados está se tornando uma prática promissora na ciência. Isso se faz necessário para que os dados sejam encontrados e acessíveis de modo coerente a todos que o buscam sem que seja necessário *login*, senha, cadastros e até mesmo pagamentos a fim de se obter uma entrada nesses ambientes. Ressaltamos ainda que, em muitos casos, os dados não são interoperáveis, ou seja, não trabalham em conjunto resultando em uma falta de comunicação e troca de informações e dados, fazendo com que eles também não sejam reutilizáveis. Esses são alguns dos entraves encontrados devido à falta de gestão de dados.

No caso dos periódicos científicos, Harnad (2004) argumenta que o acesso a muitos deles passou a ser pago e, em razão disso, um novo fenômeno aparece: a chamada crise com os periódicos, contribuindo para a pouca visibilidade dos dados publicados e limitando, assim, os usuários

em potencial. Esse episódio contribuiu para a cultura do acesso aberto, acontecimento que culminou com a Declaração de Berlim, citada anteriormente, e a Iniciativa de *Open Access de Budapeste* (BOA), essa na qual propõe uma visão mais aberta sobre a pesquisa em pares, a qual apresentou como estratégia a Via Dourada e a Via Verde, que traz a viabilização dos dados utilizados nas pesquisas.

Em vista disso, conforme Henning et al. (2019) várias organizações se reuniram em 2014, destacando-se entre eles o grupo Elixir, em cooperação com as instituições: *Netherlands Science Center*, o *Dutch Techcentre for Life Sciences* (DTL) e o Centro Lorentz, organizaram uma conferência intitulada *Jointly designing a data FAIRPORT*. Tal evento deu origem aos princípios FAIR, cujo acrônimo significa *Findable, Accessible, Interoperable, Reusable* do português “Encontráveis, Acessíveis, Interoperáveis e Reutilizáveis”. Tais princípios abordam melhorias envolvendo a gestão de dados, partindo do pressuposto da ocorrência de uma rede de colaboração científica.

Tendo em mente a prática do compartilhamento de dados, e a ciência aberta, constata-se, cada vez mais, uma crescente adesão a essa prática, como dito anteriormente, principalmente após o surto do vírus SARS CoV-2, responsável pela pandemia da COVID-19, observou-se que em pouco tempo houve uma infinidade de produção de dados científicos. Santos (2020) traz que o novo vírus conhecido como 2019-nCoV, SARS CoV-2, doença de coronavírus 2019 ou COVID-19, embora tenha sido detectado primeiramente na China, espalhou-se e continua se espalhando mundialmente. Justamente por essa particularidade pervasiva é que a Organização Mundial da Saúde (OMS), no dia 30 de janeiro de 2020, classifica a COVID-19, como uma pandemia.

A COVID-19 gerou impactos sociais, econômicos e culturais, trazendo novos costumes como, por exemplo, a necessidade do distanciamento social e de políticas restritivas de movimento. As implicações desses impactos serviram de incentivo a inúmeras pesquisas nas mais diversas áreas do conhecimento, além daquelas relacionadas com as Ciências da Saúde. Assim, podemos citar as pesquisas nos campos da Educação, da Ciência da Informação, da Economia, das Tecnologias Digitais de Informação e de Comunicação, na Ciência humanas e Sociais, em cada uma dessas áreas voltando-se para seus estudos sobre a COVID-

19.

Assim, além de pesquisas com relação à doença, a sua origem, mutação, prevenção, vacinação e a hesitação com relação às novas vacinas vêm gerando um aumento na troca de dados científicos envolvendo o vírus. Essa produção documental não se constata somente nas disciplinas das áreas da Saúde, como pôde-se notar, as Ciências sociais, englobando a Ciência da Informação, assim como em outras áreas produziram dados e pesquisas crescentes relativas às vacinas contra o vírus assim como sobre a COVID-19. Em realidade essa produção documental e de dados são resultados de pesquisas em várias temáticas, tendo os estudos sobre vacinas impulsionados em grande escala, o que intensifica o uso de plataformas tecnológicas para acelerar as pesquisas, em busca da cura dessa enfermidade.

Ocasionalmente esse panorama reforça a utilização das Tecnologias Digitais da Informação e Comunicação (TDICS), de modo que se percebe a profusão de dados gerados com a pandemia, criando assim a necessidade de padrões e ferramentas que sustentem o uso dos dados, dentro dos padrões assim afirmados. Isso posto, o presente trabalho aponta como questionamento: **De que modo se dá a aderência da ferramenta Fair Data BR do nos repositórios de estudos que abordam as vacinas contra COVID-19?**

Mediante a essa conjuntura, muito se tem discutido sobre as formas de tratamento de dados, o entendimento e o estudo desses métodos que ajudem nesse processo para que assim haja otimização quanto à localização, acesso, interoperabilidade, reutilização, encontrabilidade, acessibilidade e reuso dos dados em repositórios de documentos, se mostra como um dos impulsionadores para a realização da pesquisa. Aspectos que nos motivam a empreender essa pesquisa.

Outra motivação é a interdisciplinaridade entre assuntos distintos, mas que aproxima a Ciência da Informação, a Tecnologia da Informação e a área da Saúde.

Nesse sentido, esta pesquisa tem como propósito evidenciar as contribuições dos estudos sobre as temáticas de gestão e padronização de dados para a área da Saúde e da Ciência da Informação, com foco em utilizar métodos para observar a aplicabilidade dos princípios FAIR em dados de estudos que abordem sobre vacinas contra COVID-19. Assim,

definimos como **objetivo geral**, analisar o modo como está sendo efetivada a aplicabilidade dos princípios **FAIR Data BR** nos repositórios de estudos voltados para as áreas da Saúde e da Ciência da Informação que abordam as vacinas contra COVID-19.

Os **objetivos específicos** do estudo são:

- a) Explorar as abordagens sobre a gestão de dados científicos a fim de melhor entendimento sobre a temática;
- b) Compreender a filosofia dos princípios FAIR e a sua ligação com a área da Ciência da Informação;
- c) Verificar a aderência dos princípios FAIR em repositórios de dados, utilizando os protocolos de elegibilidade e metadados que contemplem dados referentes às vacinas contra a COVID-19; Experimentar a ferramenta de verificação de aderência aos princípios Fair Data BR, desenvolvida no âmbito do grupo de pesquisa Web, Representação do Conhecimento e Ontologias (WRCO) da Universidade Federal da Paraíba.

Com a implementação da pesquisa acreditamos que suas contribuições serão importantes, tanto para o campo da Ciência da Informação como também para as Ciências da Saúde e, naturalmente, para a sociedade.

A organização da dissertação está estruturada em oito (8) sessões. Na primeira encontra-se a introdução, na qual trouxemos uma abordagem sucinta relacionada ao objeto de estudo contemplando o problema da pesquisa, a justificativa, os objetivos a serem atingidos para execução desse trabalho.

A segunda seção é dedicada ao percurso metodológico para concretização dessa pesquisa. Nela, apresentamos todos os passos e etapas, métodos e técnicas que serão utilizadas para na construção do trabalho.

Na terceira, dedica-se a gestão de dados científicos, adentrando nos conceitos da E- *Science* e de *Big data*, trazendo uma visão de outros autores, logo interagindo esse discurso, ao ciclo de vida dos dados, a tipologia dos dados, a cauda longa e ao plano de gestão de dados (GDP) e dos repositórios de dados.

A quarta, aborda-se sobre a perspectiva de ciência da informação

diante desse novo quadro pandêmico vivido no mundo, discutindo sobre a forma como a área trata os dados e trazendo também uma ponderação sobre o impacto que a pandemia causou na produção dos trabalhos científicos do campo.

Na quinta seção tem-se cada um dos princípios FAIR abordando como eles estão atrelados à Ciência da Informação, tanto quanto o uso dos princípios na perspectiva da gestão dos dados científicos. A sexta seção é dedicada à COVID-19, trazendo um contexto acerca do cenário em que essa doença se apresenta, discute-se também sobre a origem da doença, trazendo um pouco sobre sua estrutura, tipologia, contágio, tratamentos e vacinas.

Enquanto a sétima expõe os resultados encontrados na pesquisa, na qual foram colhidos os dados e analisados a aderência da ferramenta FAIR, assim como, trazendo uma reflexão dos dados colhidos apresentando os resultados quanto às tipologias dos dados, à periodicidade, levando em conta o tempo aqui estipulado, e as áreas estabelecidas no protocolo de identificação desses dados. E, finalmente, a oitava e última seção, traz as conclusões do trabalho e o impacto que a pesquisa venha a ter em distintas áreas, assim como induzir mais estudos com relação à temática em futuras pesquisas.

2 PROCEDIMENTOS METODOLÓGICOS

A metodologia de um trabalho se volta para os caminhos a serem percorridos para chegar a um denominador comum sobre a temática abordada. Desta forma, para Minayo e Sanches (1993, p. 240), “O conhecimento científico é sempre uma busca de articulação entre uma teoria e a realidade empírica; o método é o fio condutor para se formular esta articulação”. Ou seja, a metodologia e o método usados auxiliam o pesquisador a formular hipóteses, planejar e investigar um determinado fenômeno. Com base nisso, Laville e Dionne (2008, p.11) ratificam essa ideia ao apresentar que “o método indica regras, propõe um procedimento que orienta a pesquisa e auxilia a realizá-la com eficácia”. Partindo desse conhecimento, delineiam-se aqui procedimentos metodológicos para a realização da pesquisa gerando, conseqüentemente, novos conhecimentos.

O estudo científico se constrói, em princípio, por uma pesquisa bibliográfica e documental, na qual se realiza levantamento de materiais, a exemplo de artigos, livros, teses, dissertações, relatórios técnicos, relatos de experiência etc; de modo a dar sustentação para o referencial teórico acerca o tema abordado. No caso desta pesquisa, a aplicabilidade dos princípios FAIR na Ciência da Informação e na área da Saúde, assim como uma explanação acerca da gestão de dados e uma introspecção quanto a essas temáticas envolvendo o panorama pandêmico vivido atualmente.

A escolha da pesquisa bibliográfica e documental se deve justamente para evidenciar os estudos que venham ao encontro deste projeto. Dessa maneira Gil (2008, p. 69) define a pesquisa bibliográfica como:

A pesquisa bibliográfica é desenvolvida a partir de material já elaborado, constituído principalmente de livros e artigos científicos. Embora em quase todos os estudos seja exigido algum tipo de trabalho desta natureza, há pesquisas desenvolvidas exclusivamente a partir de fontes bibliográficas. Parte dos estudos exploratórios podem ser definidos como pesquisas bibliográficas, assim como certo número de pesquisas desenvolvidas a partir da técnica de análise de conteúdo.

Por sua vez, Bentes Pinto e Cavalcante (2015, p. 19) defendem que essas “pesquisas visam identificar documentos e recuperar informações registradas em textos (verbais ou não verbais, analógicos ou digitais) a fim de que possamos ter ciência sobre o que já foi ou está sendo produzido no mundo, em matéria de estudos e pesquisas” e, assim, avançar o conhecimento no que diz respeito à temática de investigação.

A vantagem dessa metodologia está na abrangência de obter várias perspectivas quanto a um mesmo assunto, conseguindo assim diferentes propostas de soluções para o problema, tendo neste trabalho autores de distintas nacionalidades, os quais discutem a mesma questão. A pesquisa traz uma abordagem qualiquantitativa, de modo a mostrar uma perspectiva diferente e aprofundada sobre a compreensão do estudo. A metodologia qualiquantitativa é híbrida e conforme o nome sugere, combina tanto dados quantitativos quanto qualitativos. No entendimento de May (2004, p. 146, tradução nossa) ao invés de visões separatistas entre quantitativo e qualitativo, como se uma dessas abordagens “[...] produzisse automaticamente uma verdade melhor do que a outra, mas aos seus pontos fortes e fragilidades na produção do conhecimento social. Para tanto é necessário um entendimento de seus objetivos e da prática”. Trata-se de uma pesquisa de natureza exploratória, visto que, conforme Moresi (2003), sua realização se efetiva em áreas nas quais há pouco conhecimento acumulado e sistematizado. Conforme Sampieri et al. (1991, p. 60), as pesquisas exploratórias “se caracterizam por serem mais flexíveis em sua metodologia em comparação com os estudos descritivos ou explicativos, e são mais amplos e dispersos que estes dois últimos tipos (por exemplo, buscam observar tantas manifestações do fenômeno estudado quanto for possível”. Ademais, Dankhe (1986, p.44, *tradução nossa*) defende que “Os estudos exploratórios servem para nos familiarizar com fenômenos relativamente desconhecidos, obter informação sobre a possibilidade de realizar uma investigação mais completa sobre um contexto particular da vida real [...]”

Por sua natureza de sondagem, esse tipo de pesquisa não comporta hipóteses que, todavia, poderão surgir durante ou ao final do estudo. É, normalmente, o primeiro passo para quem não conhece suficientemente o campo que pretende abordar. Conforme Babbie (1995, p. 84, *tradução nossa*) os estudos exploratórios “são mais tipicamente feitas para três fins: (1) simplesmente para satisfazer a curiosidade do pesquisador e o desejo de melhor compreensão, (2) para testar a viabilidade de realizar um estudo mais cuidadoso, e (3) para desenvolver os métodos a serem empregados em um estudo mais cuidadoso”.

Então, devido aos princípios FAIR ainda serem uma temática relativamente nova, assim como a COVID-19, esse estudo se propõe a trazer uma perspectiva de melhor compreensão dessas temáticas bem

como a estruturação de novos conteúdos para estudos futuros.

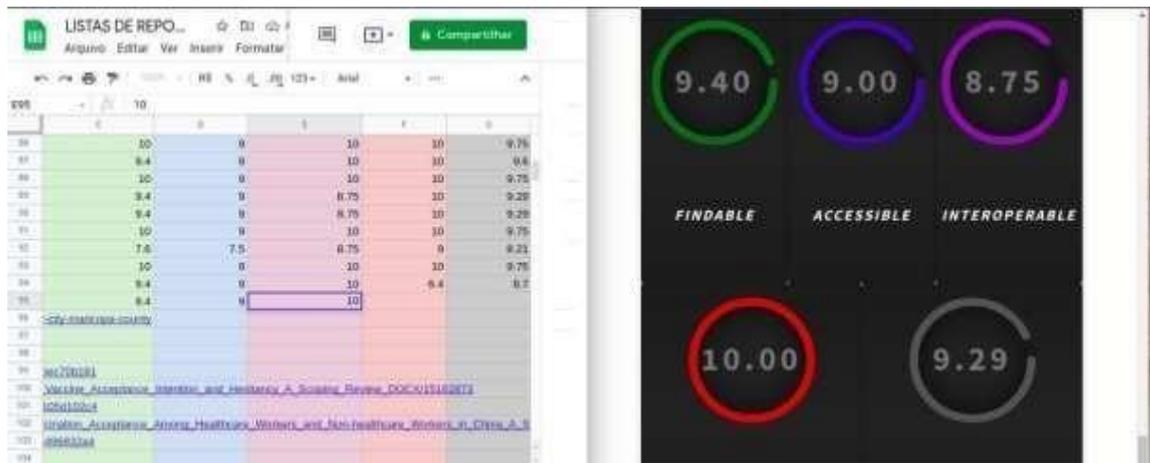
De modo que para atingir os objetivos aqui proposto, o estudo divide em duas partes: temos a **primeira, a pesquisa bibliográfica**, voltada para a pesquisa qualitativa, na qual se realizou um levantamento bibliográfico, como dito anteriormente, verificando entre artigos, teses, dissertações, livros e revistas para se obter uma visão acerca da gestão de dados científicos com o intuito de um melhor entendimento sobre a temática, o mesmo método foi aplicado para compreender a filosofia dos princípios FAIR e a sua ligação com a área da Ciência da Informação.

Para a **segunda parte, a pesquisa de campo**, tem-se o uso do método de busca, utilizou-se o mecanismo Dataset Search, em que foram adotados os termos “ VACCINE and COVID-19”, e seguindo um protocolo para a identificação e a elegibilidade dos dados, tidos como critérios de inclusão e exclusão dos dados e repositórios, esses são: a temática, de modo que na pesquisa as temáticas selecionadas englobam os assuntos aqui propostos, visto que a área da saúde está incorporada como Ciência da Vida e a área da Ciência da Informação está inserida nas Ciências Sociais.

Outro critério escolhido foi quanto aos direitos de uso, de modo a encontrar dados abertos tendo a finalidade comercial permitida e gratuita como alguns desses requisitos. Em relação ao formato de download, não se especificou precisamente para se ter diversidade quanto à tipologia dos adquiridos. O último critério adotado foi quanto a atualização do dado, de forma a escolher o último ano, especificamente entre os meses de janeiro a agosto de 2021, e assim foram extraídos os resultados correspondentes.

Partindo desses dados recuperados, em seguida aplicou-se a ferramenta de verificação de aderência aos princípios *Fair Data BR*, desenvolvida no âmbito do grupo de pesquisas Web, Representação do Conhecimento e Ontologias (WRCO), da Universidade Federal da Paraíba e cadastrado no CNPq, a qual é constituída por perguntas objetivas e de múltipla escolha pertinentes a cada princípio, tendo diante das respostas o cálculo de cada média, nas quais são somadas e seu produto dividido por quatro resultando na média FAIR, exposta na Figura-01.

Figura 01- Procedimento de análise FAIR



Fonte: Dados da pesquisa, 2021

Conforme defendem Laville e Dione (1999, p. 176) para que a observação seja qualificada como científica faz-se necessário cumprir as seguintes exigências “[...] não deve ser uma busca ocasional, mas ser posta a serviço de um objeto de pesquisa, questão ou hipótese, claramente explicitado; esse serviço deve ser rigoroso em suas modalidades e submetido a críticas nos planos da confiabilidade e da validade”.

Ainda nesse contexto, Marconi e Lakatos (2003, p. 190) definem a observação como “uma técnica de coleta de dados para conseguir informações e utilizar os sentidos na obtenção de determinados aspectos da realidade”. Não consiste apenas em ver e ouvir, porém, também, em examinar fatos ou fenômenos que se desejam estudar”.

Por meio das abordagens aqui mencionadas, o presente trabalho busca, não apenas demonstrar a importância dos princípios FAIR partindo do levantamento e da gestão de dados trazendo reflexões no tocante à interdisciplinaridade da ciência da informação, colocando-a em outras áreas.

3 GESTÃO DE DADOS CIENTÍFICOS

A gestão de dados ou “data management” tem seu nascedouro no início dos anos sessenta, com a chegada dos bancos de dados das organizações. Entretanto, ela ganha outras aplicabilidades e adentra ao contexto da exorbitante produção de dados de pesquisa e, naturalmente, emerge a necessidade de sua gestão de modo que possam vir a ser acessados, principalmente, quando se trata de pesquisas financiadas pelo setor público.

A Organização para a Cooperação e Desenvolvimento Econômico (OCDE) define **dados de pesquisa** como sendo “registros factuais (números, textos, imagens e sons), que são usados como fontes primárias para a pesquisa científica e geralmente são reconhecidos pela comunidade científica como necessários para validar os resultados da pesquisa.” (OCDE, 2007, p. 18). Por sua vez, para o “Le glossaire original de DRC” do Canadá, a gestão de dados de pesquisa compreende o “Armazenamento, consulta e preservação de dados resultantes de qualquer trabalho de pesquisa”. (RDC- DRC. 2011, online, tradução nossa). Além disso, o gerenciamento de dados afeta todo o ciclo de vida dos dados, desde o planejamento da pesquisa até a execução e desde o registro dos dados no momento de sua criação e exploração até a preservação a longo prazo dos resultados, uma vez concluída a pesquisa.” (OCDE, 2007, p.15, tradução nossa). Em realidade, a gestão de dados de pesquisa sem dúvida é “[...] muito eficaz para atender as necessidades dos pesquisadores. [...]”. Esses dados deverão ser compartilhados e as condições de acesso entre as instituições e os responsáveis pelos arquivos de dados e os produtores de dados, de modo a compartilhar as melhores práticas. (OCDE, 2007, p.20) Isto posto, nessa seção discutiremos a gestão de dados de pesquisa e repositórios.

3.1 GESTÃO DE DADOS E REPOSITÓRIOS

Embora pareça algo novo, a prática da gestão de dados e de informações está diretamente próxima a história da Biblioteconomia, da Documentação e, posteriormente, com a Ciência da Informação desde as suas origens.

Tal fato tem um desfecho fundamental para o nascimento das propostas científicas de representação, organização, recuperação da

informação e outras temáticas, de modo a favorecer as tomadas de decisões, independentemente do ponto de vista econômico, administrativo, social, cultural, científico ou do cotidiano das pessoas.

Nesse ínterim, o desenvolvimento científico e tecnológico, continuou avançando cada vez mais e, naturalmente, contribuiu para o aumento da produção científica tanto do ponto de vista de publicações: artigos, teses, dissertações, livros, relatórios técnicos, patentes e outros do gênero, porém, também, de uma imensidão de dados brutos oriundos dos trabalhos investigativos e que pouco a pouco foram se acumulando.

Com a popularização da internet e do sistema *World Wide Web*, particularmente a partir dos anos de 1990, novos rumos passam a ser pensados para um modelo vigente mais ativo de difusão do conhecimento. Tem-se uma nova atmosfera na produção científica, ocasionando mudanças que favorecem maior colaboração entre pesquisadores e possibilita maior multidisciplinaridade nas pesquisas acadêmicas, possibilitando assim, diferentes panoramas e formas de trabalhos, uma prática intitulada como *E-Science*.

A *E-Science* foi descrita por John Taylor, Diretor Geral do *Research Councils Office of Science and Technology*, como “[...] tendo a ver com a colaboração global em áreas-chave da ciência e com a próxima geração de infraestruturas que o permitirão” (Taylor, 2001, p.8, tradução nossa).

Citamos como exemplo, o primeiro projeto de banco de dados em ciências da vida denominado GenBank que foi pensado desde 1979 e está disponível publicamente em: <https://www.ncbi.nlm.nih.gov/genbank/>. Destaca-se a rede Arctic Net, que desde 2007 publica e compartilha dados de pesquisa relacionados à temática do Polo Norte. Em realidade, conforme Luce (2008, p.42, tradução nossa) a *E-Science* é vista como um novo paradigma que se caracteriza “[...] por uma colaboração global em larga escala e distribuída, utilizando tecnologias de informação distribuídas, a *E-Science* é normalmente conduzida por uma equipe multidisciplinar que trabalha em problemas que só se tornaram solucionáveis nos últimos anos, com capacidades melhoradas de recolhimento e de análise de dados.”

No entendimento de Gray (2009) a *E-Science* ocorreu quando houve a difusão da TI (tecnologia da informação) com a ciência. Esse

evento trouxe melhorias nas formas de obtenção, análise, formação, verificação e conservação de dados científicos. No que consta ao compartilhamento de dados tem-se que conforme define o glossário da FIOCRUZ (2020).

Dados compartilhados são aqueles cujo acesso, reuso e redistribuição estão limitados a determinadas pessoas ou grupos. Seja através do estabelecimento de convênios entre instituições, seja por conta de uma relação de confiança entre pesquisadores que colaboram informalmente. Tais acordos costumam determinar as condições de uso, especialmente em relação à possibilidade de exploração comercial e a atribuição de autoria. (FIOCRUZ, 2020).

No tocante à mudança na forma de interação na ciência, tem-se uma série de termos, nos quais são apresentados como ciberciência (*cyberscience*), ciberinfraestrutura (*cyberinfrastructure*) e *E-Science* (JANKOWSKI, 2007).

Em contextos brasileiros, podemos considerar que desde 2013 a ideia de *E-Science* está sendo colocada em prática no Programa FAPESP de Pesquisa *em E-Science*. Esse programa entende que a

e-Science é um nome dado em todo o mundo a pesquisas que são realizadas em todas as áreas do conhecimento e que têm necessidade de lidar com grandes volumes de dados ou de usar métodos computacionais sofisticados e computação de alto desempenho. A pesquisa em eScience aborda todas as etapas de um processo de pesquisa -- desde a criação de ferramentas computacionais que ajudam cientistas a formular problemas de pesquisa, coletar e analisar dados, até a modelagem, a simulação, a divulgação e o reuso dos resultados da pesquisa. [...]. Pesquisa em eScience pressupõe trabalho conjunto e multidisciplinar, em que cientistas da Computação auxiliem pesquisadores de outras áreas a desenvolver pesquisas de forma mais rápida e eficiente e que, nessa parceria, a Computação também seja feita de forma inovadora, transformadora (FAPESP, 2013, p. 2).

Costa e Cunha (2014) discorrem que a profusão da geração e armazenamento de dados oriundos de diversas áreas do conhecimento, ocasionando uma enorme capacidade de análise, trazem característica do que se denomina como: *Big data* e *E-Science*.

Partindo de uma noção mais ampla, tem-se que *big data* se volta para a produção e tratamento de dados em variadas tecnologias digitais e oferece uma perspectiva que aborda os volumes e as variedades em diferentes velocidades, associados a uma análise de dados em larga escala.

Normalmente, relacionado ao fenômeno da internet e das redes sociais, a *big data* surgiu em decorrência dos grandes volumes de dados fornecidos pelos usuários na *Web*, de modo que as empresas ao observar que esses dados continham informações pertinentes para o entendimento do comportamento do usuário para com sua empresa, poderiam assim

oferecer serviços direcionados às necessidades de cada cliente (SILBERSCHATZ; SUDARSHAN; KORTH, 2020).

A produção de dados científicos é algo presente na ciência, visto que essa prática se tornou ainda mais crescente com a vinculação das TDCIS no meio acadêmico, trazendo também, mudanças na percepção dos dados produzidos nas pesquisas. Conforme o *Conseil de Recherches en Sciences Humaines* (CRSH) os dados de pesquisa são constituídos por um quantitativo de dados sociais, políticos e econômicos, informações qualitativas em formato digital, dados de investigação experimental, bases de dados de imagens e sons fixos e móveis, bem como outros objetos digitais utilizados para análise analítica (CRSH, 20116, online, tradução nossa).

Por sua vez, o *National Institutes of Health Grants Policy Statement* (NIHGPS) define dados de pesquisa como sendo

[...] material factual registado, comumente aceito na comunidade científica, necessário para validar os resultados da investigação. Não inclui análises preliminares; projectos de trabalhos científicos; planos para investigação futura; revisão interpares; comunicações com colegas; objetos físicos (por exemplo, amostras de laboratório, fitas áudio ou vídeo); segredos comerciais; informação comercial; materiais necessários para serem mantidos confidenciais por um investigador até à sua publicação num periódico revisto por pares; informações protegidas por lei (por exemplo, propriedade intelectual); ficheiros pessoais e médicos e processos similares, cuja divulgação constitua uma invasão injustificada da vida privada; ou informação que possa ser utilizada para identificar uma determinada pessoa num estudo de investigação (NIHGPS, 2012, p. 68-69, tradução nossa).

Ainda nessa direção, a CRSH sustenta que a gestão de dados de pesquisa contempla

Armazenamento, consulta e conservação de dados resultantes de trabalhos ou programas de investigação. As práticas de gestão dos dados de investigação abrangem todo o ciclo de vida dos dados; da planificação dos trabalhos de investigação para a sua execução, bem como do registo dos dados no momento da sua criação e da sua exploração para a sua conservação a longo prazo após a conclusão dos trabalhos de investigação. A gestão dos dados da investigação inclui igualmente a partilha de dados, se for caso disso (CANADÁ. CRSH, 2021, online, tradução nossa).

Na mesma questão da gestão de dados de pesquisa, a *Association des bibliothèques de recherche du Canada* (ABRC) defende que essa prática vem contribuir para: acelerar o progresso científico, aumentar a visibilidade e o impacto da pesquisa, garantir o cumprimento das políticas das agências financiadoras, limitar a duplicação de trabalhos de pesquisa, facilitar a reprodução e a validação dos resultados da pesquisa e intensificar a cooperação entre pesquisadores.

Sayão e Sales (2015, p. 7) defendem dados de pesquisa como “[...] Registros factuais usados como fonte primária para a pesquisa científica

e que são comumente aceitos pelos pesquisadores como necessários para validar os resultados do trabalho científico”. De modo que, com a *E-Science*, esses dados deixaram de ser apenas um subproduto das pesquisas e passaram a ter um maior foco, tem-se hoje um cenário em que o tratamento desses se tornou uma realidade no campo científico.

Por sua vez, os dados providos de pesquisas apresentam diferentes tipos, conforme Sayão e Sales (2015) discutem e apresentam-se no Quadro-01.

Quadro 1- Tipologia de dados

TIPOLOGIA DE DADOS QUANTO À ORIGEM	
DADOS	ORIGEM DESSES DADOS
DADOS OBSERVACIONAIS	São dados obtidos por meio de observações diretas, que podem ser associadas a lugares e tempo específicos. os dados observacionais – por sua natureza instantânea – guardam uma importância crítica que os qualifica como registros históricos, pois não podem ser coletados uma segunda vez e, portanto, devem ser submetidos a processos de curadoria que os preserve para sempre.
DADOS COMPUTACIONAIS	São resultados da execução de modelos computacionais ou de simulações [...] Para esta categoria de dados a preservação por longo prazo pode não ser necessária, posto que os dados podem ser replicados ao longo do tempo. Entretanto, replicar o modelo ou a simulação no futuro pode exigir um grande número de informações que incluem descrição das dependências de hardware, software e outras dependências técnicas, e ainda os dados de entrada. É preciso notar que algumas vezes é mais conveniente preservar somente os dados de saída.
DADOS EXPERIMENTAIS	São provenientes de situações controladas em bancadas de laboratórios, em tese, dados experimentais provenientes “de experimentos que podem ser precisamente reproduzidos não necessitam ser armazenados indefinidamente; porém, na prática, nem sempre é possível reproduzir precisamente todas as condições experimentais, particularmente onde algumas variáveis experimentais não podem ser conhecidas e quando os custos de reprodução do experimento são proibitivos”
TIPOLOGIA DE DADOS QUANTO A NATUREZA	
Números, Imagens, Vídeos Ou Áudio, Software, Algoritmos, Equações, Animações ou Modelos E Simulações.	
TIPOLOGIA DE DADOS QUANTO ÀS FASES DA PESQUISA	
DADOS	FASES
DADOS BRUTOS, CRUS ou PRELIMINARES	São dados que vêm diretamente dos instrumentos científicos.

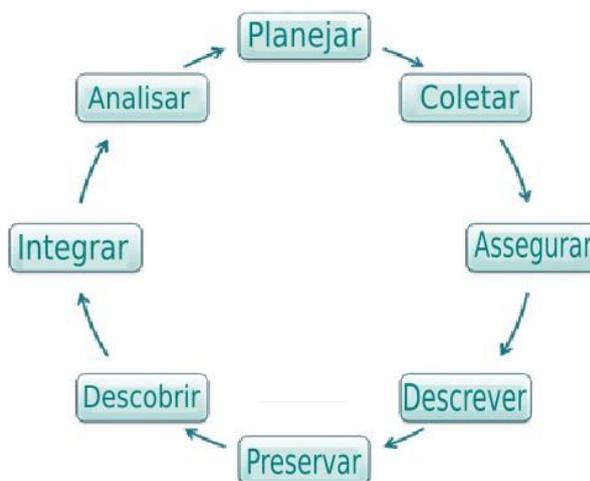
DADOS DERIVADOS	São resultados do processamento ou combinação de dados brutos ou de outros dados.
DADOS CANÔNICOS ou DADOS REFERENCIAIS	São coleções de dados consolidados e arquivados geralmente em grandes centros de dado

Fonte: Adaptado de Sayão e Sales (2015, p.8)

Logo, nota-se que os dados de pesquisa estão ligados de modo inerente à pesquisa, visto que, a partir deles pode-se ter uma ideia dos métodos aplicados ao estudo. É lícito observar a necessidade de uma gestão para melhor aproveitamento desses dados, de modo que a compreensão sobre a temática leva a estratégias e a meios de tratá-los para fins de compartilhamento.

Com base nisso, torna-se importante entender o ciclo de vida dos dados, pois com a visão ampla das etapas tem-se mais subsídios para o gerenciamento, a preservação, o uso e a reutilização dos dados. A Figura-02 evidencia o modelo idealizado pela cooperação internacional *Data ONE (Data Observation Network for Earth)*:

Figura 02- Ciclo de vida dos dados



Fonte: Traduzido de Data One (2017)

Conforme a descrição trazida pelo *Data ONE* (2017) tem-se, inicialmente, que na etapa planejar é a parte em que há a esquematização do processo de uso e reuso dos dados, de modo que o Plano do Gerenciamento de Dados (PGD) apresentará o que fazer com os dados na realização do projeto por completo. Na etapa de coleta de dados devem ser observadas as estratégias pertinentes para maior comprometimento quanto ao armazenamento de dados legíveis e interoperáveis. Para tanto,

as análises são efetuadas de forma manual ou com sensores ou outros instrumentos e os dados são colocados em formato digital.

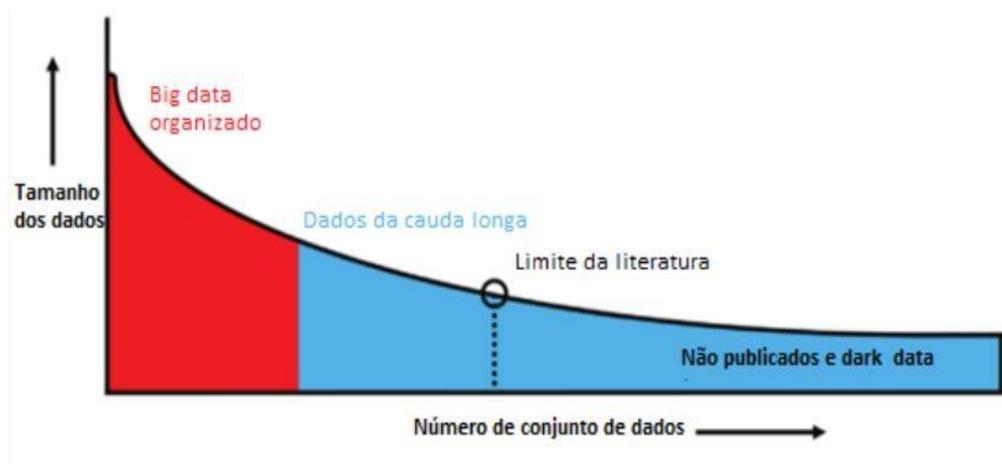
Na etapa assegurar, considera-se a valia por meio de verificações, inspeções e protocolos, tendo como objetivo evitar falhas e garantir a qualidade dos dados ali armazenados. A próxima etapa descreve-se, volta para a utilização de metadados, os quais são definidos como a forma precisa, completa e apropriada para descrever os dados.

Na etapa concernente à preservação, os dados são enviados a um arquivo de longo ou curto prazo apropriados (ou seja, *data center*). Consoante Bertin *et al.* (2017) os de curto prazo se voltam para perda de dados por acidente e a de longo prazo para usos futuros.

No que tange à descoberta dos dados, é a etapa que se atém a localizar potenciais dados úteis por intermédio de metadados, ontologia ou *thesaurus*, a fase de integração dos dados relaciona-se a combinações de fontes diferentes resultando em um conjunto homogêneo de dados que podem ser facilmente analisados. Por fim, a análise é o momento em que os dados são analisados por distintas óticas e perspectivas para que assim haja uma nova geração de dados. Em realidade, ciclo de vida dos dados, configura-se como um processo de comunicação, justamente pela dinamicidade científica que possibilita o surgimento de novos dados.

Nesse contexto, percebe-se que os dados são indicativos da veracidade do estudo, o que promove uma evolução científica, de modo que a disponibilidade desses dados permite que seu uso e reuso contribua ainda mais com novas descobertas na ciência. Em consonância a isso Córdula e Araújo (2019) trazem a concepção da *long tail* ou cauda longa que é usado para identificar a distribuição de dados nas mais variadas áreas do conhecimento, ao que concluem que “o volume de dados é distribuído de forma desigual pelos vários campos da ciência.” Pautando-nos em Almeida (2019), apresentamos na Figura 03 a cauda longa na ciência.

Figura 03- Ilustração esquemática dos dados da cauda longa



Fonte: Almeida (2019, p. 38)

Ainda nesse sentido, Almeida (2019, p. 38) versa que a “cauda longa” se constitui em um terreno fértil para novas ideias haja vista que:

Os autores apontam que a extrema direita da distribuição inclui os dados não publicados (dark data) que, apesar da riqueza, geralmente são inacessíveis. Embora possam ser resultados positivos ou negativos e existir tanto na big science quanto na small Science, grande parte dos dark data reside na cauda longa da ciência.

Fica demonstrado com a “cauda longa” que os usos de práticas para gestão de dados de pesquisa são aconselháveis para que haja transparência no uso e reuso dos dados científicos, afinal muito se produz, mas poucos fazem gestão. E essa prática faz com que alguns dados não sejam compartilhados e, conseqüentemente, não sejam reutilizados.

Dada a relevância do tratamento dos dados científicos, observa-se que o Plano de Gestão dos Dados (PGD) é uma aplicação viável durante o desenvolvimento do projeto, a fim de que eles sejam preservados e tratados de forma a serem publicados ao fim da pesquisa. Sayão e Sales (2015, P. 18) trazem que:

O PGD descreve o ciclo de vida de gestão para todos os dados que serão coletados, processados ou gerados por um projeto de pesquisa. De uma forma abreviada, ele se constitui em um documento formal que estabelece um compromisso de como esses dados serão tratados durante todo o desenvolvimento do projeto, e, também, após sua conclusão

No panorama atual, nota-se que o PGD se torna um documento cada vez mais essencial para os pesquisadores, haja vista que ele retrata os percursos dos processos e as condições dos dados, trazendo desde a coleta até a publicação desses dados. Bertin *et al.* (2017), apresenta em seu estudo os pontos de reflexão para elaboração de um PGD, indicando os tipos de dados de pesquisa, os formatos dos dados e os padrões, as políticas de acesso, o uso de dados de distribuição e a preservação de

dados e arquivamento.

Em relação à execução dos processos de gestão, tem-se a necessidade de um suporte tecnológico que entenda o funcionamento do ciclo dos dados para que assim eles sejam acessados, usados e reusados de forma correta. Dentre esses suportes há os repositórios digitais de dados de pesquisa que se tornaram recorrentes para o âmbito da pesquisa. De acordo com Sayão e Sales (2015, p. 53),

Os repositórios de dados de pesquisa são banco de dados que recebem, gerenciam e disseminam dados e conjuntos de dados de pesquisa. Eles constituem o lugar mais apropriado para que seus dados sejam preservados e possam ser recuperados, acessados e citados por outros pesquisadores, ou seja, tenham visibilidade em escala mundial.

Ainda com arrimo em Sayão e Sales (2015), percebemos que os repositórios possuem, de certa forma, uma facilidade no acesso por não possuírem barreiras geográficas nem restrições temporais, fazendo assim que sejam mais utilizados consequentemente trazendo visibilidade.

À margem disso tem-se que, de acordo com *Open Access* (2021), que os repositórios se classificam como repositórios temáticos ou disciplinares, repositórios institucionais e repositórios de dados científicos. Os repositórios temáticos se voltam para os conteúdos de uma determinada disciplina; os institucionais são criados por instituições de investigação científica e, por fim, os repositórios de dados científicos são aqueles pertencentes às distintas organizações para recolher, preservar e facilitar a partilha de conjuntos de dados resultantes de pesquisas científicas.

Tendo essa pesquisa como foco de estudo nos dados alocados em repositórios de dados, Rodrigues et al. (2010, p. 23) asseveram que “Nos repositórios de dados pode ir-se muito além desta visão de repositório de objetos, uma vez que cada conjunto de dados tem características próprias e por isso pode requerer um tratamento diferenciado. ” Seguindo esse pensamento a pesquisa em questão observou as características de cada conjunto de dados dentro dos princípios FAIR.

4 PRINCÍPIOS FAIR

Na esteira da reutilização de dados abertos de pesquisa nascem preocupações sobre como ter acesso a eles de modo mais racional e que possam contribuir para outras pesquisas. Nesse entorno é que os princípios FAIR foram propostos em 2014 no *workshop* "Design conjunto de um Data Fairport", realizado em Leiden na Holanda¹. Esses princípios se constituem em um conjunto de diretrizes visando que os dados sejam encontráveis (*Findable*), acessíveis (*Accessible*), interoperáveis (*Interoperable*) e reutilizáveis (*Reusable*) tanto por humanos como por máquinas.

Os princípios FAIR abordam melhorias quanto ao tratamento de dados partindo do pressuposto da ocorrência de uma rede de colaboração científica em consequência das mudanças na cultura da pesquisa. Assim, esses princípios oferecem aprimoramentos e modificações da cultura e prática de pesquisa de forma a ser coerente à criação de um ecossistema com serviços e componentes de dados como políticas, plano de gestão e dados, identificadores padrões e repositórios, resultando na abordagem do FAIR (MONTEIRO E SANT'ANA, 2020).

Henning (2019), ao trazer a discussão das vantagens dos dados abertos e a influência dessa ação, tanto para a sociedade, quanto para comunidade acadêmica, destaca que diante desse cenário, as devidas proteções e licenciamento e tendo como a Ciência da Informação como a área gestora, na qual ocupar-se-ia com o tratamento dos dados, tem-se assim um progresso na organização e no direcionamento de pesquisas pertinentes a cada área.

Ao falar sobre a gestão de dados, trazendo o ciclo de vida dos dados bem como o arquivamento e o cuidado a longo prazo. Para Wilkinson *et al.* (2016) existe uma necessidade de manter uma uniformização de orientações simples que auxiliem os autores a preservar os dados acadêmicos. Logo, os princípios FAIR, trazidos no FORCE11¹ *Community*, apresentam diretrizes que auxiliam os pesquisadores na conservação dos seus dados e no compartilhamento desses.

Quanto aos princípios FAIR, observa-se que não se aplicam

¹ [Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1.0 | FORCE11](#)

somente a dados convencionais, mas a algoritmos, ferramentas e fluxos destes repositórios com o propósito de capturar, tratar e ajustar os dados para melhorar a produção acadêmica. À vista disso Henning et al. (2018, p. 177) defende que

Esses princípios foram considerados o instrumento norteador para a descoberta, o acesso, a interoperabilidade global, o compartilhamento e a reutilização dos dados de pesquisa, pois se propõe estabelecer orientações estruturais para a publicação de recursos digitais, conjunto de dados, códigos e objetos de pesquisa de maneira FAIR.

Entende-se nesse sentido que esses princípios possuem um propósito de tornar os dados mais fáceis de serem encontrados, interoperáveis, compartilhados, usados e reusados. Partindo desse pressuposto, tem-se o Quadro-02, adaptado e traduzido do FORCE 11, trazendo o FAIR *Guiding Principles*:

Quadro 02- Princípios FAIR

PRINCÍPIO	ORIENTAÇÕES
Findable (Ser Localizável)	F1. metadados são atribuídos a um identificador global único e eternamente persistente.
	F2. os dados são descritos com metadados ricos.
	F3. (meta) dados são registrados ou indexados em um recurso pesquisável.
	F4. metadados especificam o identificador de dados.
Accessible (Ser Acessível)	A1 são recuperáveis por seu identificador usando um protocolo de comunicação padronizado.
	A1.2 o protocolo permite um procedimento de autenticação e autorização, quando necessário.
	A1.3 Os metadados A2 são acessíveis, mesmo quando os dados não estão mais disponíveis.
Interoperable (Ser Interoperável)	I1. metadados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento.
	I2. metadados usam vocabulários que seguem os princípios FAIR.
	I3. metadados incluem referências qualificadas a outros metadados.
Reusable (Ser Reusável)	R1. metadados têm uma pluralidade de atributos precisos e relevantes
	R1.1. metadados são liberados com uma licença de uso de dados clara e acessível.
	R1.2. metadados estão associados à sua proveniência.
	R1.3. (metadados atendem aos padrões da comunidade relevantes para o domínio.

Fonte: Adaptado e traduzido do FORCE 11 (2018)

Primeiramente, temos um identificador global ligado aos

metadados contextuais pesquisáveis, o que entra como princípio da **Facilidade**, logo esses identificadores devem agir com os metadados usando um protocolo aberto, o tornando **Acessível**, por conseguinte esses dados e metadados serão manipulados por um vocábulo controlado e/ou ontologia tornando-os assim **Interoperável**, já que por fim os dados devem ser ricamente descritos e possuir um alto cruzamento de referências para que as informações possam ser **Reutilizadas** dentro dos padrões de licença (DÍAS; ANJOS; RODRIGUES, 2019).

Com isso, os autores Dias, Anjos e Rodrigues (2019), refletem que não unicamente a ampla disseminação, o acesso aberto e a facilidade de encontrar os dados científicos devem ser levados em consideração, o controle sobre a padronização e a utilização também são pontos importantes a serem levados em conta.

Desse modo e mediante a dificuldade em tornar os princípios FAIR global o *Dutch Techcentre for Life Sciences* (DTL), apoiado pelos governos da Holanda, da Alemanha e, mais recentemente o da França, criou a iniciativa GO FAIR, um empreendimento de alguns países que visam ao *open access*, *open data* e ciência aberta.

Baseados nos princípios FAIR trouxeram como diretrizes o “*go change*”, no qual se volta para a padronização do FAIR gerando assim mudanças culturais quanto às pesquisas. O *go train*, que estabelece um treinamento para o gerenciamento de dados e o *go build* que propõe a criação de uma infraestrutura (HENNING *et al.*, 2019). Desta forma, segue na Figura- 04 a ilustração dos princípios FAIR com suas respectivas diretrizes.

Figura 04- Fair Principles



Fonte: Crabtree, 2020

Ao observar essa proposta, tem-se um interessante argumento que poderia se tornar mais geral e padronizada na comunidade acadêmica, conseguindo assim aplacar os problemas de pouco acesso à grande massa de dados produzidos.

Com referência a isso existem algumas ferramentas que tratam dos princípios FAIR para dados e metadados, sendo assim, observou-se a necessidade da construção de uma ferramenta nacional em língua portuguesa, denominada de FAIR Data BR, desenvolvida no âmbito do grupo de pesquisas *Web, Representação do Conhecimento e Ontologias (WRCO)* certificado pela UFPB e cadastrado no CNPq. Para o seu desenvolvimento foi elaborado um conjunto de perguntas que ao serem respondidas, mostram se o dado e os metadados estão disponibilizados ou não segundo os princípios de *Findable, Accessible, Interoperable e Reusable*. De acordo com isso, Menezes *et al.* (2021), ao tratar da ferramenta expõe que

A ferramenta é composta por 19 questões de múltipla escolha de caráter fechado. Foram verificadas as devidas relações existentes entre as questões ressaltando a interdependência entre elas. Ao final da avaliação, a ferramenta informa o Percentual FAIR Data BR de aderência do conjunto de dados analisado aos Princípios FAIR. Para cálculo do percentual atribuiu-se peso a cada uma das alternativas das questões da ferramenta, de modo que, cada princípio obtivesse uma nota, a partir dessas notas calcula-se a média por princípio e posteriormente a média das médias, sendo essa multiplicada por 10 e obtém-se o Percentual FAIR Data BR (Menezes *et al.*, 2021, p. 3).

Diante disso nota-se que o Percentual FAIR age como um indicador para saber se determinados repositórios de dados, de acordo com os dados e metadados disponibilizados ou se são adequados para compartilhamento mediante os princípios FAIR.

Em linhas gerais essa ferramenta oferece um controle maior, assim como uma adequação aos dados e metadados perante à necessidade do usuário, visto que esse compartilhamento auxilia na criação de pesquisas, fazendo com que gere mais resultados satisfatórios para a Ciência. Visto assim é importante entender e saber mais sobre os princípios e o que ele aborda.

4.1 FINDABLE

Uma das premissas do princípio é a localização dos dados, o **ser achado** é um dos principais pontos na gestão do dado, Teixeira (2003, p. 189) ao debater sobre a localização do dado aborda que

O uso de medidas múltiplas, mas independentes, é denominado como triangulação, que significa o uso de três pontos de referência para se verificar a localização de um objeto[...] Existem quatro categorias de triangulação: teórica (envolve o empréstimo de modelos de uma disciplina e seu uso para explicar situações em outra disciplina; de dados (refere-se a pesquisar a coleta de dados em diferentes momentos ou de fontes diferentes; por investigador (acontece quando pessoas diferentes coletam dados sobre a mesma situação e os resultados são comparados); e metodológica (ocorre quando o uso de métodos qualitativos e quantitativos de coleta de dados configura uma forma imaginativa de maximizar a quantidade de dados obtidos).

Levando em consideração essa premissa observa-se que o princípio *findable* procura incorporar as dificuldades que venham a ser encontradas na localização dos dados, de modo que a afirmativa da citação traz uma triangulação sobre a localização do objeto. Pode-se observar que os critérios atribuídos nesse princípio são pertinentes para que se encontre os dados.

Vale ressaltar que tal princípio traz uma reflexão no que diz respeito à relação homem- máquina, pois torna-se válido, ao implantar os princípios, a observação quanto ao usuário perante a prática da pesquisa de dados de forma a torná-la mais humana e menos mecanizada, algo estudado na Ciência da Informação.

Em vista disso, Wilkinson *et al.* (2016, *online*) ao abordar que a primeira diretriz desse princípio retrata que “Os metadados são atribuídos a um identificador persistente, único e global”, ou seja, para que não haja falhas ou erros é necessário que os metadados possuam um único identificador para que se possa ser encontrado.

Ainda na visão de Wilkinson *et al.* (2016, *online*), a outra diretriz, trata que os dados são descritos com metadados ricos, convergindo com outro princípio, o primeiro tópico do princípio do reuso, ao que se pode observar tem-se a importância de uma descrição detalhada, pois é ela que irá auxiliar na localização do dado a ser procurado.

Ao tratar que os metadados incluem, de forma clara e explícita, o identificador dos dados que descrevem, o autor neste tópico busca trazer que nem sempre o identificador está junto ao dado, portanto é imprescindível que haja uma associação entre eles. Henning *et al.* (2019) apresenta como exemplo quando os metadados identificadores dos dados que descrevem são indexados por um mecanismo de busca e, portanto, estão em uma plataforma diferente dos dados.

A última diretriz do princípio aborda que, os metadados são registrados ou indexados em um recurso pesquisável, ou seja, para que

haja a obtenção do dado é necessário que ele seja anexado aos metadados, em, por exemplo, um mecanismo de busca para que assim seja mais fácil achar os dados.

4.2 ACCESSIBLE

Em se tratando do princípio *Accessible*, é pertinente trazer em pauta o *open access* ou acesso aberto. Oliveira e Silva (2016) discorre que

O acesso aberto aos dados da investigação refere-se ao direito de acesso e reutilização de dados de pesquisa digitais sob os termos e as condições estabelecidas no contrato de concessão. Contudo é necessária uma política que descreva a tipologia dos dados, norteie as práticas científicas e normas acadêmico-científicas e institucionais (OLIVEIRA; SILVA, 2016, p.14).

Posto isso, observa-se que os princípios trazem uma ligação com o acesso aberto, contudo, para ser aplicados é pertinente uma política de acesso de modo a direcionar o usuário ao objeto buscado.

Portanto, a primeira diretriz desse princípio, com base em Wilkinson *et al.* (2016, *online*), traz que os metadados são recuperáveis por seu identificador, usando-se um protocolo de comunicação padronizado, ou seja, o usuário pode recuperar os dados através de um protocolo de comunicação estabelecido, advindo do identificador recuperável dos metadados.

O protocolo é aberto, gratuito e universalmente implementável, nessa premissa Wilkinson *et al.* (2016, *online*), aborda que seja qual for o protocolo implementado e as políticas de uso aderidas, tem que ser aberto, gratuito e passível de execução para quem queira acessá-lo. Ao que cabe refletir que necessariamente os protocolos têm que ser abertos, não os dados.

A última premissa do princípio *Accessible*, aborda que, os metadados são acessíveis, mesmo quando os dados não estão mais disponíveis, ao que se pode ver a importância dos metadados e a rica descrição destes a vista disso Henning (2019) discute que “[...] é preciso existir um conjunto de estratégias de preservação para os dados e metadados. Os metadados devem ser sempre acessíveis, possibilitando a criação de índices para conjuntos de dados atuais e passados (HENNING, 2019, p. 178).”

Ao trazer que o protocolo possibilita um procedimento de autenticação e autorização, quando necessário, a premissa desse princípio remonta que conforme as limitações existentes para o acesso, o protocolo

de comunicação deve ser o meio que viabilize o acesso aos dados e metadados.

4.3 INTEROPERABLE

No que se refere, esse princípio associa-se à interação entre os dados existentes no sistema, ao refletir sobre a interoperabilidade. Mello *et al.* (2015, p. 5) defende que “a interoperabilidade é a capacidade de diversos sistemas e organizações trabalharem em conjunto (interoperar), de modo a garantir que pessoas, organizações e sistemas computacionais interagem para trocar informações de maneira eficaz e eficiente.”

Em função disso, tem-se que a interoperabilidade também vai além do princípio sendo vista por outras áreas como um campo pertinente para estudo, com base nisso Mucheroni e Silva (2011, p. 5) ressaltam que

No âmbito da Biblioteconomia, Documentação e da Ciência da Informação, a interoperabilidade, de maneira geral, é entendida como a capacidade que os sistemas de hardware e software têm para se comunicar e operar com outros sistemas no intercâmbio de dados. Esses sistemas, geralmente, são de diferentes tipos, modelos e marcas comerciais distintas.

Partindo desse pressuposto, a primeira premissa desse princípio aponta que os metadados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para representar o conhecimento, sendo assim é correto afirmar que para realizar a representação dos dados e metadados deve-se empregar uma linguagem formal.

Os metadados usam vocabulários que seguem os Princípios FAIR, neste tópico observa-se que os metadados têm que seguir os vocabulários controlados, já esses por sua vez têm que estar condizentes com os princípios FAIR, de modo que os vocabulários controlados estão ativamente ligados à Ciência da Informação. Cabe afirmar que os profissionais dessa área são aptos para gerenciar esse tipo de tarefa.

Os metadados incluem referências qualificadas para outros metadados, desta forma percebe-se que no princípio Interoperable ele se preocupa com as referências estarem alinhadas aos dados para que assim haja uma conexão com o conjunto dos dados o que oferece uma ligação semântica entre eles.

4.4 REUSABLE

O último pilar do princípio FAIR, o *reusable*, também entendido como o princípio do **ser reutilizável** trata do reuso dos dados de forma a

ajudar o pesquisador em seu propósito. Visto isso, Curty (2019, p. 179, tradução nossa) expõe que

A reusabilidade é determinada por características técnicas de conformidade, preconizadas pelo FAIR, e também representa o potencial de reuso aferido, em última instância, pelo pesquisador, considerando os propósitos de sua pesquisa e com base em seu melhor julgamento sobre os atributos e condições de qualidade dos dados. Em outros termos, os dados precisam possuir certas características para serem considerados reutilizáveis, incluindo aspectos como relevância, completude, compreensibilidade e confiabilidade. Tais atributos são melhor verificáveis julgados quando os dados estão documentados, acompanhados de informações suplementares e descrições que informem sobre características como, por exemplo, a proveniência dos dados e o contexto em que eles foram obtidos, os métodos para coleta, registros, transcrição, codificação e/ou processamento dos dados, suas diferentes versões, entre outros.

Observa-se assim que, nas recomendações do FAIR, para se ter uma reutilização dos dados é necessário que ele siga alguns requisitos que os fazem ser aptos. Diante dessa premissa o primeiro tópico desse princípio traz que, os metadados são ricamente descritos com uma pluralidade de atributos precisos e relevantes, ou seja, como visto anteriormente para que se haja uma boa gestão dos dados um dos pré-requisitos é a sua descrição rica e que tenha ligação com os dados, pois tanto a interoperabilidade, a encontrabilidade, quanto o acesso e o reuso se projetam nessa prática.

Ao tratar a disponibilidade dos metadados com uma licença de uso de dados clara e acessível, o próximo item do princípio refere-se à importância de se ter uma licença de uso bem definida e clara para que o usuário entenda quais os dados que se pode ter acesso e sobre quais condições.

O próximo item, no qual se volta para a associação dos dados a uma proveniência detalhada Henning (2019) detalha que as informações pertinentes ligas a proveniência são:

(a) A linhagem dos dados, ou seja, o processo de obtenção do dado (gerado ou coletado). (b) Particularidades ou limitações sobre os dados que outros usuários devem conhecer. (c) Data da geração do conjunto de dados, condições de laboratório, quem preparou os dados, configurações de parâmetros, nome e versão do software utilizado. (d) Explicitar se são dados brutos ou processados. (e) A versão dos dados arquivados e/ou reutilizados deve ser claramente especificada e documentada (HENNING et al., 2019, p. 179).

De modo que essas diretrizes auxiliam tanto o usuário, quanto o gestor dos repositórios de dados. O último ponto trazido contempla que, os metadados deverão estar de acordo com padrões comunitários relevantes para o domínio, o que indica que devem estar dentro das normas de cada área atendendo assim uma boa prática de armazenamento e compartilhamento com a comunidade acadêmica.

5. ABORDAGENS DA CIÊNCIA DA INFORMAÇÃO E A PANDEMIA DA COVID 19

Em vista da atual situação, percebe-se que muitas áreas do conhecimento estão mudando suas abordagens no que diz respeito aos campos que ocupam, dentre elas têm-se a Ciência da Informação, na qual, atualmente, vem abrangendo seu foco em decorrência de um aumento na perversão da informação e de dados reforçado pela então pandemia da Covid-19.

Com efeito, tem-se um crescimento quanto à produção de pesquisas científicas ligadas às formas como essas informações e esses dados estão sendo gerenciados. Nota-se que todos esses aspectos estão diretamente ligados às pesquisas e estudos em Ciência da Informação que vem se dedicando à gestão de dados de pesquisa, tanto do ponto de vista teórico como nas ações de trabalho dos profissionais da área, está trazendo contribuições ímpares nesse contexto.

Desta forma, nessa seção abordamos sobre os dados e a Ciência da Informação e a Ciência da Informação no panorama da pandemia.

5.1 DADOS E A CIÊNCIA DA INFORMAÇÃO

No prisma da Ciência da Informação é importante observar que sua vertente multi e interdisciplinar possibilita essa transição entre conteúdos de modo a gerar uma ligação entre diversos assuntos resultando em uma alta carga de conhecimento produzido. Em vista disso, ao trazer um olhar mais aprofundado sobre a Ciência da Informação, Borko (1968, p. 5) a define como

Uma ciência interdisciplinar que investiga as propriedades e o comportamento da informação, as forças que governam os fluxos e os usos da informação, e as técnicas, tanto manual quanto mecânica, de processamento da informação, visando sua armazenagem, recuperação, e disseminação ideal.

Uma vez que a informação é o foco dessa área, muito se discute sobre os conceitos de informação e dado, posto que ao se falar de dados remete-se logo às Ciências exatas e as TIDCS, e ao se falar em informação associa-se às Ciências Sociais Aplicadas, todavia, observa-se que o dado ao se relacionar com outro dado produz uma informação, desta maneira pode-se dizer que o dado também é o foco das Ciência da Informação (Souza e Almeida, 2021).

Com base nisso vários autores da Ciência da Informação trazem

pesquisas envolvendo o dado, a informação e o conhecimento, Santos e Santana (2002, p. 3) por exemplo ao se basearem em outros autores definem que o termo dado como “um elemento básico, formado por signo ou conjunto finito de signos que não contém, intrinsecamente, um componente semântico, mas somente elementos sintáticos”.

Outros, como Semeler e Pinto (2019, p. 117) destacam que os dados são peças de informação digital distintas, geralmente formatadas de maneira particular. Os dados digitais são considerados um tipo de informação de natureza digital. ” Perante às afirmativas, observa-se que a Ciência da Informação possui essa correlação acerca dos dados de modo geral.

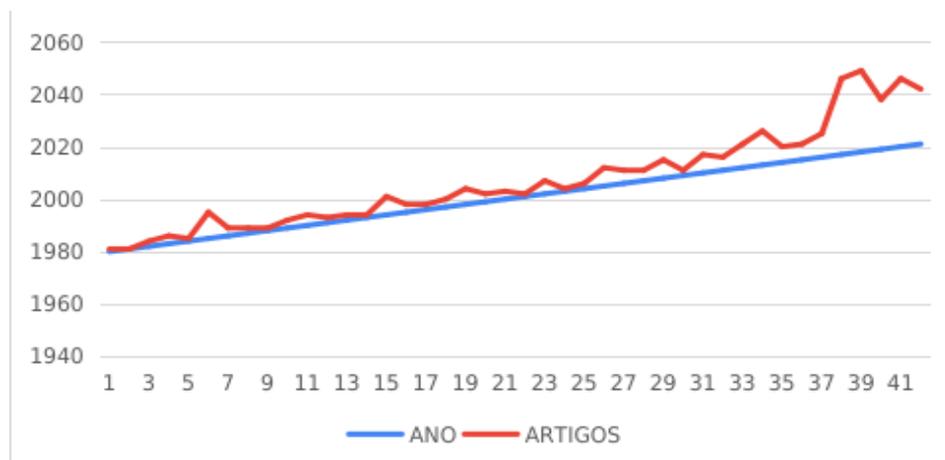
Convém ressaltar que os dados constituem o centro da pesquisa e precisam ser organizados e assimilados, demandando de um processo de gestão, com efeito, tem-se os dados e a Ciência da Informação atrelados.

A ciência da Informação pode e deve contribuir para que este cenário de acesso e uso intenso de dados se desenvolva da melhor maneira possível, buscando identificar e estudar fatores e características que propiciem ampliação do equilíbrio entre os atores envolvidos no processo e a máxima otimização do uso dos dados (SANT’ANA, 2016, p.119).

À guisa disso tem-se que a linha tecnológica da CI abarca também estudos com enfoque sobre dados, em decorrência disso, ao observarmos o cenário nacional em relação a estudos decorrentes da coleta, da organização e da disseminação de dados associado a base de dados.

Ao realizar uma busca sobre banco de dados na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI) verificou-se que 276 artigos, entre os anos de 1972 a 2021, como mostra no Gráfico-1.

Gráfico 01 -Pesquisa sobre banco de dados na BRAPCI

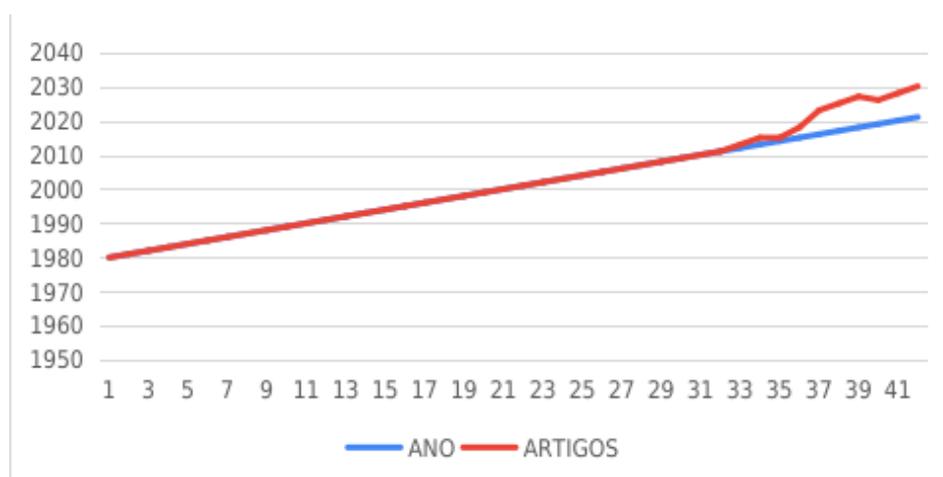


Fonte: Dados da pesquisa ,2021

Com isso, nota-se que a temática vem sendo abordada desde os anos 80 ganhando força com o passar do tempo, chegando a uma grande produtividade de estudos referentes ao tema, sendo assim esse crescimento representado no gráfico pela linha vermelha, tendo subsequente a linha azul na qual denota o passar dos anos. Resultado parecido foi encontrado no estudo de Santa Anna (2018, p. 73), no qual realizou seu estudo em 2017, trazendo um recorte de 1972 a 2017, ele declara que considerando o recorte temporal adotado para esta pesquisa, qual seja de 1972 a 2017, nota-se que o primeiro artigo sobre periódico científico aparece em 1986; o segundo trabalho no ano posterior; o terceiro, apenas 11 anos depois, em 1998; o quarto artigo foi publicado em 1999 e, no ano de 2000, não houve nenhuma publicação. No início do século XXI, o número de publicações oscila, havendo publicações em todos os anos, com crescimento mais elevado no ano de 2010 (com nove artigos), e por fim, nos dois últimos anos, em 2015 e 2016, manifestam-se três e duas publicações, respectivamente.

Apresentando uma perspectiva associada aos repositórios de dados, observa-se que ao processar a busca na BRAPCI acerca dos repositórios de dados, tem-se 59 resultados entre os anos de 1972 a 2021, como mostra no Gráfico-02.

Gráfico 02- Pesquisas sobre repositório de dados na BRAPCI



Fonte: Dados da pesquisa ,2021

Mediante a isso, nota-se que apesar da área possuir uma grande base literária para as bases de dados, os repositórios ainda estão em processo de evolução quanto a estudos nacionais na área da Ciência da Informação, representado assim no gráfico as pesquisas sobre repositório

de dados pela linha vermelha trazendo o crescimento da produção desses trabalhos, tendo também a linha azul que denotando o passar dos anos e mostrando que apesar da temática de gestão de dados ser vista como uma novidade, a Ciência da Informação há tempos possui envergadura para a curadoria dos dados, em seus estudos sobre as bases de dados e agora tendo um aumento quanto às pesquisas sobre repositórios, em consideração a afirmativa a seguir:

Verifica-se, então, que a curadoria digital para a área de Ciência da Informação vem sendo usada em especial com foco em três abordagens: na filtragem e seleção de dados na Web (Curadoria de Conteúdo ou de Informações), na agregação de valor, gestão ativa e preservação de dados digitais (Curadoria Digital) e na curadoria de dados de pesquisa (e-Science) (SIEBRA, BORBA E MIRANDA, 2016, p. 27)

Em consonância a isso, Márdero-Arellano e Andrade (2006) trazem a concepção que para a realização de uma boa curadoria digital, dependendo do contexto, os profissionais da informação são considerados os mais adequados, tendo em vista as bibliotecas digitais. Deste modo, observa-se que o profissional da informação poderá ter mais facilidade quanto à preservação dos dados de pesquisa. Corroborando essa ideia, Costa (2017, p. 5) retrata que

Na Ciência da Informação, a e-Science traz implicações relevantes sobre a comunicação científica, afinal os dados oriundos da e-Science são de fato dados científicos primários. Também traz implicações para os serviços e produtos de informação, bem como afeta diretamente as bibliotecas digitais, porque exige reflexões sobre a preservação digital e o planejamento das bases de dados.

Pode-se, então, dizer que a Ciência da Informação, por ter como enfoque a geração, coleta, representação, organização, interpretação, armazenamento, recuperação, disseminação, transformação e uso da informação, como apresentado por Capurro e Hjørland (2007), se torna uma área propensa a aderir à gestão de dados, visto que há uma preocupação semelhante por se tratar do tratamento dos dados, desde sua coleta ao descarte.

Com essa percepção, Santos e Santana (2013, p. 200) alegam que “A ciência da Informação refere-se à atividade direcionada à pesquisa de princípios e métodos que são partes da análise, do projeto e da evolução dos sistemas de informação”, sem embargo, pode-se observar, que a área, pelas suas características e pela sua vertente tecnológica faz com que os profissionais sejam aptos para a realização da gestão dos dados, sendo assim, no que tange ao profissional da Ciência da Informação na prática da gestão.

Devido à Ciência da Informação possuir essa vertente interdisciplinar, assim como vertentes tecnológicas, observa-se então esse alinhamento no que concerne à gestão de dados, visto que ambas se voltam para práticas que auxiliem na resolução de problemas no campo científico, à vista disso Costa e Sousa (2021) tratam que:

Assim, um dos pontos que justificam esta pesquisa corresponde a identificação de que maneira a CI vem tratando à Gestão de Dados, haja vista a relevância do gerenciamento de dados para solução de problemas informacionais e, conseqüentemente, para promoção de acessibilidade da informação (COSTA; SOUSA, 2021, p.61).

Quanto à gestão de dados nas áreas de conhecimento, convém ressaltar que não são somente as empresas comerciais ou de serviços que se beneficiam de seu uso, porém, também outros ramos de atividades, como a área da saúde também já está inserindo esse recurso em suas práticas, de maneira a realizar melhor proveito dos dados produzidos.

Na área de saúde, há uma tendência também por produção de dados semi ou não estruturados, representados por imagens, documentos, sinais de monitoração etc. O crescimento de sistemas de análise digital de imagens médicas tem também sugerido um aumento exponencial de uso de Big Data nesta área. Equipamentos e softwares têm sido desenvolvidos para atuar de forma integrada nos exames de imagens, com possibilidade cada vez maior de aplicações de algoritmos de IA no apoio às interpretações desses elementos. (BARBIERI, 2020, p. 107)

Com base nessas reflexões, pode-se dizer que tanto o *E-Science* quanto a *Big Data* geraram um impacto não apenas na cultura científica, bem como na aplicabilidade de algumas áreas, em que estas se viram propensas a mudar seus serviços e suas relações para uma adaptação desse novo modo de produção e de tratamento científico.

5.2 CIÊNCIA DA INFORMAÇÃO NO PANORAMA DA PANDEMIA

Muito tem-se discutido acerca dos efeitos causados pelo surto da COVID-19, como mostrado anteriormente, muitas áreas sofreram com o impacto da pandemia, e a Ciência da Informação foi uma delas.

Observando o cenário atual, mesmo não sendo objeto de estudo desta pesquisa, ainda assim, fez-se um pequeno levantamento no Portal de Periódicos da CAPES, e detectou-se um alto nível de produções científicas da área da Ciência da Informação voltadas para a temática da COVID-19. De posse das respostas, selecionou-se alguns trabalhos, dos anos de 2020 a 2021 (tendo em vista que não se constatou trabalhos na área da Ciência da Informação que tratem sobre a COVID-19 nos anos anteriores) que representam uma pequena amostra da repercussão da

produção científica que a pandemia ocasionou na área. Vejam-se o

Quadro-3.

Quadro 3- Pesquisa sobre Covid-19 na Ciência da Informação

TRABALHOS	ANO	REVISTAS	AUTORES
Desinformação, infodemia e caos social: impactos negativos das fake News no cenário da COVID-19	2021	Em Questão	FERREIRA, J. R. S.; LIMA, P. R. S.; SOUZA, E. D.; SOUZA, E. D.
Informação em tempos de Covid-19 e patrimônio cultural de ciência e tecnologia: de onde viemos e para onde vamos	2020	Liinc em revista	ABALADA, V. E. T. M.; ARAÚJO, B. M.; GRANATO, M.; RIBEIRO, E. S.
Arquitetura da informação no contexto da informação em saúde: um olhar para o website do COVID-19 no Brasil	2020	AtoZ: Novas Práticas em Informação e Conhecimento	BRITO, J. F.; SILVA, R. C.; SANTOS, B. R. P.; MELLO, M. R. G.; MARTÍNEZ-ÁVILA, D.
Desinformação sobre o Covid-19 no WhatsApp: a pandemia enquadrada como debate político	2021	Ciência da Informação em Revista	SOARES, F. B.; RECUERO, R.; VOLCAN, T.; FAGUNDES, G.; SODRÉ, G.
A Ciência da Informação e a COVID-19: análise de sentimentos dos usuários no Twitter	2020	Páginas A&B, Arquivos e Bibliotecas (Portugal)	FRANÇA, A. L. D.; ALVES, I. N. C. N.; SILVA, D. E. L. S. E.
Infodemia no contexto da pandemia de COVID-19 no Brasil: uma política de contaminação?	2021	Liinc em Revista	CEZAR, L. S.; MACIEL, A. J. N.
A representação temática da Covid-19 na literatura de cordel	2020	Informação & Informação	LAGE, S. R. M.; LUNARDELLI, R. S. A. A
Dimensões da competência em informação: reflexões frente aos movimentos de infodemia e desinformação na pandemia da Covid-19	2020	Liinc em revista	MATA, M. L.; GRIGOLETO, M. C.; LOUSADA, M.
REPOSITÓRIOS BRASILEIROS SOBRE A COVID-19: colaboração para o enfrentamento da pandemia	2021	Páginas A&B, Arquivos e Bibliotecas (Portugal)	CRUZ-RIASCOS, S. A.; SILVA, P. R. F.
BIBLIOTECÁRIO CLÍNICO EM AÇÃO NA PANDEMIA DA COVID-	2020	Revista Bibliomar	SOUZA, A. D.; JAVIER JUNIOR, G. F.; FERNANDES, M. R

19: recursos de informação em saúde para tomada de decisão			
A iniciativa digital CONVIDE-i9 no combate à infodemia de COVID-19: breves apontamentos de atuação	2020	AtoZ: Novas Práticas em Informação e Conhecimento	FRANÇA, F. P.; ARAÚJO, D. O.; SILVA, M. B. A
COVID-19 e a circulação de informações em redes sociais: análise em um grupo brasileiro no Facebook sobre o Coronavírus	2021	Em Questão	CORDEIRO, D. F.; ROCHA, A. S.; VIEIRA, L. M.; CASSIANO, K. K.; SILVA, N. R.
Profissionais da informação: atuação resiliente durante a pandemia de Covid-19: atuação resiliente durante a pandemia de Covid-19	2021	Em Questão	FARIAS, M. G. G.; FARIAS, G. B.; SOUSA, F. L. M.; ABREU, P. M. H.; BARRETO, R. B.; CABRAL, V. H
Acesso aos serviços e produtos da biblioteca em tempos de pandemia da COVID-19: possibilidades de uso do Facebook	2021	Informação & Informação	LESSA, B.
Não faça a “egípcia” em tempos de COVID 19: LGBTI+ competentes em informação	2020	Revista Brasileira de Educação em Ciência da Informação	VALÉRIO, E. D.; SILVA, A. L. A.
Iniciativas de acesso aberto no combate à pandemia: dados abertos e propriedade intelectual na disseminação da informação e conhecimento	2021	Revista Digital de Biblioteconomia & Ciência da Informação	ROSA, S. S.; SILVA, F. C. C.; PAVÃO, C. M. G.
Sala de aula invertida no enfrentamento fake News, desinformação e infodemia em época de Covid-19	2021	Revista ACB: Biblioteconomia em Santa Catarina,	SANTOS, L. R. D.; ANDRADE, E. L. M.; LIMA, E. F.; FERNANDES, J. C. C.
Desinformação e comportamento informacional nas mídias sociais: a divulgação científica na prevenção ao novo coronavírus	2021	Informação & Informação	CORRÊA, M. V.; CAREGNATO, S. E.
A Educação em Ciências e Saúde e o enfrentamento à desinfodemia: um relato de experiências críticas no	2021	Liinc em revista	CORDEIRO, J. D. R.; FONSECA, A. B.; LESSA, L. R.; LIMA, A. G. I.; NOBILE,

ensino online			
O papel das bibliotecas durante a pandemia da Covid-19: em busca da emancipação humana	2021	Ciência da Informação em Revista	SANTOS, L. R.; ANDRADE, E. L. M.; GUIMARÃES, M. B. H.; ISAÍAS, P. H.; SILVA, L. H.; LIMA, E. F.; FERNANDES, J. C. C.; MATOS, F. B.
As fakes news em um contexto de pandemia pelo coronavírus	2021	Informação em Pauta	GUIMARÃES, J. A. C.; DALESSANDRO, R. C.

Fonte: Dados da pesquisa, 2022

Pode-se afirmar que, em razão do quadro atual, as temáticas das pesquisas estão passando por algumas mudanças, de forma que muitos trabalhos trazem enfoque sobre o combate à infodemia, desinformação e fake news. Como a pesquisa do Guimarães e Dalessandro (2021) em seu trabalho intitulado **As Fakes News em um contexto de pandemia pelo coronavírus** trazem uma categorização temática das fake news relativas ao COVID-19 no universo brasileiro, fazendo um *fact-checking*, acerca das informações disseminadas sobre a pandemia, de modo que assim como houve um crescimento na produção de dados acerca da doença, também houve um crescimento de produção de informações sobre a doença.

Outro trabalho que envolve o enfrentamento das fake news é o das autoras Santos, Andrade, Lima e Fernandes(2021) intitulado **Sala de aula invertida no enfrentamento a Fake News, Desinformação e Infodemia em época de Covid-19**, o qual com o propósito de trazer umavisão com um viés sobre a infodemia, fake news e desinformação objetivou tratar sobre a competência em informação, e na perspectiva da educação profissional e tecnológica, para estudantes do curso técnico, integrado ao ensino médio, do Instituto Federal Goiano, ao que mostra a importância desse trabalho diante da avalanche de informações ao retratarsobre senso crítico e a buscas corretas de informações.

De modo similar tem-se o trabalho do Ferreira, Lima e Souza (2021) denominado **Desinformação, infodemia e caos social: impactos negativos das fake news no cenário da COVID-19**, no qual tem como propósito evidenciar a correlação entre essas temáticas, tomando como exemplo o cenário caótico marcado pela pandemia da COVID-19.

Outros trabalhos como **Dimensões da competência em**

informação: reflexões frente aos movimentos de infodemia e desinformação na pandemia da Covid-19, das autoras Mata, Grigoletto e Lousada(2020), **Desinformação e comportamento informacional nas mídias sociais: a divulgação científica na prevenção ao novo coronavírus**, dos autores Corrêa e Caregnato(2021), **A Ciência da Informação e a COVID-19: análise de sentimentos dos usuários no Twitter** de França, Alves e Silva (2020), **COVID-19 e a circulação de informações em redes sociais: análise em um grupo brasileiro no Facebook sobre o Coronavírus**, bem como **Desinformação sobre o Covid-19 no WhatsApp: a pandemia enquadrada como debate político** de Soares *et al.* (2020) preocuparam-se em estudar as ações do indivíduo em meio a esse caos informacional de modo a observar suas impressões, mobilizações e o sentimento de polaridade dos usuários levando em conta a nova situação enfrentada, como por exemplo o isolamento social, a hesitação quanto às novas vacinas entre outrotópicos e assim analisar como ocorre o quadro de desinformação e infodemia, ao que nota assim a importância do profissional da informação para discutir e auxiliar no enfrentamento desses novos problemas informacionais.

Em consonância a isso, pesquisas como: **Profissionais da informação: atuação resiliente durante a pandemia de Covid-19: atuação resiliente durante a pandemia de Covid-19** de Farias *et al.*(2021), **Educação em Ciências e Saúde e o enfrentamento à desinfodemia: um relato de experiências críticas no ensino online** de Cordeiro *et al.*(2021).

Não faça a “egípcia” em tempos de COVID 19: LGBTI+ competentes em informação de Valério e Silva (2020), tratam sobre o seguimento de profissionais da informação de forma que o impacto causado pela pandemia gerou novos focos de estudo e interesse para o profissional da área, sendo visto como agente auxiliador contra a nova onda da desinformação, infodemia e fake news.

Com relação aos meios que auxiliam o enfrentamento desses novos problemas informacionais têm-se estudos mostrando como as bibliotecas e os bibliotecários fazem parte dessa nova situação, pesquisas como: **Bibliotecário Clínico Em Ação Na Pandemia Da Covid-19: recursos de informação em saúde para tomada de decisão** de Souza,

Javier Junior e Fernandes (2020), **O papel das bibliotecas durante a pandemia da Covid-19: em busca da emancipação humana** de Santos *et al.*(2021) e **Acesso aos serviços e produtos da biblioteca em tempos de pandemia da COVID-19: possibilidades de uso do Facebook** de Lessa (2021) que trazem o papel fundamental das bibliotecas nessa empreitada, mostrando como socialmente a biblioteca e o bibliotecário possuem uma responsabilidade e uma contribuição com informações de fontes seguras, colaborando com o senso crítico e consciente dos cidadãos.

Outros pontos trazidos na área são acerca da colaboração das TIDICS nesse quadro pandêmico, o que é mostrado nas pesquisas **REPOSITÓRIOS BRASILEIROS SOBRE A COVID-19: colaboração para o enfrentamento da pandemia** de Cruz-Riascos e Silva(2021), **A iniciativa digital CONVIDE-i9 no combate à infodemia de COVID-19: breves apontamentos de atuação** de França, Araújo e Silva(2020), **Iniciativas de acesso aberto no combate à pandemia : dados abertos e propriedade intelectual na disseminação da informação e conhecimento** de Rosa, Silva e Pavão(2021), e **Arquitetura da informação no contexto da informação em saúde: um olhar para o website do COVID-19 no Brasil** de Brito *et al.*(2020) isso mostra que , como dito anteriormente, as tecnologias são vistas como recursos viáveis para o enfrentamento da pandemia.

Por fim tem-se a repercussão cultural que esse novo panorama trouxe, isso fica explícitos nos trabalhos como **Informação em tempos de Covid-19 e patrimônio cultural de ciência e tecnologia: de onde viemos e para onde vamos** de Abalada *et al.* (2020) e **A representação temática da Covid-19 na literatura de cordel** de Lage e Lunardelli(2020) que mostram a inserção de novos horizontes e novos termos sendo inseridos em nosso meio.

Como pode se vê o cenário da pandemia trouxe impactos sociais, tecnológicos e culturais, trazendo não apenas mudanças na disseminação da informação, mas também novos caminhos para o profissional da informação, as bibliotecas e assim mudando o rumo da Ciência da Informação, entretanto observou-se que até o momento não há trabalhos envolvendo a COVID-19 e bases de dados, como trazemos nesta pesquisa.

6 COVID 19: UM OLHAR SOBRE O IMPACTO PANDÊMICO

A pandemia da COVID-19 tem-se mostrado como um enorme desafio, com a necessidade de estabelecer novos rumos e novas formas de convivência, não precisamente melhor que a anterior, o quadro pandêmico acarretou o caos nos sistemas de saúde do mundo inteiro e provocou debates que impulsionaram um olhar especial em muitos aspectos sociais.

Artigos sobre as mudanças educacionais perante a pandemia vem sendo bastante produzidos, pois com as novas regras de restrições, como o uso de máscaras faciais, distanciamento físico e a não aglomeração por conta da transmissão dos vírus, forçou o interrompimento das aulas nas escolas e universidade fazendo assim com que os alunos tivessem que se adaptar a um novo quadro: aulas remotas. Sob o ponto de vista de Pokhrel e Chhetri (2021, p.2, tradução nossa):

O aprendizado online, a distância e a educação continuada se tornaram uma panaceia para essa pandemia global sem precedentes, apesar dos desafios impostos tanto aos educadores quanto aos alunos. A transição da aprendizagem presencial tradicional para a aprendizagem online pode ser uma experiência totalmente diferente para os alunos e educadores, à qual eles devem se adaptar com pouca ou nenhuma outra alternativa disponível.

Considera-se assim que a pandemia gerou um grande impacto não apenas na área da saúde, mas em outras áreas e aspectos da sociedade, houveram transições que acrescentaram algumas melhorias por meio da exigência do momento, tal como as implicações e as evoluções tecnológicas que realizam a conexão entre as pessoas assim como auxiliam no aceleração da criação das vacinas, fazendo com que um grande volume de dados fossem compartilhados e produzidos assim como traz a premissa: “ A comunidade científica respondeu rapidamente à pandemia do COVID-19, lançando mais de 125.000 artigos científicos relacionados ao COVID- 19 dentro de 10 meses após o primeiro caso confirmado, dos quais mais de 30.000 foram hospedados por servidores de pré-impressão” (Fraser *et al.*, 2021, p 1, tradução nossa).

Levando em consideração o impacto gerado pelo surto da doença nos diversos ramos, é pertinente conhecer melhor a doença e o vírus, assim como sobre as vacinas e dados produzidos acerca desse assunto, sendo esses os tópicos que discutiremos nessa seção.

6.1 ENTENDENDO A COVID-19

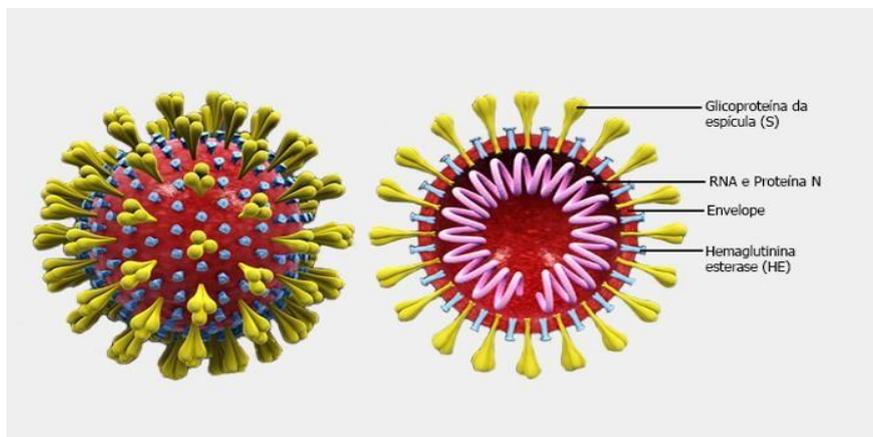
Os primeiros casos evidentes da doença COVID-19, como é conhecida, foram inicialmente identificados em dezembro de 2019, com um surto na China, especificamente em Wuhan, província de Hubei. A princípio as primeiras ocorrências eram ligadas apenas à pneumonia, até a realização de estudos sobre a doença, que logo se alastrou. Em relação às pesquisas sobre o entendimento do vírus revelam-se que:

A inoculação de amostras respiratórias em células epiteliais das vias aéreas humanas, linhagens celulares vera E6 e Huh7, levou ao isolamento de um novo vírus respiratório cuja análise do genoma mostrou ser um novo coronavírus relacionado ao SARS-CoV e, portanto, chamado de coronavírus da síndrome respiratória aguda grave 2 (SARS-CoV-2). O SARS-CoV-2 é um betacoronavírus pertencente ao subgênero *Sarbecovirus*. (CIOTTI, MARCO *et al.*, 2020, p. 365, tradução nossa).

Diante a isso, a proliferação do vírus provocou a pandemia, o que causou milhares de mortes e uma emergência de saúde global, conforme foi declarado pela Organização Mundial da Saúde (OMS) no dia 30 de janeiro de 2020.

Em sua estrutura o vírus contém em seu material genético RNA de fita simples positiva, envolvido em uma cápsula lipoproteica, contendo nesta estrutura uma proteína Spike ou proteína S unida a enzima ACE 2 (enzima de conversão de angiotensina tipo 2), diante disso o termo *coronavírus* originou da observação das projeções da membrana e do aspecto dos vírions de Cov, nos quais assemelha-se a uma coroa, partindo dessa análise tem-se que o termo coroa ou coroa em latim, como mostrado na Figura-05.(NOGUEIRA, 2020).

Figura 05 – Representação esquemática da estrutura do Coronavírus.



Fonte: Site da Sociedade Brasileira de Análises clínicas, 2020, online

Convém ressaltar que segundo Nogueira (2020, p. 119) “a análise genômica e filogenética completa do vírus indicam que este é do tipo betacoronavírus, mesmo subgênero do SARS que causa a Síndrome

Respiratória Aguda Grave. ” Além do mais, o RNA positivo não segmentado e envelopado está normalmente associado a síndromes entéricas e respiratórias em animais e humanos (SILVA *et al.*, 2021).

Inicialmente surgiu a teoria de que o vírus foi resultado de pesquisas laboratoriais, porém após algumas análises observou-se que o vírus é oriundo de uma transmutação de um vírus já conhecido pela alta contaminação de animais.

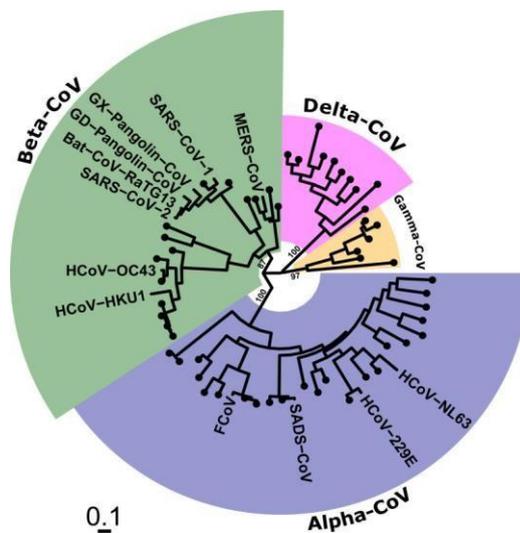
Os coronavírus pertencem à ordem Nidovirales e à família Coronaviridae. A subfamília Coronavirinae é composta pelos gêneros Alphacoronavirus e Betacoronavirus, cujos membros infectam mamíferos e Gammacoronavirus e Deltacoronavirus, os quais infectam tanto aves quanto mamíferos. O SARS-CoV-2 é um β -coronavírus (subgênero Sarbecovirus, Subfamília Orthocoronavirinae), possuindo como material genético RNA de sentido positivo não segmentado (DUARTE, 2020, p.3586)

Isso levou a crer, e posteriormente foi confirmado, que o vírus possui muitas variáveis e que é de fácil mutação, tendo em vista que sua estrutura genética sofreu mudanças para se adaptar ao hospedeiro, chegando assim à contaminação de humanos. Silvia *et al.* (2021, p. 3) traz em seu estudo que:

Tinham conhecimento de seis espécies de coronavírus que causam doenças humanas. Quatro deles causam sintomas comuns de resfriado. Entretanto os outros dois, de origem zoonótica, mais graves, relacionados com o surgimento de infecções com alto impacto na saúde pública devido a sua letalidade, são eles: coronavírus da síndrome respiratória aguda grave (SARS-CoV), o agente responsável pelos surtos em 2002 e 2003 na província de Guangdong, China, e o outro coronavírus da síndrome respiratória do Oriente Médio (MERS-CoV) responsável pelos graves surtos de doenças respiratórias em 2012 no Oriente Médio

Ou seja, já se tinham estudos sobre a ação desse vírus, os tipos e a sua contaminação, de modo que surtos do vírus já foram registrados em países do oriente médio anteriormente. Todavia, o que difere dos surtos anteriores ao surto acometido em 2020 é a alta letalidade, ocasionada pela mutação genética do vírus, esse no qual continua modificando-se, acarretando variáveis do vírus como evidenciado na Figura-6.

Figura 06 - Relações filogenéticas de corona vírus.



Fonte: Morens,2020, p.995

De acordo com a OMS alguns estudos apontam que a reprodução do vírus é em média R 0 1,4 a 2,5 segundo, de modo que pesquisas detectaram o vírus em superfícies, como por exemplo plástico, metal, ou até mesmo aço, Triggler *et al.* (2020) pontua que o SARS-CoV 2 possui meia- vida de 1,1 a 1,2h podendo essa medida se alterar dependendo do meio em que se instala (TRIGGLE CR *et al.*, 2020).

Apresentando como sintomas febre, tosse e dificuldade de respirar, que em casos graves podem evoluir para uma pneumonia com insuficiência respiratória aguda grave, podendo levar a morte, sobre os sintomas Ciotti (2020, p. 369, tradução nossa) trata que:

Pacientes com infecção por SARS-CoV-2 podem apresentar sintomas que variam de leves a graves, sendo grande parte da população portadores assintomáticos. Os sintomas mais comuns relatados incluem febre (83%), tosse (82%) e falta de ar (31%). Em pacientes com pneumonia, a radiografia de tórax geralmente mostra múltiplas manchas e opacidade em vidro fosco. Sintomas gastrointestinais como vômitos, diarreia e dor abdominal são descritos em 2- 10% dos pacientes com COVID-19, e em 10% dos pacientes, diarreia e náusea precedem o desenvolvimento de febre e sintomas respiratórios. Os pacientes com COVID-19 geralmente apresentam diminuição da contagem de linfócitos e eosinófilos, valores medianos mais baixos de hemoglobina, bem como aumentos nos leucócitos, na contagem de neutrófilos e nos níveis séricos de PCR, LDH, AST e ALT.

O vírus se propaga tanto por meios diretos (gotículas e transmissão entre humanos) como indiretas (objetos contaminados e contágio pelo ar), quando a pessoa infectada pelo vírus tosse, espirra ou fala tem-se que as gotículas contaminadas atingem em média 1 m, e permanecem suspensas no ar por até três horas (LOTFI, HAMBLIN e

REZAEI, 2020). Quanto a transmissão da COVID, Freitas, Napimoga e Donalísio (2020) apresenta que

Para a dimensão de transmissibilidade, a pontuação varia de 1 a 5, os indicadores são: taxa de ataque de sintomático em diferentes cenários; R0 (número reprodutivo básico); e o pico de percentual de consultas de síndrome gripal em prontos socorros. Por sua vez, para a dimensão da gravidade clínica, o escore varia de 1 a 7, as variáveis utilizadas são: letalidade; proporção de casos hospitalizados entre os atendidos (proporção de hospitalização); e taxa de mortalidade hospitalar, considerando-se apenas casos de influenza, na situação atual da Covid-19 (FREITAS; NAPIMOGA; DONALÍSIO, 2020, p. 1).

Logo faz com que a população mundial adote diferentes formas de convivência, como por exemplo, o uso de máscaras protetoras, uso de álcool em gel e distanciamento social são algumas das medidas protetivas contra a doença.

Ao entender a tipologia, a forma de transmissão e como ela age, pôde-se então realizar estudos relacionados à prevenção do vírus, estudando viáveis para que o grau de impacto causado pelo vírus se torne menor, algo que só é possível por intermédio de vacina.

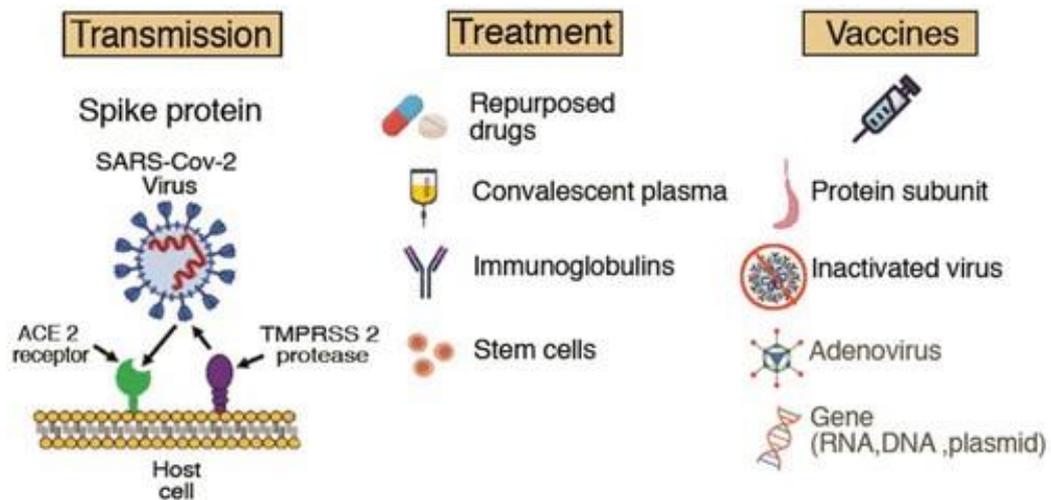
6.2 VACINAÇÃO E DADOS CIENTÍFICOS

Como foi visto anteriormente, não se sabia muito sobre o SARS-CoV-2 e o fato de os sintomas serem bastante semelhantes ao de outras doenças foi o que inicialmente auxiliou com o tratamento contra o vírus.

Essas semelhanças forneceram várias oportunidades para tratar pacientes com COVID-19 usando abordagens clínicas comprovadamente eficazes contra a SARS. É importante ressaltar que a identificação de semelhanças em como o SARS-CoV e o SARS-CoV-2 acessam o host, replicam, e desencadear condições patológicas com risco de vida revelaram oportunidades para redirecionar medicamentos que se mostraram eficazes contra a SARS (SALIAN et al., 2021, p.745 tradução nossa).

Inicialmente observou-se como o vírus se comportava, sua combinação genética ao entrar em contato com o receptor, então pensou-se em tratamento como remédios usados para tratamento de doenças com sintomas similares, assim como foram realizados estudos com plasma convalescentes, imunoglobulinas e até mesmo células tronco para auxiliarem no combate à doença, até chegarem nas vacinas por meio de pesquisas com subunidade de proteína, inoculação do vírus inativo, adenovírus e gene (RNA, DNA, plasmídeo) com como apresenta a Figura -07:

Figura 07- Transmissão, tratamento e vacinas contra a covid-19



Fonte: Salian *et al.*, 2021,p.745

Ao observar-se que para combater o surto pelo vírus seria necessário entender e estudar sobre sua ação e as formas de inativá-lo, desta maneira chegou-se à conclusão e a criação de uma variedade de vacinas.

A maioria das vacinas estudadas para a COVID-19 visa induzir anticorpos neutralizantes contra as subunidades virais, a maior parte delas tendo como alvo a região RBD (do inglês domínio de ligação do receptor) da proteína mais conservada do vírus, a Spike (S), impedindo assim, a captação do vírus pelo receptor ACE2 (enzima conversora da angiotensina 2) humano (LIMA, ALMEIDA E KFOURI, p.22, 2021).

O impacto causado pela pandemia e o auxílio das novas tecnologias, fez com que o desenvolvimento das vacinas se tornasse mais rápido, em menos de um ano já se tinham testes clínicos em humanos, de modo que no ano anterior iniciaram as vacinações na população, um feito com rapidez sem precedente.

Muitos estudos foram realizados com intuito de descobrir formas para fabricar o imunizante, tendo como aliado as TDICS, os pesquisadores conseguiram realizar o feito de criar um variado número de vacinas em pouco tempo. Lima *et al.* (2020) traz

Cerca de 120 equipes de pesquisa em todo o mundo estão estudando o vírus por vários ângulos para a elaboração das vacinas. Programas de desenvolvimento de vacinas são continuamente atualizados pela OMS, com projetos identificados a partir de fontes públicas e privadas (LIMA *et al.*, 2020, p. 1).

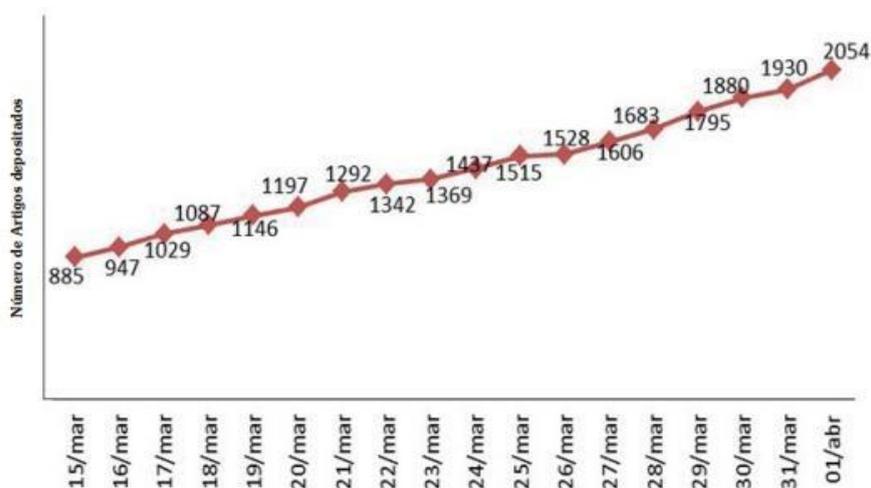
Diante disso, observa-se que com esses estudos há uma geração de dados acerca da vacina, de modo que com a análise e a interação entre os dados dos estudos tem-se a identificação de diferentes quadros clínicos, tratamentos, prevenções e criação de imunizantes como a vacina. Mediante a isso Chiavegatto (2015) reconhece

Não se trata de um problema de fácil solução, já que a quantidade de dados usada

pelas pesquisas aumenta a cada ano. Na metade do século passado, encontrar os parâmetros de uma regressão linear com 500 observações era uma tarefa que levava alguns dias. Hoje, são necessários também alguns dias para rodar modelos bayesianos com centenas de milhares de observações (CHIAVEGATTO, 2015, p. 326).

Nesse panorama tem-se hoje um crescimento na produção científica, de modo que estudos realizados com a finalidade de compreender o funcionamento da COVID-19, assim como seu impacto em diferentes contextos e quadros clínicos vem se mostrando cada vez mais crescente no âmbito acadêmico, assim como apontado por Diniz *et al.* (2020). Em sua pesquisa e evidenciado na Figura-08.

Figura 08- Curva de crescimento da produção científica acumulada sobre COVID-19 no PubMed em 18 dias, de 15 de março a 1º de abril de 2020



Fonte: Diniz et al, 2020, p.362

O crescimento do compartilhamento desses trabalhos, reflete também na expansão de dados, os quais eles produzem, ao observar que, no cenário pandêmico vivido atualmente, os surgimentos de novas práticas de pesquisa e a intensificação da produção científica estão cada vez mais presentes na cultura e no âmbito acadêmico.

Em razão disso, nota-se que não apenas a produção de trabalhos acerca do entendimento da dinâmica da doença e do vírus estão em alta, mas produções científicas, como patentes e artigos, trabalhos pré-impresos (pré-print), sobre novos tratamentos, produtos assim como em relação à prevenção estão em progresso, de modo a se ter um enfoque maior na criação e na produção de vacinas que combatam a doença.

Isso faz crer na importância da gestão de dados na área da saúde, de modo a trazer um melhor planejamento para o uso, interoperabilidade, localização e acessibilidade dos dados, de modo que os princípios FAIR aplicados a dados e os metadados voltados para as vacinas contra COVID

irão auxiliar os profissionais a distinguir quais dados e metadados estão dentro dos padrões de gestão, o que os faz ser confiáveis para uso.

7 PESQUISA DE CAMPO

Pautando-se nos objetivos específicos concernentes à: verificação da aderência dos princípios FAIR em repositórios de dados, observando-se os protocolos de elegibilidade, e metadados referentes às vacinas contra a COVID-19; experimentação do Fair Data BR-UFPB. Para tanto, efetuou-se a empiria no mecanismo de busca Dataset Search, inicialmente, foi feita uma busca no mecanismo *Dataset Search*, ligado à empresa Google. Ele se configura como um mecanismo de pesquisa para um conjunto de dados no qual oferece filtros que auxiliam em uma busca mais direta e simplificada dentro de milhares de repositórios de acesso aberto, hospedados na web, como evidenciado na interface mostrada na Figura-9

Figura 09 – Dataset Search



Fonte: datasetsearch.research.google.com

Conforme já mencionamos na seção 2, referente à metodologia, utilizou-se como estratégia de busca, o termo “**VACCINE AND COVID**” e protocolos para a identificação e elegibilidade dos dados a saber.

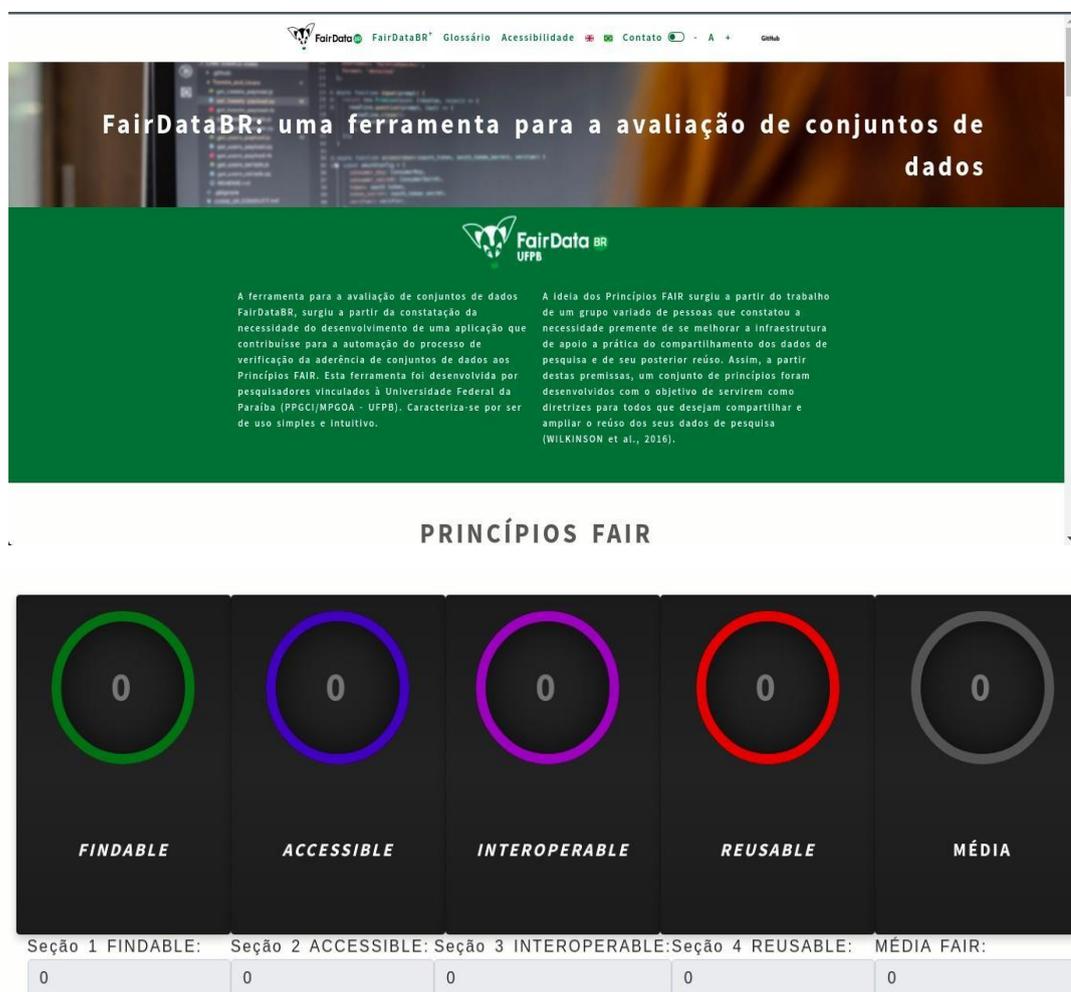
- a. Tema (Ciências da vida, em que se insere a área da saúde, e Ciências Sociais). Embora a ferramenta *Dataset Search* não separe Ciências Sociais e Ciências Sociais Aplicadas, ainda assim, tomamos a decisão de incluir aqui a Ciência da Informação.
- b. Direitos de uso, tendo dados com uso comercial permitido;
- c. Formato de download, no qual engloba todos os tipos de formatos (tabular, documento, imagem, texto, arquivo e outros) e;
- d. Cobertura dos dados de pesquisa: janeiro a agosto de 2021.

Tendo como resultado 122 conjuntos de dados, como mostrado na

figura 10. Em seguida, esses conjuntos foram submetidos à ferramenta Fair Data BR, ferramenta produzida por pesquisadores: Dias, Guilherme Ataíde; Moro, Anderson Rodrigues; Menezes, Ana Karoline França de; Simões, Anderson Rafael Castro; Campos, Arthur Ferreira; Silva, Lucas Henrique Alves da; Anjos, Renata Lemos dos; Bezerra, Tanise Dantas, vinculados à Universidade Federal da Paraíba (PPGCI/MPGOA - UFPB), sob o endereço: wrco.ufpb.br/fair/index.

Tal ferramenta possui um conjunto de perguntas para cada princípio e de acordo com as respostas obtém-se o parâmetro de médias, dividido entre os 4 princípios que resulta na média final: MÉDIA FAIR, como representado na Figura-10.

Figura 10- Ferramenta Fair Data BR



Fonte: wrco.ufpb.br/fair/index.html

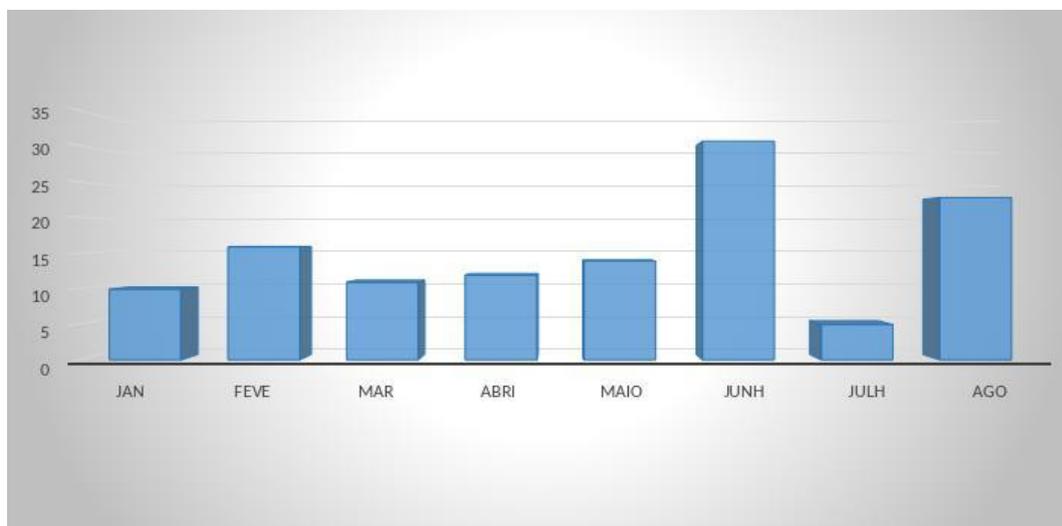
Em decorrência disso, na próxima seção tratamos sobre os conjuntos de dados obtidos nesta pesquisa e a aderência destes à ferramenta Fair Data BR, assim como discutimos sobre esses resultados com outros autores.

7.1. ANÁLISE DOS DADOS E DISCUSSÕES DOS RESULTADOS

Tendo extraídos 122 conjuntos de dados, o primeiro passo foi observar detalhadamente, a data de publicação, entre os meses de janeiro a agosto de 2021, a área, Ciência da Vida ou Ciências Sociais. Também se realizou uma análise da tipologia dos conjuntos de dados, quanto à origem, a natureza e quanto às fases da pesquisa, assim como disposto por Sayão e Sales (2015), conforme evidenciado no referencial teórico.

Para tanto, ao examinar o quesito da temporalidade dos conjuntos de dados coletados, levando em consideração as últimas publicações datadas entre os meses de cobertura da pesquisa, chegou-se ao seguinte Gráfico-03.

Gráfico 03 - Gráfico de temporalidade



Fonte: Dados da pesquisa, 2021

Pode-se notar que o mês de junho foi o que teve mais publicações, seguido pelos meses de agosto e fevereiro. Logo, entende-se que com o despontar da pandemia em muitos estudos sobre o vírus foram se desenvolvendo para se chegar a um tratamento ou a vacinas, de modo que com o acesso aberto, houve uma enxurrada de dados o que com as redes cooperativas resultaram em uma variedade de vacinas contra o vírus. Tais resultados encontram-se com o discurso de Andreadakis *et al.* (2020, p. 305, tradução nossa) ao afirmar que,

Em 8 de abril de 2020, a vacina global R&B landscape da COVID-19 incluiu 115 vacinas candidatas, das quais 78 são confirmados como ativos e 37 não confirmados (status de desenvolvimento não pode ser determinado na forma de publicidade disponível ou fontes de informação proprietárias).

Constata-se que nos primeiros períodos do ano de 2021, com o

início das vacinações, surgiram assim muitos trabalhos acerca dos resultados desse assunto, como por exemplo um dos conjuntos de dados que trata sobre o rastreamento de vacinas da Covid-19, evidenciado pelo Centro Europeu de Prevenção e Controle de Doenças (ECDC), de modo que em sua descrição tem-se:

[...] fornece uma visão geral do progresso na implantação de vacinas COVID-19 em adultos (com 18 anos ou mais) nos Estados-Membros da UE/EEE. Apresenta o número de doses de vacinas distribuídas pelos fabricantes a cada Estado-Membro para as diferentes vacinas autorizadas para utilização na UE e o número de primeiras, segundas ou não especificadas doses administradas a indivíduos adultos. (ECDC, online, tradução nossa)

Outros conjuntos abordam sobre a eficácia e a segurança das vacinas, a necessidade de outras doses, a hesitação da população ao tomar essas vacinas, assim como trabalhos os quais englobam a disseminação da informação, e até mesmo a demografia de cidades e países envolvendo os fluxos de vacinação.

De modo que as patologias ligadas à coronavírus, vem surgindo antes mesmo do surto em 2019, como bem evidenciam Vargas e Lawall (2020, p.16, tradução nossa)

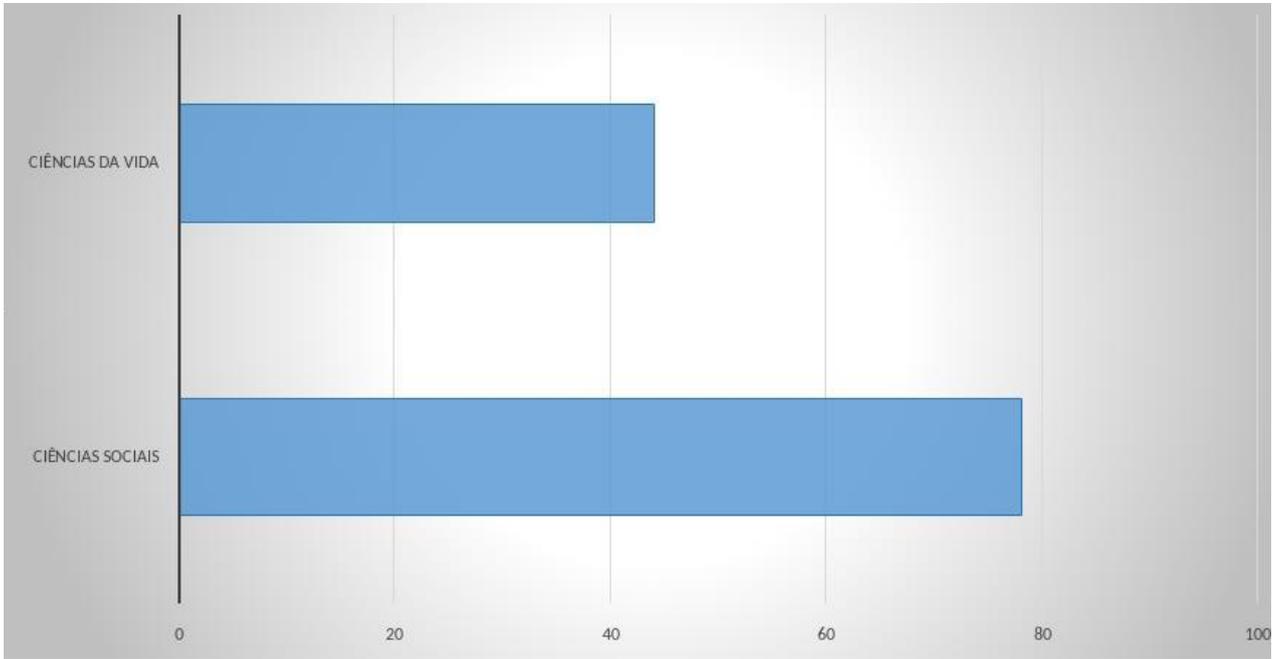
Sars-cov-1 agente da pandemia da Sars com surto na Ásia Continental, entre os anos de 2002-2003; o Mers-cov, causador da Mers no ano de 2012, com surto na Arábia Saudita. As outras quatro espécies de coronavírus HKU1, NL63, OC43 e 229E estão associadas a doenças com sintomatologia leve.

Não é demais lembrar que, essa pesquisa possui como foco empírico, a aderência dos princípios FAIR no contexto da literatura envolvendo as vacinas contra a covid-19. Também conhecida como “Síndrome Respiratória Aguda Severa” ou (Sars-cov-2), e as vacinas contra a doença. Ao retratar sobre as análises das produções científicas tem-se uma alta nos estudos sobre a COVID-19 desde 2019, que eram mais voltados para as Ciências da Vida, e poucos com foco em Ciências Sociais. No trabalho desenvolvido por, Vasconcelos e Nascimento (2020, p.3) “Foram encontrados 5.570 artigos sobre COVID-19 na plataforma Scopus, em que apenas 10 foram publicados no ano de 2019 e os outros 5.560 no ano de 2020”.

Ainda nessa linha de pensamento, Vasconcelos e Nascimento (2020, p.3), ressaltam que “as cinco áreas que mais publicaram sobre COVID-19 foram: Medicina, Bioquímica, Genética e Biologia Molecular, Ciências Sociais, Imunologia e Microbiologia, e Ciência ambiental. ” Esses dados não poderiam ser diferentes, afinal, a pervasividade da COVID-19, concretiza-se a partir de 2020 em que os

artigos geraram muitos dados e informações sobre o assunto. Ao que corrobora com os resultados encontrados em nossa pesquisa (Gráfico-04) acerca das áreas em que os conjuntos de dados pertencem.

Gráfico 04 - Média por área



Fonte: Dados da pesquisa, 2021.

Observando-se o Gráfico-04, percebe-se que a predominância sobre dados de pesquisas envolvendo as Ciências da Vida/ da Saúde, destacam-se entre eles aqueles relativos à COVID, às vacinas, à distribuição das vacinas, à porcentagem das vacinações entre outros.

Figura 11 -Nuvem de tag dos dados da área da Vida.



Fonte: dados da pesquisa,2021

Dados circundam as vacinações, as características sociais e demográficas, os *tweets* envolvendo as vacinas contra o covid-19 também são alguns dos assuntos de maior alcance encontrados dentre os dados extraídos. Pôde-se notar então que essas são as temáticas mais estudadas nesse período, ao se falar sobre a hesitação das vacinas investigadas em alguns países tem-se, segundo Machingaidze e Wiysonge (2021, p.1338, tradução nossa).

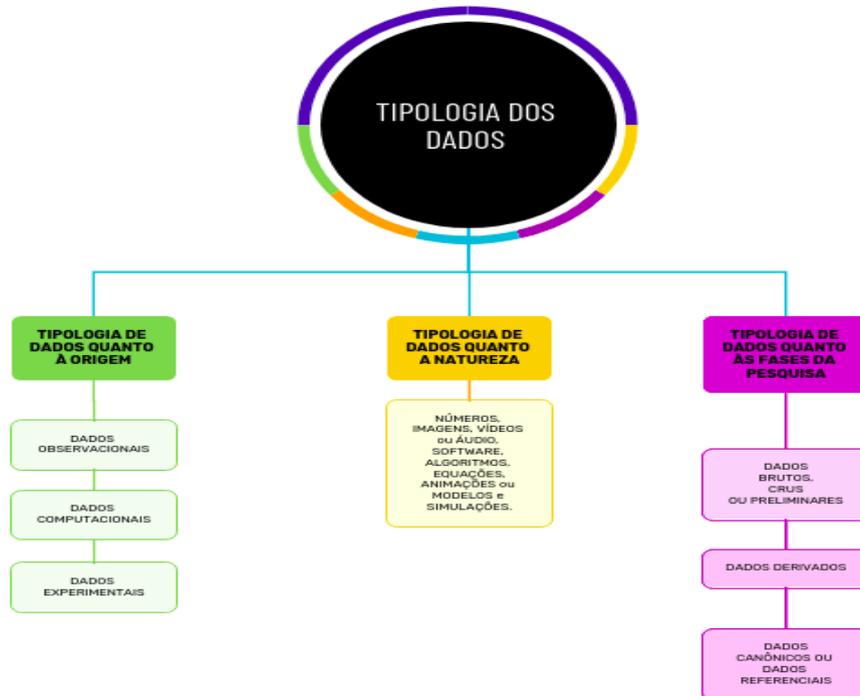
Promover a adoção de vacinas (particularmente aquelas contra a COVID-19) exigirá entender se as pessoas estão dispostas a ser vacinadas, as razões pelas quais estão dispostas ou não a fazê-lo e as fontes de informação mais confiáveis em sua tomada de decisão. Solis Arce et ai. investigaram essas questões usando um conjunto comum de itens de pesquisa implantados entre junho de 2020 e janeiro de 2021, em 15 estudos realizados na África, Sul da Ásia, América Latina, Rússia e Estados Unidos

Nesse contexto observa-se que apesar de termos um elevado número de dados sendo compartilhados, ainda há uma área se sobressaindo, deste modo podemos nos remeter a *long tail* ou cauda longa abordado anteriormente e apresentado por Córdula e Araújo (2019) a cauda longa observa “o volume de dados é distribuído de forma desigual pelos vários campos da ciência”.

Na mesma linha de pensamento Sayão e Sales (2018, p. 4184) trazem que “Os dados da cauda longa representam a maior parcela de dados produzida pela ciência e constituem um território de constante criatividade e inovação que precisam ser revelados, integrados e compartilhados.” Posto isso, fica explícito que ainda há a necessidade de maior implementação da gestão dos dados em algumas áreas.

Dada a relevância da pesquisa, outro aspecto foi analisado perante os dados obtidos, de modo a seguir as tipologias trazidas por Sayão e Sales (2015) tem-se: os tipos de dados, fase (dados brutos, dados referenciais e dados derivados), a origem (observacionais, experimentais e computacionais), quanto a sua natureza (mapas, números, textos etc.). Na Figura-13, ilustram-se esses achados.

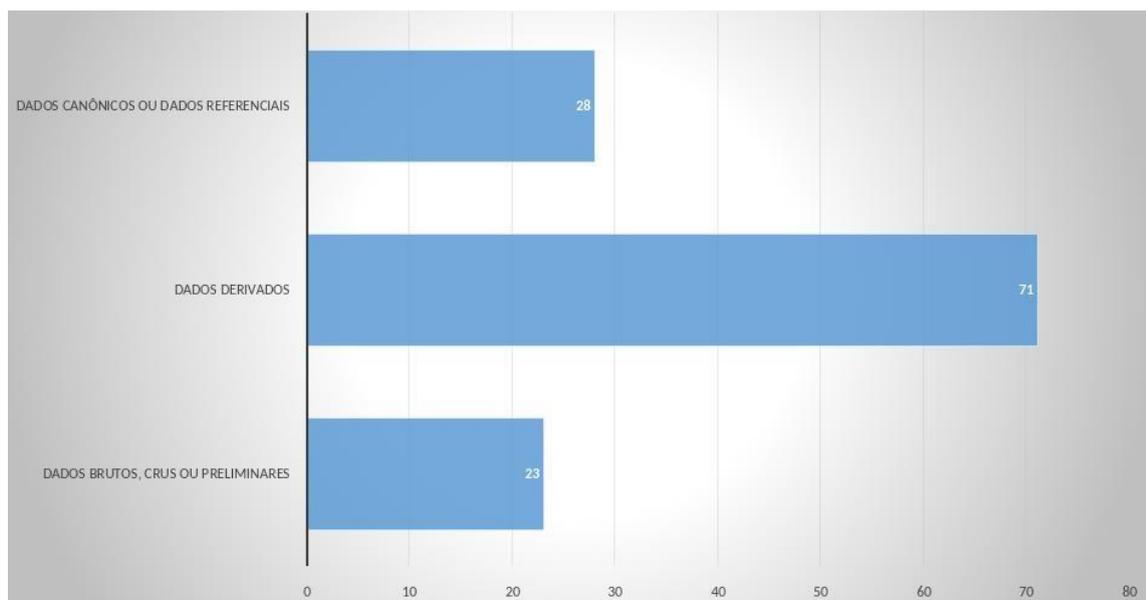
Figura 13 - Tipologia os dados de pesquisa



Fonte: Elaborado pela autora e adaptado de Sayão e Sales (2015, p.8)

Acresce que segundo Silva (2019, p. 27) “os dados podem ser classificados mediante diferentes critérios: pelo procedimento de coleta, por seu caráter primário ou secundário, pelo grau de estruturação, pelo seu nível de abertura e pelo seu formato”, com isso tem-se que a nossa pesquisa, no que diz respeito à tipologia dos dados por fases, registou-se o Gráfico-05

Gráfico 05- Tipologia de dados (FASE)



Fonte: Dados da pesquisa, 2021.

Fica demonstrado que uma grande parte dos dados são derivados, tendo assim dados oriundos de um processamento ou combinação de outros dados brutos, de modo que em sua maioria, os dados se mostraram continuações ou combinações entre outros dados por exemplo dados secundários de tweets brasileiros sobre COVID-19, os quais são dados oferecidos por entidades diferentes trazendo um panorama geral criando mais um conjunto de dados.

Esses conjuntos de dados contêm dados secundários extraídos de tweets brasileiros sobre a pandemia de COVID-19. Esses dados foram gerados entre março de 2020 e maio de 2021. Devido às restrições do Twitter, não foi possível compartilhar os dados brutos. Além disso, respeitando a privacidade dos usuários e de acordo com as recomendações do Comitê Brasileiro de Ética em Pesquisa, nenhum dos tweets ID pode ser compartilhado aqui. Assim, um conjunto de scripts foi aplicado em dados brutos para extrair informações úteis e criar esses conjuntos de dados para futuras pesquisas. Os arquivos iniciados com o prefixo superior contêm as palavras mais citadas por dia em três categorias: tweets gerais, tweets relacionados a vacinas e tweets de contas verificadas. Já os arquivos que começam com o prefixo assunto contêm a contagem diária de menções de acordo com as seguintes categorias: sintomas, medicamentos, vacinas, marcas ou fabricantes de vacinas e a contagem por dia. Como os dados estão associados à data de postagem, estudos podem ser desenvolvidos considerando o aspecto temporal para comparar a percepção dos usuários sobre determinado assunto ao longo do tempo. É importante notar que algumas questões se tornaram mais importantes ao longo do tempo, principalmente em relação às vacinas. (XAVIER, ACOSTA E SARAIVA, 2021, ONLINE, tradução nossa).

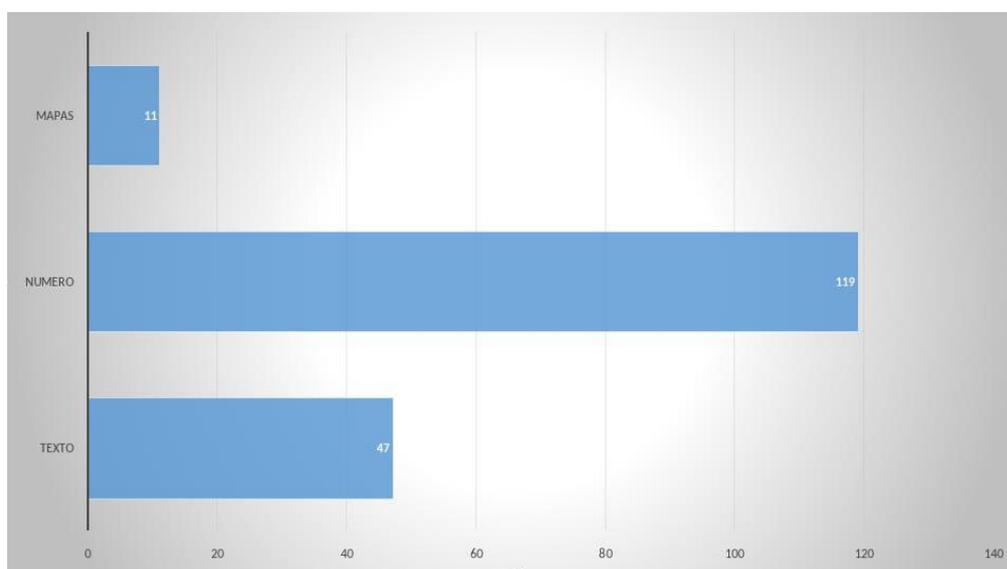
De forma a corroborar com a afirmativa Green, Macdonald e Rice (2009, p. 6, tradução nossa) apresentam que:

Dados derivados: resultantes do tratamento ou da combinação de matérias brutas ou outros dados (se for necessário respeitar os direitos dos proprietários dos dados brutos) Dados canônicos ou de referência: por exemplo, sequências genéticas, estruturas químicas etc.

Observou-se que os dados brutos ou crus, vieram de estudos que

não tiveram influência ou combinação com outros, exemplificando o conjunto de dados que traz sobre número acumulado e porcentagem de pessoas que receberam uma vacina COVID-19 no Canadá por número de doses, faixa etária, sexo e jurisdição. Os dados Canônicos ou referenciais pouco encontrados, são mais sólidos. Citamos como exemplo, o conjunto de dados sobre a Linfadenopatia após a vacina BNT162b2 Covid-19: achados preliminares de ultrassom. No Gráfico-06, apresentamos outra tipologia de dados analisadas quanto a sua natureza (mapas, números e textos verbais).

Gráfico 06 – Tipologia dos dados (Natureza)



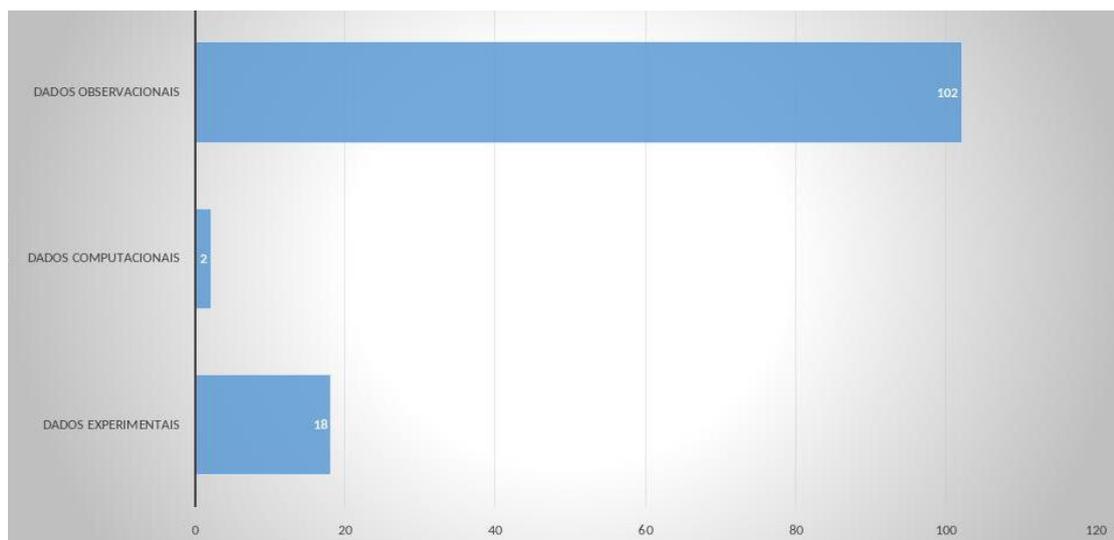
Fonte: Dados da Pesquisa, 2021.

Quanto à natureza dos dados, observou-se que em sua maioria eles são expressos por números, textos verbais e mapas, assim como gráficos. De acordo com Silva (2019, p. 21),

São as informações registradas ou produzidas através de qualquer forma ou meio durante o decurso de uma pesquisa. Podem ser de tipos numérico, descritivo ou visual e reproduzir-se em formato de papel (incluindo notas de pesquisa em cadernos, fotografias, etc.) ou digital. O conceito de dados científicos também faz referência às distintas ferramentas como protocolos, códigos numéricos, gráficos e tabelas que são necessárias para recolher e organizar os dados, tanto em trabalhos de campo quanto em laboratório.

Em decorrência da publicação de dados, analisou-se que em alguns casos houve o uso de dados numéricos e textuais, assim como em outros casos foram utilizados gráficos e mapas. Por fim, no Gráfico-07, expõem-se a última tipologia observada em relação à origem dos dados capturados na pesquisa.

Gráfico 07- Tipo de Dado (Origem)



Fonte: Dados da Pesquisa,2021.

Ao se analisar o Gráfico- 07- tem-se a predominância dos dados observacionais, os quais são voltados para ponderações diretas, que podem ser associadas a lugares e tempo específicos, de modo que muitos estudos tratam sobre as hesitações e atitudes tomadas perante a vacinação da COVID-19.

Em seguida, temos os dados experimentais, esses possuem uma origem laboratorial, são provenientes de pesquisas com experimentos e ensaios clínicos, como por exemplo dados que tragam informações sobre testes de vacinas ou tratamentos realizados em laboratório, portanto, todos os achados relacionam-se, com as Ciências da Vida, naturalmente.

Por fim, evidencia-se os dados computacionais, que apesar de terem sido poucos, ainda assim, são de grande valia para outras pesquisas. Segundo Silva (2019, p.27) “são produtos da execução de modelos de ordenação, simulações ou fluxo de trabalho”.

De modo geral, pode-se dizer que houve uma extração de dados bastante variada e rica com relação aos termos utilizados, logo, espera-se que não sejam apenas variados, mas encontrados, acessíveis, interoperáveis e reutilizáveis.

7.2 RESULTADOS E DISCUSSÕES SOBRE AS MÉDIAS FAIR

Para adentrar sobre os princípios FAIR é importante trazer em pauta uma visão, a qual reforça o que foi evidenciado no referencial da pesquisa, sobre o compartilhamento e a gestão de dados, desta forma,

segundo Lili (2018, p. 774, tradução nossa)

Como alma das atividades de investigação científica, os dados científicos não são apenas o ponto de partida para estimular a inovação da investigação científica, mas também uma parte indispensável das ricas realizações das atividades de investigação científica. A gestão e partilha de dados científicos é realizada a partir dos dois aspectos da “orientação ativa e moderada da política de dados científicos” e da “prática abrangente e meticulosa de gestão de dados científicos”.

Diante da afirmativa torna-se necessário evidenciar que nem todos os dados estão disponibilizados como dados abertos, alguns por questões contratuais permanecem “restritos ou fechados” e outros encontram-se em período de embargo ou disponíveis por meio de solicitação direta de sua autoria.

Em razão disso, entende-se que o compartilhamento de dados, independentemente do campo de investigação, e dos grupos de interesse, é visto como uma prática benéfica de maneira a fazer com que as pesquisas científicas sejam impulsionadas ao que traz vantagens para que haja uma promoção maior da ciência, redução quanto à duplicação de esforços e aumento nas produções (LILI, 2021).

Ao abordar sobre os dados abertos Royal Society, (2012, p.7, tradução nossa) argumenta que

Perceber os benefícios dos dados abertos requer uma comunicação eficaz através de uma abertura mais inteligente: os dados devem estar acessíveis e prontamente localizados; eles devem ser inteligíveis para aqueles que desejam examiná-los; os dados devem ser avaliados para que possam ser feitos julgamentos sobre sua confiabilidade e a competência daqueles que os criaram; e eles devem ser utilizáveis pelos outros. Para que os dados atendam a esses requisitos, eles devem ser suportados por metadados explicativos (dados sobre dados).

Refletindo-se acerca do compartilhamento de dados percebe-se a relevância da gestão dos dados científicos que conforme Costa (2017, p. 148) trata-se do “[...] conjunto de atividades intrínsecas ao processo de tratamento técnico (curadoria), armazenamento, recuperação, disseminação e preservação dos dados coletados pela e-science”.

Dada a importância de se ter o compartilhamento e a gestão dos dados científicos, é notório a inserção de políticas que envolvam as práticas da gestão de dados, quanto a isso segundo Veiga et al (2019, p. 227)

Um Plano de Gestão de Dados (PGD) se faz necessário por ser considerado um instrumento das boas práticas de gestão possibilitando que os dados sejam acessados, preservados e reusados em novas pesquisas. Outro aspecto importante a ser considerado é que o PGD possa ser legível e acionável por máquina, permitindo a integração automatizada com outros sistemas de informação, facilitando o preenchimento e otimizando o tempo do pesquisador.

Partindo dessa premissa, fica evidente a importância de se ter um

plano de gestão de dados, pois para que se tenham dados encontrados, reutilizáveis, interoperáveis e acessíveis é necessário que ele siga um padrão para uma maior comodidade entre o autor e o usuário.

Ao trazer em questão a encontrabilidade, acesso, interoperabilidade e reuso tem-se em mente os princípios FAIR, em que segundo Moura, Aragão e Soares (2018, p. 124).

Estes princípios ajudam a maximizar o valor adquirido nas publicações acadêmicas contemporâneas em formato digital. Em resumo, todos os objetos de pesquisa digitais dos dados para a análise de conduta acadêmica, também devem aplicar os princípios FAIR, pois todos os componentes do processo da pesquisa deverão estar disponíveis para garantir a transparência, reprodução e reuso dos dados.

Em virtude disso, constata-se que esses princípios ao serem utilizados trazem maior adequação no quesito do uso e reuso dos dados de pesquisa científica, ao observar os componentes que cada princípio aborda pode-se obter um filtro dos dados que possuem ou não as características pertinentes para serem compartilhadas. Da perspectiva de outros autores tem-se que para Cruz-Riascos, Hott e Rezende (2018, p.20)

[...] os Princípios FAIR são aceitos como diretrizes essenciais dentro da área científica e são considerados como uma estrutura para a condução correta e responsável da ciência. Estes princípios devem ser aplicados a todas as pesquisas que utilizam dados digitais e/ou criação de software e, portanto, incluídos em todos os planos de dados de pesquisa.

Criados com o propósito de propor características fundamentais para os ambientes digitais de publicação de dados científicos, os princípios evidenciam o uso de metadados, dos identificadores persistentes, das autorizações, licenças para publicação e acessibilidade, como também o uso de vocabulários compartilhados e ontologias para o acesso em geral aos dados.

Em concordância com Dias, Anjos e Simões (2021, p. 98) “[...] Tais princípios incentivam, dentre outras práticas, o uso de metadados que, quando devidamente empregados, podem contribuir para aumentar a encontrabilidade, acesso, interoperabilidade e o reuso dos diferentes conjuntos de dados”.

Considerando assim o aumento da utilização dos princípios tem-se o surgimento de métricas que avaliam se há ou não a utilização desses instrumentos, diante disso Macario *et al.* (2021, p. 54) traz que

Aplicações e métricas para adoção concreta dos princípios FAIR têm sido desenvolvidas em todo o mundo, por meio da adoção de padrões, metadados, vocabulários controlados, ontologias e identificadores persistentes que proporcionam significado preciso aos dados e aos demais objetos a eles

vinculados.

Segundo pesquisas recentes muitos autores se voltam para a realização de métricas que auxiliam na incorporação dos princípios aos dados, em face disso Wilkinson et al (2018, online) traz alguns pontos importantes a serem observados pelas métricas, como espelha o Quadro-04.

Quadro 04- Requisitos para as métricas

REQUISITOS PARA AS MÉTRICAS
➤ As métricas devem abordar a multidimensionalidade dos princípios FAIR e abranger todos os tipos de objetos digitais.
➤ As métricas universais podem ser complementadas por métricas adicionais específicas de recursos que refletem as expectativas de comunidades específicas.
➤ As próprias métricas e quaisquer resultados decorrentes de sua aplicação devem ser JUSTOS.
➤ Padrões abertos em torno das métricas devem promover um ecossistema vibrante de ferramentas de avaliação de FAIRness.
➤ Várias abordagens para a avaliação FAIR devem ser habilitadas (por exemplo, autoavaliação, forças-tarefa, crowdsourcing, automatizado), no entanto, a capacidade de escalar as avaliações FAIRness para bilhões, senão trilhões de diversos objetos digitais, é fundamental.
➤ As avaliações de FAIRness devem ser mantidas atualizadas, e todas as avaliações devem ser versionadas, ter um carimbo de data/hora e ser publicamente acessíveis.
➤ As avaliações FAIRness apresentadas como uma simples visualização, serão uma modalidade poderosa para informar os usuários e orientar o trabalho dos produtores de recursos digitais.
➤ O processo de avaliação e a avaliação de FAIRness resultante devem ser concebidos e divulgados de forma a incentivar positivamente os fornecedores de recursos digitais; ou seja, eles devem ver o processo como justo e imparcial e, além disso, devem se beneficiar dessas avaliações e usá-las como uma oportunidade para identificar áreas de melhoria.
➤ A governança sobre as métricas, e os mecanismos para valia-las, será necessária para permitir sua evolução cuidadosa e abordar desacordos válidos.

Fonte: Wilkinson *et al.* (2018, online)

Ao refletir acerca desses requisitos pode-se constatar que eles

possuem uma interação com as propostas apresentadas nos princípios, de forma que cada tópico contempla uma preocupação tanto quanto aos resultados dos mecanismos serem justos e condizem com o cada conjunto de dados traz quanto a sua identificação, acesso, interoperabilidade e reuso, quanto a forma como essas métricas realizam suas avaliações. Por fim o autor conclui, ao falar sobre a FAIR que

Aqui relatamos a estrutura que desenvolvemos, que abrange a primeira interação de um conjunto básico de indicadores de FAIRness que podem ser medidos objetivamente por um processo semi automatizado e um modelo que pode ser seguido dentro de domínios acadêmicos individuais para derivar métricas que avaliam aspectos FAIR importantes para eles (WILKINSON et al, 2018, online) .

Presume-se então que as ferramentas sigam não apenas as premissas trazidas nos princípios, mas também tenham em mente as variadas condições em que os dados se encontram, posto isso Henning (2018, p. 9) diz que “[...] Para amenizar interpretações ambíguas sobre o entendimento dos princípios FAIR e monitorar o nível de FAIRness dos recursos digitais, algumas iniciativas vêm desenvolvendo, de forma colaborativa, métricas que descrevem características, atributos e comportamentos. ”

Consoante a isso, observa-se a importância das métricas serem estabelecidas e utilizadas, de modo que com o aumento do compartilhamento de dados, assim como o uso e reuso desses dados tem-se assim uma necessidade de garantir a segurança quanto aos resultados das pesquisas. Visto assim a utilização de ferramentas de avaliação que empregam os princípios FAIR como base são percebidas como norteadoras para se ter dados seguros e exatos.

Figura 14- Infográfico FAIR



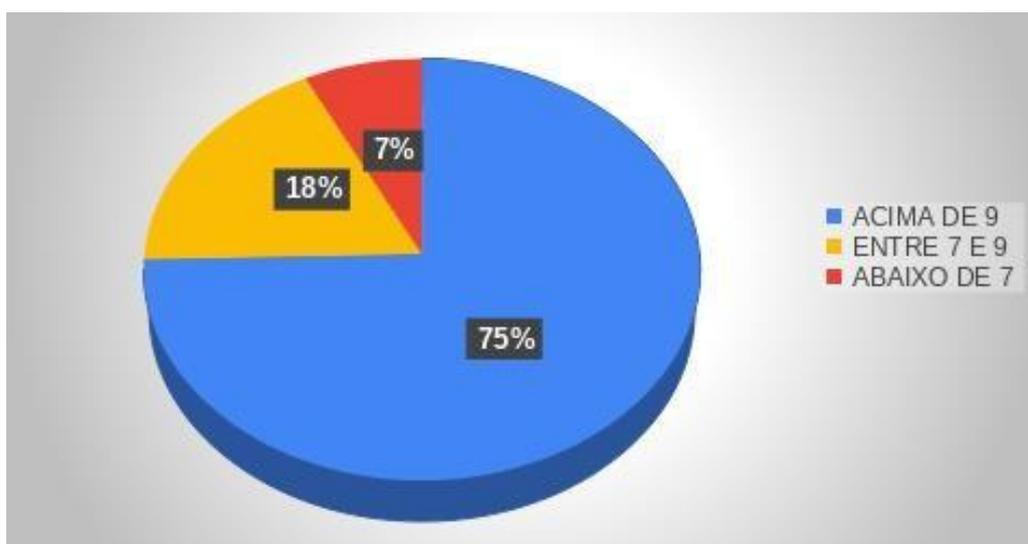
Fonte: Elaborado pela autora e adaptado de Henning, 2018.

A vista disso tem-se a Figura-15 que evidencia no infográfico FAIR e apresenta algumas observações acerca dos países que os seguem, por quem são utilizados, o que são os princípios e as métricas.

Percebendo-se esses aspectos, em nossa pesquisa utilizou-se a Ferramenta FAIR data BR, para fazer uma análise dos dados de pesquisa. Para tanto, foi aplicada uma métrica qualitativa baseada em perguntas norteadoras que envolvem intrinsecamente cada princípio e que se dispôs a realizar uma avaliação nos dados relacionados ao surto de coronavírus e acerca das vacinas contra a doença.

Como resultado foram apurados e analisados 122 conjuntos de dados, nos quais encontram-se enunciados no Gráfico-08.

Gráfico 08- MÉDIA FAIR



Fonte: Dados da pesquisa, 2021

Assim, constataram-se que boa parte 75% dos dados estavam em consonância com os padrões exigidos pela ferramenta FAIR Data BR e se apresentavam com a da média FAIR acima de 9, outros, com médias entre 7 e 8 resultam em 18% e aqueles com médias abaixo de 7 correspondem a 7%, como ilustrado no gráfico abaixo.

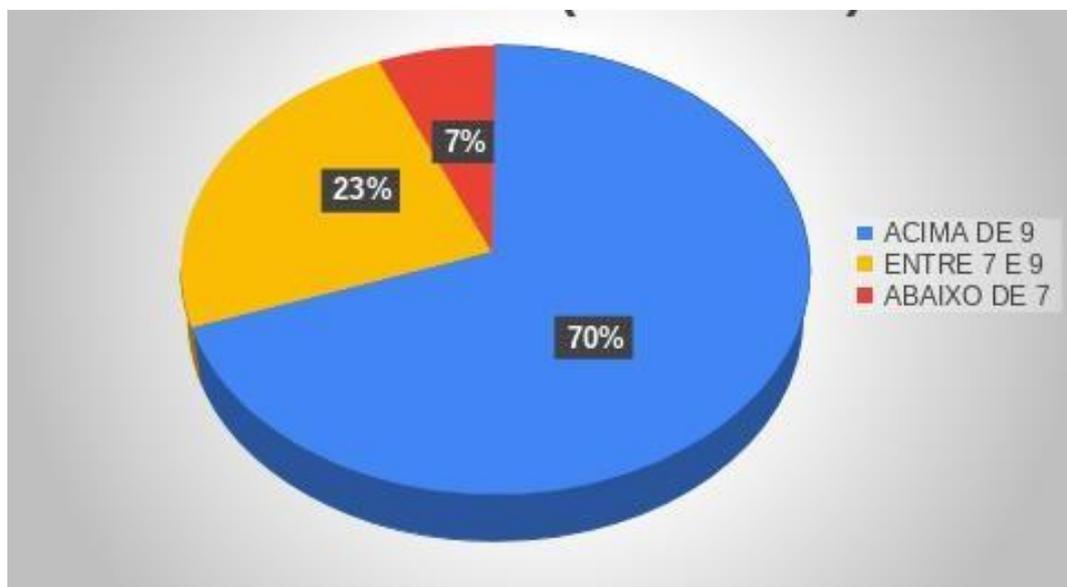
Pode-se observar que mais da metade dos dados estão em consonância com os princípios, de acordo com a média apresentada pela ferramenta. Entretanto, ainda há alguns dados que não atingiram uma pontuação suficiente, e outros não estão enquadrados nas médias, de modo que a média é realizada com os resultados dos outros princípios, os quais são vistos como “pesos”, tendo a média final FAIR.

Percebe-se que houve um crescente quanto à produção de dados, principalmente a respeito do assunto aqui abordado, com isso nota-se que há uma disponibilização dos dados de forma que eles apresentaram uma fácil localização, acesso, utilização e em sua maioria uma interoperabilidade entre outros dados. Diante disso tem-se que FAIR conforme Mons et al (2017, p. 50) :

FAIR refere-se a um conjunto de princípios, focados em garantir que os objetos de pesquisa sejam reutilizáveis, e realmente serão reutilizados, tornando-se tão valiosos quanto possível. Eles deliberadamente não especificam requisitos técnicos, mas são um conjunto de princípios orientadores que fornecem um continuum de reutilização crescente, por meio de muitas implementações diferentes. Eles descrevem características e aspirações para sistemas e serviços para apoiar a criação de resultados de pesquisa valiosos que poderiam ser avaliados rigorosamente e amplamente reutilizados, com crédito apropriado, para benefício do criador e do usuário.

Com base nisso, analisou-se também as médias de cada princípio, de modo a observar a aderência dos conjuntos de dados em cada um deles. Vejam-se o Gráfico-09.

Gráfico 09- Média *FINDABLE*



Fonte: Dados da pesquisa, 2021

Verifica-se, no estudo aqui realizado, que a média *FAIR findable* está em consonância com a média *FAIR* total, em que 70% dos dados possuem uma média acima de 9, 24%. Ainda se verificou que outros dados se encontram entre 7 e 9 possuem uma média inferior ao esperado resultam em 6%, sendo que alguns dados apresentaram o identificador local, metadados simples e até mesmo não apresentavam o identificador em todos os registros / arquivos de metadados que descrevem.

Ao que se pode constatar, a maioria dos dados são encontrados possuindo um identificador persistente, único e global, metadados ricos e alguns estruturados, eles incluem, de forma clara e explícita, o identificador dos dados que descrevem e são registrados ou indexados em um recurso pesquisável, de modo que nem todos estão indexados, apenas registrados, visto que os dados foram achados em mecanismos de pesquisa na web.

Logo, nota-se que o primeiro princípio abordado é o *FINDABLE* (ser encontrado), referindo-se que os conjuntos de dados devem ser encontrados quando o usuário realiza a pesquisa nos ambientes voltados para dados nesse princípio, Desse modo, se volta para os identificadores utilizados nos quais tornam possível a localização. Com isso Souza e

Monteiro (2020, p.507) debatem que

os Princípios FAIR e a Encontrabilidade contribuem para o aumento na qualidade e valor agregado aos dados (e metadados) disponíveis para a sociedade, além de proporcionar o desenvolvimento de padrões de metadados que trazem maior cobertura para os dados abertos de instituições por meio dos seus repositórios digitais, de maneira mais interativa e dinâmica.

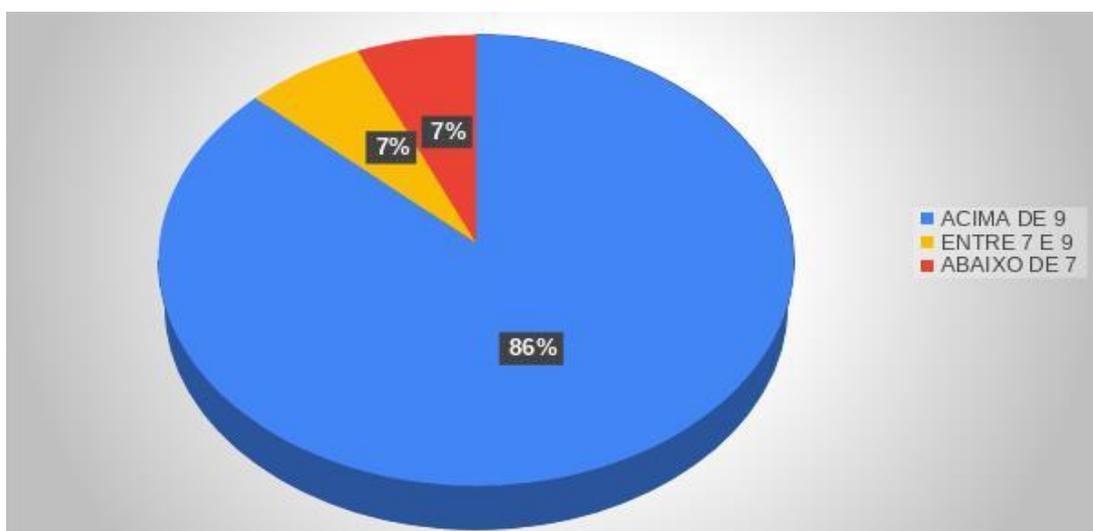
É importante ressaltar que esse princípio representa o ser encontrável, tendo em mente que os dados e os documentos precisam ser detectados utilizando-se de um verificador e partindo de metadados, assim como do uso de palavras-chave, tendo assim tanto os dados quanto os documentos localizados.

Acerca disso Moura *et al.* (2018, p.113) debatem que “[...] eles devem ser encontrados em quanto tempo, os dados digitais devem ser consistentes, ter seus metadados, deve ter metadados processáveis por máquina, que permitam interagir com outros objetos de dados, uso de identificadores únicos e persistentes.” Para se entender mais sobre esse princípio Veiga *et al.* (2019, p. 278) apresentam seus tópicos declarando que

Localizável: (Findable) a) dados e metadados precisam ter um identificador único persistente; b) os dados devem ser descritos com metadados ricos; c) ter o identificador persistente para o conjunto de dados descrito nos metadados; d) metadados e dados devem ser recuperáveis por meio de repositórios confiáveis;

Ao constatar o que o *findable* retratado, tem-se assim uma visão que auxilia em uma melhor observação quanto aos dados aqui tratados e apresentados na ferramenta. No que tange ao segundo princípio, o *ACCESSIBLE* e a aderência dos dados a esse princípio o estudo apontou a seguinte média mostrada no Gráfico 10.

Gráfico 10- Média *ACCESSIBLE*



Fonte: Dados da pesquisa, 2021.

Face ao exposto verifica-se que uma grande percentagem dos dados é acessível, com 87% de sua totalidade tendo apenas 6% entre a média de 7 e 9, e 7% abaixo da média. Observou-se que a maioria possuía os metadados e os dados acessíveis um protocolo aberto e gratuito. Entretanto, alguns dados o acesso somente era possível, por meio de *logins* e pagamentos, se consolidando como dados fechados, a maioria dos dados possuía *download* para seus metadados ou API de *Web Service* padrão.

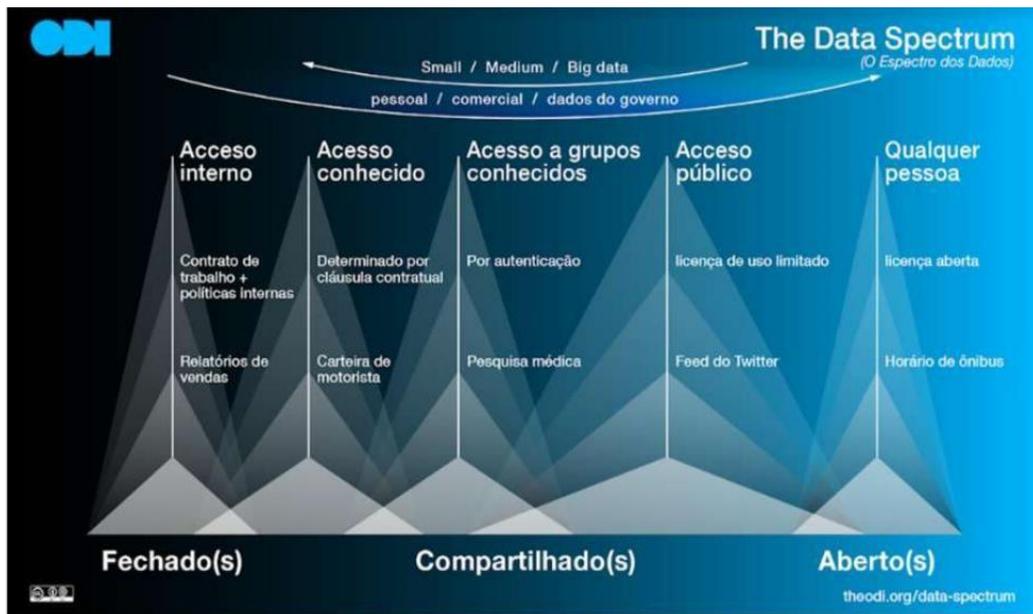
Foi observado também que a maioria dos metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis, eles ficarão hospedados de modo online ou para download e que os identificadores fornecidos são correspondentes com o conjunto dos (meta) dados. Fica claro que os metadados são recuperáveis por seu identificador, usando-se um protocolo de comunicação padronizado, seguindo assim um dos princípios FAIR.

O princípio abordado se preocupa com a acessibilidade dos dados proveniente de um identificador persistente e de protocolos que auxiliem em seu acesso, portanto em conjunto com a análise. Nesse âmbito, Veiga et al (2019, p. 278) destaca que um dado para ser acessível deve possuir

Acessível (Accessible) a) dados e metadados devem ser recuperados pelo seu identificador usando protocolos de comunicação padrão; b) os protocolos devem ser gratuitos, abertos e suportar autenticação e autorização; c) os metadados devem estar acessíveis mesmo quando os dados não estiverem mais disponíveis.

O segundo princípio no qual trata acerca do acesso, em que os conjuntos de dados devem ser acessíveis ao público, remonta ao pensamento sobre os dados abertos, pois para que se tenha um acesso aos dados é necessário que esses sejam abertos. Levando isso em consideração, tem-se que o ODI (*open data institute*) elaborou uma figura ilustrando como os dados podem ser distribuídos quando a dinâmica de uso e o compartilhamento de forma máxima pelos usuários.

Figura 15- Uso e compartilhamento dos dados pela ODI



Fonte: ODI (2020).

Logo, entende-se que a prática do compartilhamento de dados é algo muito relativa, pois em muitos casos eles podem ser compartilhados, embora não acessíveis para todos. Desse modo, ao que consta tem-se o acesso interno, acesso conhecido, acesso a grupos conhecidos, acesso ao público e acesso a qualquer pessoa.

Tal fato evidencia que as formas de acesso são estabelecidas de acordo com as entidades que o compartilham. Ficando evidente que os dados de pesquisas financiadas pelos governos são os que mais aderem a uma real abertura dos dados, como demonstrado nessa pesquisa, em que os repositórios de dados governamentais e institucionais são os que mais disponibilizaram seus dados.

A aderência e o aumento quanto à prática de compartilhamento vêm sendo percebida cada vez mais, após o surto da COVID-19, haja vista que para a produção das vacinas, foi necessário que as entidades envolvidas na criação compartilhassem seus estudos para que assim fosse produzida uma vacina eficaz em pouco tempo. O *Open Data Institute*(ODI) ratifica tal compreensão ao afirmar que

“A resposta inicial à pandemia de COVID-19 foi de colaboração em massa. Nasceram novas redes, iniciativas, parcerias e grupos de trabalho formados por governos, empresas, academia e outros. Muitas dessas abordagens colaborativas tinham dados e compartilhamento de dados em seu núcleo (ODI, 2021).”

Em consonância com essa afirmativa é pertinente observar que com a ciência aberta e o compartilhamento dos dados surgem políticas

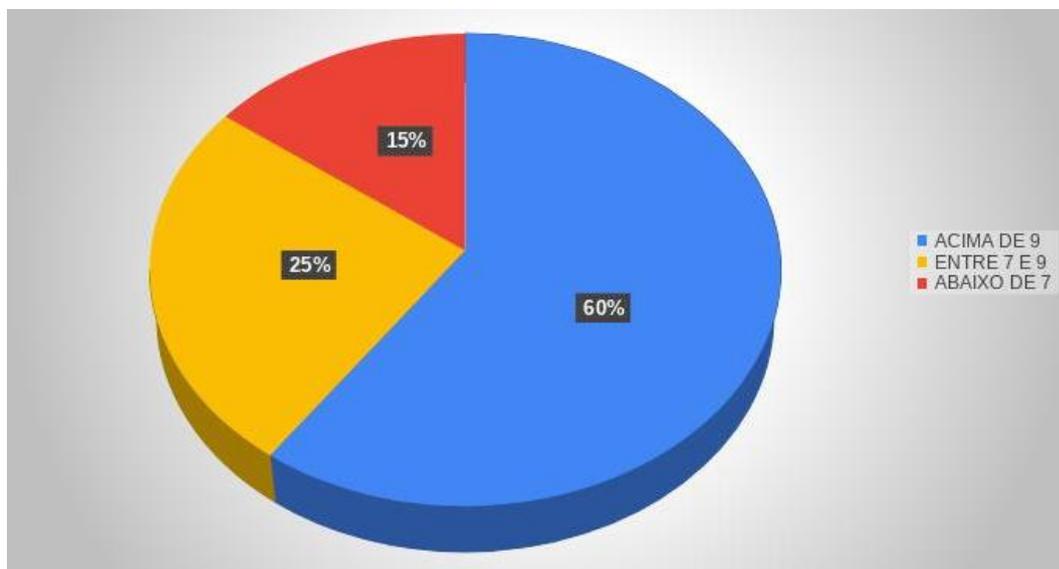
públicas que envolvem o acesso aberto, como por exemplo, as classificações previstas na Lei Acesso à Informação (Lei nº 12.527, de 18 de novembro de 2011) em que Lecardelli (2020, p. 47) traz que

Nos casos de restrição de acesso, os metadados podem ofertar informações sobre os responsáveis e como conseguir os dados, em casos especiais. Ex: atualmente é fundamental o acesso, compartilhamento e reuso de dados na área da saúde, no menor tempo possível para acelerar as pesquisas no intuito de desenvolver uma vacina contra o vírus SARS-CoV-2, a fim de mitigar o avanço da Pandemia (COVID-19). Nesse caso, o tempo é fator decisivo na preservação de milhares de vidas e a ciência está sendo pressionada a avançar de forma mais rápida.

Em diálogo com a afirmativa, percebe-se seu interesse quanto aos metadados, de modo que eles são de grande importância para que se tenha não apenas um acesso, mas, também, para se ter uma ligação entre os dados. Logo tem-se a possibilidade da interoperabilidade como o próximo princípio a ser tratado aqui.

Ao analisar a média dada pela ferramenta referente à aderência acerca do terceiro princípio tem-se as porcentagens mostradas no Gráfico 11.

Gráfico 11- Média *INTEROPERABLE*



Fonte: Dados da pesquisa, 2021.

Quanto à média FAIR *Interoperable*, tem-se que 58% dos dados deram acima da média 9, 28% entre 7 e 9 e 14% ficaram abaixo de 7. Isso nos leva a crer que os dados, apesar de possuírem um alto fator localizável e acessível, não possuem tanta interoperabilidade. Talvez essa baixa tenha se dado pela falta de representação dos dados ou conjuntos de dados, sem a utilização de recursos/instrumentos de controle terminológico, mesmo que uma quantidade substancial utilize dos controles/ recursos

terminológicos ainda existam outros que não se utilizam desse recurso.

Foram analisados também o uso de esquemas explícitos nos conjuntos de dados e esquemas não formalizados. Alguns dados não estavam disponíveis em formatos preferenciais e outros não tinham os conjuntos de metadados vinculados a outros metadados usando identificadores, o que resultou em uma interoperabilidade mediana entre os dados encontrados.

Ao refletir sobre o terceiro princípio, Veiga *et al.* (2019, p. 278, grifos no original), trata que “ [...] Interoperável (‘Interoperable’) a) dados e metadados devem ser codificados usando padrões de representação acordados; b) dados e metadados devem usar vocabulários alinhados aos princípios FAIR e incluir referências relevantes. ”

Convém ressaltar que a interoperabilidade é a capacidade de trabalhar em conjunto promovendo a interação entre as pessoas, sistemas ou organizações trazendo assim uma troca de informação, uma forma mais produtiva e eficaz, com isso em mente Guizzardi (2005, online) esclarece que

[...] devemos ser lembrados que a interoperabilidade não é sobre encontrar maneiras de conectar artefatos de dados, mas, finalmente, sobre a interação dos seres humanos mediada por esses artefatos. Artefatos de informação são instrumentos usados por humanos para harmonizar suas concepções e, portanto, abordagens de interoperabilidade têm sucesso na medida em que podem conectar com segurança essas concepções.

Quanto aos critérios abordados pelo princípio tem-se que um dos mais significativos é o da representação temática Ribeiro, Santos e Moreira (2020, p.35) ao tratarem sobre a representação temática no princípio da interoperabilidade remonta que:

[...]ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento – ontologias de domínio, modelos semânticos de dados, modelos semânticos de metadados, além de vocabulários e tesouros que seguem os princípios FAIR;

De forma a auxiliar na representação de elementos do mundo real, a representação temática se torna um dos pontos significativos para se ter a interoperabilidade, visto que facilita a interação entre o homem e a máquina. Ribeiro e Pereira (2014 , p.3, grifo nosso) trazem que “[...] Esta representação contribuirá para utilizá-los [os elementos do domínio] em projetos futuros, de forma a viabilizar o processo de interoperabilidade

semântica entre ambientes na Web.”

Dada a relevância da representação temática, nota-se seu uso não apenas na recuperação de dados, mas também de informações, já que áreas como a Ciência da Informação, Biblioteconomia e Arquivologia fazem uso dela, ao tratar sobre a interoperabilidade da informação. Diante disso Oliveira (2005, p. 34) ressalta que

Um dos aspectos importantes para a interoperabilidade de informações é a organização do conhecimento e representação da informação. Para que duas pessoas, duas bases de dados ou até duas instituições possam trocar informações de forma eficaz é necessário o entendimento dos códigos utilizados por ambos e que eles tenham o mesmo entendimento quanto ao significado destes códigos. Para isto é necessária a implementação de padrões e normas que possibilitem o entendimento entre eles, rompendo barreiras trazidas pela hiperespecialização e pela fragmentação da informação, desenvolvendo e reforçando trabalhos cooperativos

Com referência a isso, tem-se que a interoperabilidade atrelada à representação de dados, que por sua vez se liga à semântica, de modo que estudos envolvendo a semântica em ambientes informacionais, como por exemplo na *web*, já vem sendo trabalhado há um tempo. Tim Berners Lee *et al.* (2001) propôs uma *web* em que as informações façam sentido perante a sua combinação, nomeando-se como *web semântica*.

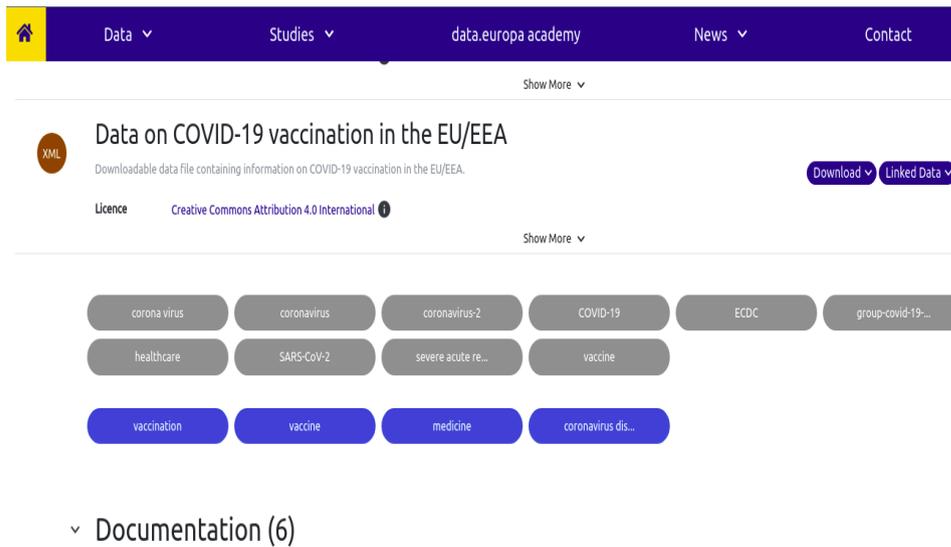
O objetivo da *Web Semântica* é, nessa medida, explicitar as relações entre os signos e o que estes significam, de modo formal e explícito para que essas relações possam ser extraídas e processadas automaticamente. Com isso, o computador poderá fazer inferências simples e recuperações mais precisas ao buscar elementos específicos e seus conteúdos, e não apenas pela sequência descrita (MOTA, 2015, p. 53).

Presume-se então que a interoperabilidade de dados segue a mesma compreensão da interoperabilidade da informação, assim como trazida na afirmativa acima, a sequência descritiva dos dados o torna mais interoperável, de modo a deter sua aplicabilidade em repositórios de dados assim como percebida por Paganine *et al.* (2020, p. 01) nos quais mostram que

Os repositórios de dados científicos são apontados como ferramenta para organização, tratamento, preservação e disseminação de dados científicos. Porém, com a vasta quantidade de modelos de descrição de conjuntos de dados, que são gerados nas mais variadas áreas do conhecimento, surge a necessidade de um estudo sobre os padrões existentes, com vistas a se obter um conjunto de descrição central, assim como o tratamento de especificidades via pequenas extensões deste conjunto, para que se viabilize a interoperabilidade dos repositórios de dados científicos de forma geral em diversos níveis.

Visto assim, nesta pesquisa observou-se que alguns repositórios utilizam desse método, de modo que uns possuem descritores mais ricos e persistentes que outros, conforme enunciado na Figura -16.

Figura 16- Interoperabilidade nos repositórios de dados



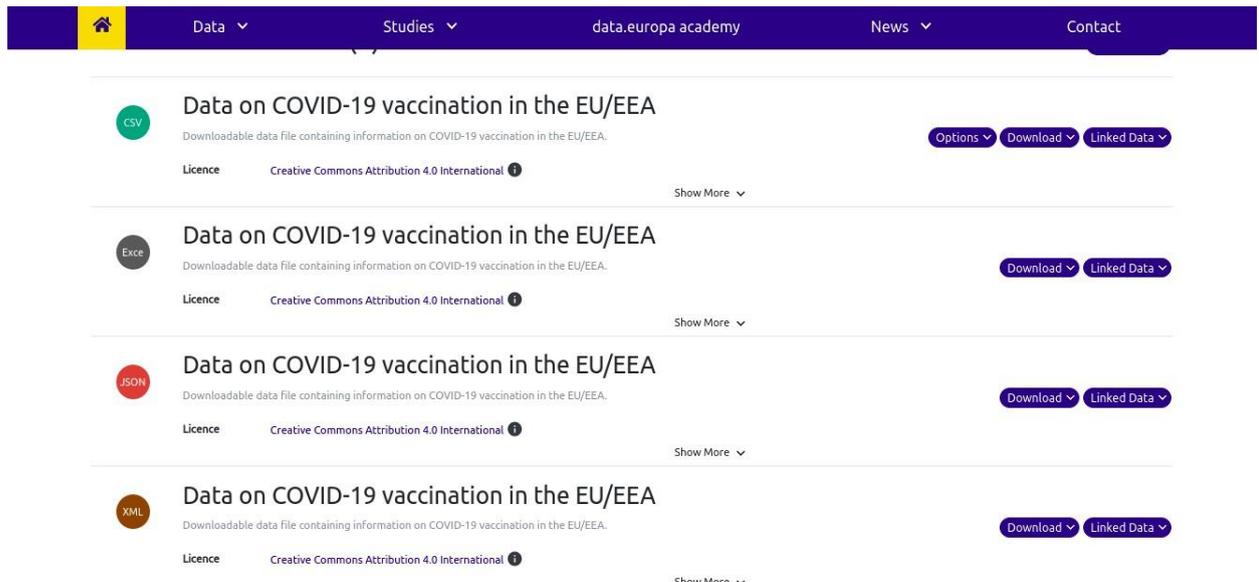
Fonte: ECDC-European Centre for Disease Prevention and Control

Em razão disso, evidenciou-se a importância dos metadados para que a interoperabilidade ocorra, pois assim como para os processos de acesso e reuso tem-se como elemento significativo os metadados condizentes com os conjuntos de dados, para que haja a interoperabilidade não poderia ser diferente, de modo que os metadados possuem algumas variações quanto aos formatos, a vista disso Lecardelli (2020, p.45) apresenta que:

Alguns exemplos de esquemas de metadados são citados, tais como: Dublin Core (DCMS), b) Learning Object Metadata (LOM), c) Visual Resources Association (VRA), d) Encoded Archival Description (EAD), e) Conceptual Reference Model (CIDOC/CRM), f) Metadata for Images in XML (NISO/MIX), g) Machine-Readable Cataloging (MARC), h) Metadata Object Description Standard (MODS), i) Metadata Encoding & Transmission Standard (METS), j) Preservation Metadata.

Deste modo tem-se que alguns dos conjuntos de dados utilizam esses formatos, como mostra a figura-17:

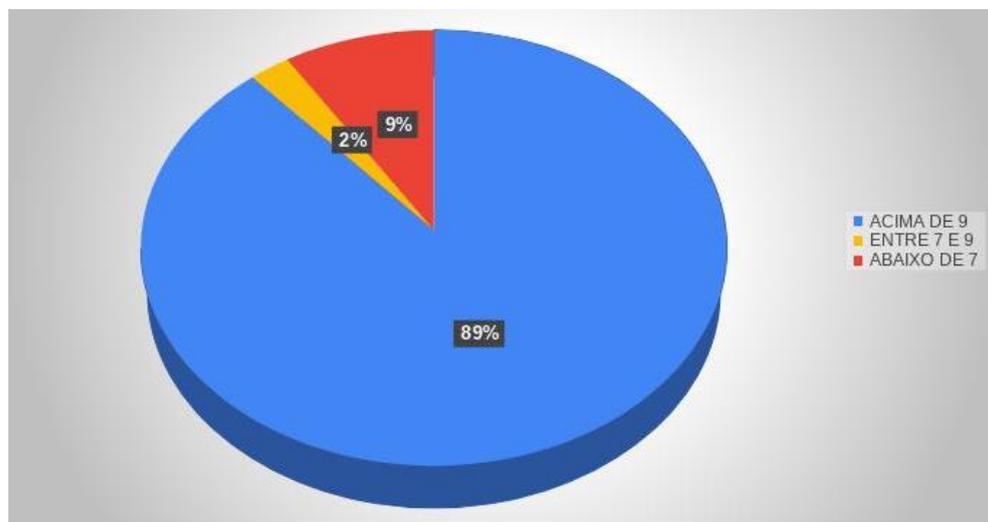
Figura 17- formato dos dados nos repositórios



Fonte: ECDC-European Centre for Disease Prevention and Control

Nesse contexto, nota-se que um metadados em um formato aceitável pela comunidade também é essencial para uma boa interoperabilidade, para o compartilhamento de dados em repositórios para finalizar as observações sobre os princípios, tem-se por fim -gráfico- 12- o último princípio a ser analisado, tendo abaixo o resultado da média calculada pela ferramenta relacionada quanto à aderência do reuso dos dados.

Gráfico 12- Média REUSABLE



Fonte: Dados da pesquisa, 2021.

No que se refere ao princípio *reusable*, pode-se observar que ao trazer os elementos desse princípio Veiga *et al.* (2019, p. 278) referem

que

Reutilizável (Reusable) a) dados e metadados precisam estar associados a atributos relevantes; b) dados e metadados devem ser liberados com licenças de uso claramente definidas; c) metadados e dados devem estar associados à sua proveniência de forma detalhada; d) dados e metadados devem atender aos padrões da comunidade

Isso autoriza a dizer que há um grande reuso dos dados com 89 % tendo a média acima de 9, 9%, outros entre 7 e 9 e apenas 2% com média abaixo de 7. Mesmo com esses resultados, eles se apresentaram aptos para o reuso. Referente ao reuso dos dados, Lecardelli (2020, p. 18) reflete que “[...] O reuso de dados é considerado tendência para acelerar o avanço das ciências de forma mais equitativa. Beneficia a sociedade com o uso dos recursos e investimentos de forma inteligente e economicamente viável, preservando os dados como ativos úteis para novas gerações.”

Em relação às licenças dos dados, foram observadas que a maioria continha a licença creative Commons, em que segundo Branco e Britto (2013, p.19) o objetivo do projeto da licença do Creative Commons “ [...] é fornecer instrumentos legais padronizados para facilitar a circulação e o acesso de obras intelectuais tanto na internet quanto fora dela”.

Quer dizer, é uma licença aberta padrão que oferece alguns termos de condição entre os autores e outros titulares, mas que propicia o compartilhamento dos dados abertamente, entretanto observou-se que em alguns casos houve a restrição e até mesmo a necessidade de pagamento para uso dos dados.

Alguns dados possuíam uma proveniência detalhada assim como um acompanhamento ou vínculo a uma descrição da origem do fluxo de trabalho que produziu os metadados, sendo alguns detalhados ou rastreados. Por fim observou-se que um maior número de dados possuía os padrões de domínio.

Acerca da proveniência é preciso destacar que, segundo Silva (2017, p. 62) a proveniência de dados se define como “Na proveniência de dados científicos, os processos considerados são bastante variados. Assim como na arte, conhecer os processos que geram um objeto digital garantem uma maior confiabilidade e permitem um melhor entendimento dos dados e dos processos que o gerou”.

Presume-se que para a utilização dos dados é necessário ter uma confiança sobre como ele foi gerado, de modo que a credibilidade de

proveniência se torna um dos critérios-chave para que os usuários se sintam seguros em reutilizar aquele conjunto de dados, quanto ao fluxo observa-se que é importante explicitar as entidades envolvidas na criação e geração dos dados, como também as diferentes versões dos conjuntos de dados.

Como resultado final do estudo observou-se que não é de hoje que os estudos sobre os princípios seguem em alta, e que o entendimento e a realização da gestão de dados, assim como a implementação dos princípios FAIR, impulsionaram a ciência que temos hoje, com resultados mais rápidos, com menos custos e seguros.

8 CONCLUSÃO

O final de uma pesquisa, demanda um certo retorno ao ponto de partida a fim de verificar se a questão da pesquisa e os objetivos foram respondidos e alcançados. Assim, esta pesquisa proporcionou reflexões concernentes à gestão dos dados, seguindo pelos princípios FAIR, trazendo um olhar sobre a pandemia da COVI-19, sobre o impacto que ela causa e, naturalmente, a pesquisa apresentou os resultados da empiria sobre a aderência do FAIR feita no mecanismo *Dataset Search*.

Mediante a essa conjuntura e com apoio na pesquisa bibliográfica, pode-se trazer como reflexão que a gestão de dados é uma atividade que aporta contribuições fundamentais ao ambiente científico, inclusive com a alta produtividade de dados que podem ser reaproveitados. Para tanto, se faz necessário que seja feito o plano de gestão de dados que exige os conhecimentos sobre a sua tipologia e seu ciclo de vida a fim de que possam auxiliar, tanto o pesquisador, quanto o usuário no compartilhamento, localização, acesso e reuso dos dados, daí a necessidade de se ter bons repositórios de dados.

Ainda no contexto da pesquisa bibliográfica, ficou evidente que a Ciência da Informação se mostra avançada em seus estudos sobre a gestão de dados e também na crescente produção de trabalhos envolvendo a temática na área. Tal fato, ratifica que há uma preocupação com o compartilhamento, a recuperação, a representação, a organização e a disseminação da informação. Esses indícios sinalizam que o campo se tornou apto para poder realizar a curadoria dos dados, ou seja, a gestão dos dados, de modo que o seu viés tecnológico traga uma afinidade maior ainda. No que diz respeito aos princípios FAIR a pesquisa abordou a proposta de sua origem, a compreensão de cada princípio assim como os sub tópicos trazidos neles para a implementação nos repositórios de dados. Concernente ao impacto causado pela pandemia da COVID-19, ficou demonstrado que há repercussão na produção acadêmica em pesquisas, no caso em baila, nas Ciências da Vida e nas Ciências Sociais Aplicadas. Ressalta-se, nessa última área a presença de várias investigações no campo da Ciência da Informação, focalizando as

mudanças nos parâmetros das produções científicas desde o surto da pandemia. Nesse ambiente, observou-se então, a crescente produção de trabalhos acerca de novos problemas que a área vem enfrentando, a desinformação, as fakes news, a infodemia, entre outros aspectos, que vêm sofrendo implicações perante a esse novo quadro que o mundo enfrenta.

A literatura relativa às Ciências da Vida notabiliza que, concernente ao vírus, os estudos tratam sobre sua origem, estruturação, variantes, sintomas e contágio mostrando que já havia estudos e casos referentes ao SARS-CoV. Entretanto, a mutação do vírus para SARS-CoV-2 ocasionou maior letalidade e transmissão da doença, vitimando assim milhares de pessoas no mundo inteiro.

Ainda nessa narrativa apresentou-se assim que os estudos acerca dos tratamentos e das vacinas impulsionaram o crescimento da produção de dados científicos pelos mundos, pois em pouco tempo já se tinham milhares de pesquisas e estudos realizados a respeito dessa temática, o que ajudou para que houvesse assim uma rapidez quanto à produção das vacinas, pois o compartilhamento de dados consequentemente conduziu ao ganho de tempo entre um estudo e outro.

Em função disso, a pesquisa realizou as análises dos dados acerca das vacinas e da COVID-19 trazendo assim os resultados inerente aos dados colhidos, logo, constata-se, portanto, que o **objetivo geral**, de analisar o modo como está sendo efetivada a aplicabilidade dos princípios FAIR nos repositórios de estudos voltados para as áreas da Saúde e da Ciência da Informação que abordam as vacinas contra COVID-19 foi assim alcançado. Do mesmo modo, os objetivos **específicos** voltados para: **a)** análise das abordagens sobre a gestão de dados científicos a fim de melhor entendimento sobre a temática; **b)** compreensão da filosofia dos princípios FAIR e a sua ligação com a área da Ciência da Informação; **c)** verificação da aderência dos princípios FAIR em repositórios de dados, utilizando os protocolos de elegibilidade e metadados que contemplem dados referentes às vacinas contra a COVID-19 e; **d)** experimentação da ferramenta de verificação de aderência aos princípios Fair Data BR, desenvolvida no âmbito do grupo de pesquisa Web, Representação do Conhecimento e Ontologias (WRCO) da Universidade

Federal da Paraíba. Conclui-se que a ferramenta de análise proposta consegue englobar as premissas de cada princípio trazendo pontos pertinentes quanto à localização, acesso, interoperabilidade e reuso dos dados.

No que concerne aos dados observou-se algumas dificuldades como por exemplo, o difícil ou inexistente acesso a alguns, o acesso por meio de pagamentos assim como uma baixa interoperabilidade entre os dados, em contrapartida observou-se que o reuso dos dados estão em alta, assim como a localização, analisou-se também as áreas que os dados encontrados pertenciam de modo que a área da saúde se sobressai ainda contaram muitos dados referentes à área das ciências sociais de modo que os trabalhos relacionados a essa área trazem um olhar sobre a hesitação quanto às vacinas e até mesmo sobre as fake news, ao que remonta à discussão trazida pela ciência da informação.

Recomenda-se assim maior uso de uma gestão de dados baseados nos princípios FAIR por parte de outras áreas, assim como a aderência da gestão, dos princípios e da ferramenta para os pesquisadores brasileiros. Lembra-se que no patamar nacional poucos dados foram encontrados, portanto, faz necessário acentuar a importância dessa prática assim como a inserção de uma nova cultura de pesquisa no país. Visto que as produções estão em alta, faz-se necessário esse compartilhamento de dados.

Vale ressaltar que estas recomendações não encerram o estudo pois, ainda há muito para se observar nesta temática, tendo novos caminhos para a área da Ciência da Informação, agindo com a curadoria de dados, assim como a implementação dos princípios FAIR no contexto dos repositórios de dados brasileiros, fazendo valer a ferramenta aqui apresentada.

Portanto, espera-se que esta pesquisa traga contribuições para estudos futuros que abordem os assuntos aqui tratados, pois assim, quem sabe se possa chegar a um maior entendimento sobre a relevância dos princípios FAIR na gestão de dados. Ademais, acredita-se que a competência informacional dessa temática ratifica a necessidade da Ciência da Informação e suas subáreas adentrarem com mais presença nos estudos teóricos e nas atividades práticas da gestão de dados, de modo a

auxiliar os profissionais da informação, os pesquisadores e os usuários para uma nova ciência.

Não podemos deixar de mencionar que levar a cabo esta pesquisa, foi um grande desafio, pois, adentrar no caminho desconhecido, significa pisar em espinhos, porém, a força de vontade foi maior para vencer cada obstáculo como uma motivação que não termina aqui, pois pretendemos dar continuidade com a produção de artigos e quem sabe em estudos de maior nível.

REFERÊNCIAS

ABALADA, V. E. T. M.; ARAÚJO, B. M.; GRANATO, M.; RIBEIRO, E. S. Informação em tempos de covid-19 e patrimônio cultural de ciência e tecnologia: de onde viemos e para onde vamos. *Liinc em revista*, v. 16, 2020. DOI: 10.18617/liinc.v16i2.5341 Acesso em: 13 fevereiro de 2022.

ALMEIDA, F. G. **Suporte à gestão de dados de pesquisa: uma ampliação dos serviços oferecidos pelas bibliotecas.** Tese de Doutorado. (Doutorando em Gestão & Organização do Conhecimento) Programa de Pós- Graduação em Gestão & Organização do Conhecimento da Universidade Federal de Minas Gerais. Belo Horizonte. 2019. Disponível em:
https://repositorio.ufmg.br/bitstream/1843/31596/4/Tese_PPGGOC_Fernanda%20Gomes%20Almeida_11out.pdf. Acesso em 24 de junho de 2021.

ANDREADAKIS, Z. et al. O cenário de desenvolvimento de vacinas COVID-19. *Nat. Rev. Drug Discov* , v. 19, p. 305-306, 2020.

BABBIE, E. **The practice of social research.** Belmont:Wadsworth Publ., 1995.

BENTES PINTO, V.; CAVALCANTE, L. E. **Pesquisa bibliográfica e documental: o fazer científico em construção** In: BENTES PINTO, Virgínia; VIDOTTI, Silvana Aparecida B. G. *Aplicabilidades metodológicas em Ciência da Informação.* Fortaleza: Edições UFC, 2015. p. 15-34.

BARBIERI, C. **Governança de Dados.** Rio de Janeiro: Editora Alta Books, 2020.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities.** New York: Scientific American, 2001.

BERTIN, P. R. B.; VISOLI, M. C.; DRUCKER, D. P. A gestão de dados de pesquisa no contexto da e-Science: benefícios, desafios e oportunidades para organizações de P & D. **Ponto de Acesso**, v. 11, n. 2, p. 34-48, 2017. DOI: [10.9771/rpa.v11i2.21449](https://doi.org/10.9771/rpa.v11i2.21449) . Disponível em: <https://brapci.inf.br/index.php/res/download/98667>. Acesso em: 16 junho 2021.

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

BRANCO, S.; BRITTO, W. **O que é Creative Commons? novos modelos de direito autoral em um mundo mais criativo.** Editora FGV, 2013.

Disponível

em: <https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/11461/O%20que%20c3%a9%20Creative%20Commons.pdf?sequenc e=1&isAllowed=y> Acesso em: 23 janeiro 2022

BRITO, J. F.; SILVA, R. C. SANTOS, B. R. P.; MELLO, M. R. G.; MARTÍNEZ-ÁVILA, D. **Arquitetura da informação no contexto da informação em saúde: um olhar para o website do covid-19 no brasil.** AtoZ: Novas Práticas em Informação e Conhecimento, v. 9, n. 2, p. 183-195, 2020. DOI: 10.5380/atoz.v9i2.75091 Acesso em: 13 fevereiro de 2022.

CAPURRO, R.; HJORLAND, B. O conceito de informação. **Perspectivas em ciência da informação**, v. 12, p. 148-207, 2007.

Disponível

em:

<https://www.scielo.br/j/pci/a/j7936SHkZJkpHGH5ZNYQXnC/?lang=pt&format=pdf>. Acesso em: 14 de junho de 2021.

CEZAR, L. S.; MACIEL, A. J. N. Infodemia no contexto da pandemia de covid-19 no brasil: uma política de contaminação? Liinc em revista, v. 17, 2021. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/160974>. Acesso em: 13 fevereiro de 2022

CHIAVEGATTO, A. D. P. Uso de big data em saúde no Brasil: perspectivas para um futuro próximo. **Epidemiologia e Serviços de Saúde**, v. 24, p. 325-332, 2015. Disponível em: <http://scielo.iec.gov.br/pdf/ess/v24n2/v24n2a15.pdf> . Acesso em: 12 de maio 2021

CIOTTI, M. *et al.* The COVID-19 pandemic. **Critical reviews in clinical laboratory sciences**, v. 57, n. 6,

p. 365-388, 2020. Disponível em:

<https://www.tandfonline.com/doi/full/10.1080/10408363.2020.1783198>

Acesso em 9 de dezembro 2021

CORRÊA, P. L. P. Modelo organizacional para gestão integrada de dados da biodiversidade Brasileira. In: DIAS, Guilherme Ataíde; OLIVEIRA, Bernardina Maria Juvenal Freire de (org.). **Dados Científicos: perspectiva e desafios.** João Pessoa: Editora UFPB, 2019. 153 – 175.

CORRÊA, M. V.; CAREGNATO, S. E. **Desinformação e comportamento informacional nas mídias sociais: a divulgação científica na prevenção ao novo coronavírus.** Informação &

Informação, v. 26, n. 1, p. 161-185, 2021. DOI: 10.5433/1981-8920.2021v26n1p161 Acesso em: 13 fevereiro de 2022

CORDEIRO, J. D. R.; FONSECA, A. B.; LESSA, L. R.; LIMA, A. G. I.; NOBILE, M. A

educação em ciências e saúde e o enfrentamento à desinfodemia: um relato de experiências críticas no ensino online. *Liinc em revista*, v. 17, 2021. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/160977>. Acesso em: 13 fevereiro de 2022

CÓRDULA, F. R.; ARAÚJO, W. J. O compartilhamento de dados científicos na era do E- Science. In: DIAS, Guilherme Ataíde; OLIVEIRA, Bernardina Maria Juvenal Freire de (org.). **Dados Científicos: perspectiva e desafios**. João Pessoa: Editora UFPB, 2019. P. 189-207.

CORDEIRO, D. F.; ROCHA, A. S.; VIEIRA, L. M.; CASSIANO, K. K.; SILVA, N. R. Covid- 19 e a circulação de informações em redes sociais: análise em um grupo brasileiro no facebook sobre o coronavírus. *Em Questão*, v. 27, n. online, n. 3, p. 42-67, 2021. DOI: 10.19132/1808-5245273.42-67 Acesso em: 13 fevereiro 2022.

COSTA, M. M; LEITE, F. C. L. Princípios e recomendações basilares para a comunicação dos dados de pesquisa. **Em Questão**, Porto Alegre, v. 23, n. 1, p. 87-112, jan/abr. 2017. Disponível em: <https://www.redalyc.org/journal/4656/465658944018/html/>. Acesso em 25 de agosto de 2020.

COSTA, M. M. **Diretrizes para uma política de gestão de dados científicos no Brasil**. 2017. 288 f. 2017. Tese de Doutorado. Tese (Doutorado em Ciência da Informação). Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília. Brasília, 2017. Disponível em: <http://repositorio.unb.br/handle/10482/24895>. Acesso em: 13 de maio 2021

COSTA, M. M.; CUNHA, M. B. O bibliotecário no tratamento de dados oriundos da e-Science: considerações iniciais. **Perspectivas em Ciência da Informação**, v. 19, n. 3, p. 189-206, 2014. Disponível em: <https://doi.org/10.1590/1981-5344/1900>. Epub 21 Out 2014. ISSN 1981-5344. <https://doi.org/10.1590/1981-5344/1900>. Acesso em: 21 de maio 2021

COSTA, L. C. A. da; SOUSA, M. R. F. de. Análise da literatura sobre a temática "Gestão de dados" no âmbito da Ciência da Informação. **In CID: Revista de Ciência da Informação e Documentação**, [S. l.], v. 11, n. 2, p. 60-81, 2020. DOI: 10.11606/issn.2178-2075.v11i2p60-81. Disponível em: <https://www.revistas.usp.br/incid/article/view/168549>. Acesso em: 13 de maio 2021

COSTA, M. M. **Diretrizes para uma política de gestão de dados científicos no Brasil**. 2017. 288 f., il. Tese (Doutorado em Ciência da Informação). Programa de Pós-graduação em Ciência da informação da Universidade de Brasília, Brasília, 2017. Disponível em : <https://repositorio.unb.br/handle/10482/24895> Acesso em : 7 fevereiro de 2022.

CRABTREE, J. Preservação dos dados de pesquisa: uma parte crítica da fair! **Revista Brasileira de Preservação Digital**, v. 1, n. 2020, 2000. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/158958>. Acesso em: 15 de abril 2021

CRSH- Conseil de Recherches En Sciences Humain's. **Definitions**. Disponível em: https://www.sshrc-crsh.gc.ca/about-au_sujet/policies-politiques/statements-enonces/edata-donnees_electroniques-fra.aspx Acesso em: 26 de agosto 2020.

CRUZ-RIASCOS, S. A.; SILVA, P. R. F. Repositórios brasileiros sobre a covid-19: colaboração para o enfrentamento da pandemia. Páginas A&B, Arquivos e Bibliotecas (Portugal), n. Especial, p. 189-191, 2021. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/157313>. Acesso em: 13 fevereiro 2022.

CRUZ-RIASCOS, S. A.; HOTT, D. F. M.; REZENDE, L. V. R. Análise relacional entre princípios fair de gestão de dados de pesquisa e normativas internacionais de acessibilidade às pessoas com deficiência. **Cadernos BAD (Portugual)**, n. 1, p. 15-29, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/109991>. Acesso em: 7 julho 2021

CURTY, R. Abordagens de reuso e a questão da reusabilidade dos dados científicos. **Liinc em Revista**, [S. l.], v. 15, n. 2, 2019. DOI: [10.18617/liinc.v15i2.4777](https://doi.org/10.18617/liinc.v15i2.4777). Disponível em: <http://revista.ibict.br/liinc/article/view/4777>. Acesso em: 15 de abril 2021

DANKHE, O. **Investigación y comunicación**. México: McGraw Hill. 1986.

DATAONE. **Current member nodes summary**. 2017. Disponível em: <https://www.dataone.org/>. Acesso em: 7 julho 2021

DIAS, G. A.; ANJOS, R. L.; RODRIGUES, A. A. Os princípios FAIR: viabilizando o reuso de dados científicos. In: **Dados científicos: perspectivas e desafios**. João Pessoa: Editora UFPB, p. 177-187, 2019.

DINIZ, M. C. *et al.* Crise global coronavírus: monitoramento e impactos.

Cadernos

de

Prospecção, v. 13, n. 2 COVID-19, p. 359, 2020. Disponível em: <https://portais.univasf.edu.br/covid-19/covid-19/producao-cientifica/35937-130928-2-pb.pdf>. Acesso em: 13 junho 2021.

DOUGLAS, K. *et al.* , "Dados de replicação para: Atitudes públicas em relação à vacinação COVID-19: O papel dos atributos, incentivos e desinformação da vacina", 2021, **Harvard Dataverse**, V2. Disponível em: <https://doi.org/10.7910/DVN/ZYU6CO> ,

DUARTE, P. M. COVID-19: Origem do novo coronavírus. **Brazilian Journal of Health Review**, v. 3, n. 2, p. 3585-3590, 2020. Disponível em: <https://www.brazilianjournals.com/index.php/BJHR/article/view/9131> Acesso em: 15 janeiro 2022.

ECDC-European Centre for Disease Prevention and Control. COVID-19 Vaccine Tracker, 2021. Disponível em [:https://data.europa.eu/data/datasets/covid-19-vaccine-tracker?locale=en](https://data.europa.eu/data/datasets/covid-19-vaccine-tracker?locale=en) . Acesso em 10 dezembro de 2021.

FAPESP- Fundação de Amparo à Pesquisa do Estado de São Paulo. **Programa FAPESP de Pesquisa em eScience**. São Paulo. 2013. Disponível em: https://fapesp.br/publicacoes/2015/folder_escience.pdf. Acesso em 19 outubro de 2019.

FARIAS, M. G. G.; FARIAS, G. B.; SOUSA, F. L. M.; ABREU, P. M. H.; BARRETO, R. B.; CABRAL, V. H. Profissionais da informação: atuação resiliente durante a pandemia de covid- 19: atuação resiliente durante a pandemia de covid-19. *Brazilian Journal of Information Science*, v. 15, 2021. DOI: 10.36311/1981-1640.2021.v15. e02117 Acesso em: 14 fevereiro 2022.

FERREIRA, J. R. S.; LIMA, P. R. S.; SOUZA, E. D.; SOUZA, E. D. Desinformação, infodemia e caos social: impactos negativos das fake News no cenário da covid-19. *Em Questão*, v. 27, n. online, n. 1, p. 30-53, 2021. DOI: 10.19132/1808-5245271.30-53 Acesso em: 12 fevereiro de 2022.

FIOCRUZ – Fundação Oswaldo Cruz. **Glossário do acesso aberto**. Disponível em: <https://portal.fiocruz.br/glossario-acesso-aberto/V>. Acesso em: 10 janeiro de 2022.

FORCE 11- Guiding principles for Findable, Accessible, Interoperable and Reusable Disponível em: <https://www.force11.org/fairprinciples#top>. Acesso em :9 de julho 2021

FRANÇA, F. P.; ARAÚJO, D. O.; SILVA, M. B. A iniciativa digital

convide-i9 no combate à infodemia de covid-19: breves apontamentos de atuação. **AtoZ: Novas Práticas em Informação e Conhecimento**, v. 9, n. 2, p. 248-252, 2020. DOI: 10.5380/atoz.v9i2.76472
Acesso em: 13 fevereiro de 2022.

FRANÇA, A. L. D.; ALVES, I. N. C. N.; SILVA, D. E. L. S. E.; A ciência da informação e a covid-19: análise de sentimentos dos usuários no twitter. **Páginas A&B, Arquivos e Bibliotecas (Portugal)**, n. 14, p. 3-9, 2020. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/152820>. Acesso em: 13 fevereiro de 2022.

FRASER , N. ;BRIERLEY L.; DEY G.; POLKA J. K.; PÁLFY M. ; NANI F. ; COATES J. A. Preprinting the COVID-19 pandemic. **BioRxiv** , p. 2020.05. 22.111294, 2021. Disponível em: <https://www.biorxiv.org/content/10.1101/2020.05.22.111294v3.abstract>. Acesso em:12 janeiro de 2022.

FREITAS, A. R. R.; NAPIMOGA, M.; DONALISIO, M. R. Análise da gravidade da pandemia de Covid-19. **Epidemiologia e serviços de saúde**, v. 29, 2020. Disponível em: <https://www.scielo.br/j/ress/a/TzjkrLwNj78YhV4Bkxg69zx/?lang=pt&format=pdf>. Acesso em: 14 de junho de 2021.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. Editora Atlas SA, 2008.

GOUVERNEMENT DU CANADA. **Déclaration de principes des trois organismes sur la gestion des données numériques.** Disponível em: http://www.science.gc.ca/eic/site/063.nsf/fra/h_83F7624E.html. Acesso em 20 agosto de 2019.

GRAY, J.; SZALAY, E. eScience-a transformed scientific method. Presentation made to the NRC-CSTB. 2007. Disponível em : <http://itre.cis.upenn.edu/myl/JimGrayOnE-Science.pdf>. Acesso em : 10 dezembro 2021

GREEN, A.; MACDONALD, S.; RICE, R. **Policy-making for research data in Repositories: a guide**. May 2009. Disponível em: <<https://www.coar-repositories.org/files/guide.pdf>>. Acesso em: 12 fevereiro de 2022.

GUIMARÃES, J. A. C.; DALESSANDRO, R. C. **As fakes news em um contexto de pandemia pelo coronavírus**. Informação em Pauta, v. 6, n. Especial, p. 24-44, 2021. DOI: 10.36517/2525-3468.ip.v6iespecial.2021.78015.24-44 Acesso em: 13 fevereiro de 2022.

GUIZZARDI, G. Ontological Foundations for Structural Conceptual Models. Enschede, The Netherlands, CTIT PhD Thesis Series, No. 05-74 **Telemática** **Instituto** **Fundamental** **Research** **Series**, No. 015 (TI/FRS/015). 2005. Disponível em: http://doc.utwente.nl/50826/1/thesis_Guizzardi.pdf. Acesso em: 10 dezembro 2021

HARNAD, S. *et al.* The Access/Impact Problem and the Green and Gold Roads to Open Access. **Serials Review**, v. 30, n. 4, p. 310-314, 2004. DOI: [10.1080/00987913.2004.10764930](https://doi.org/10.1080/00987913.2004.10764930). Acesso em: 30 novembro 2021

HENNING, P. C. *et al.* Desmistificando os princípios FAIR: conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados Fair. in: XIX Encontro Nacional de Pesquisa em Ciência da Informação (XIX ENANCIB); v. 24, n. 2, 2018, João Pessoa. **Anais [...]**. João Pessoa, 2018. Disponível em: <https://periodicos.ufpb.br/index.php/pscib/article/view/46969> Acesso em: 2 de julho 2021

HENNING, P. C. *et al.* GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da Ciência Aberta. **Em Questão**, v. 25, n. 2, p. 389-412, 2019. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/84753>. Acesso em: 2 julho 2021.

JANKOWSKI, N. W. Exploring e-Science: an introduction. **Journal of Computer-Mediated Communication**, v. 12, n. 2, p. 549-562, 2007. Disponível em: <https://academic.oup.com/jcmc/article/12/2/549/4583044>. Acesso em: 26 de julho 2019

LAGE, S. R. M.; LUNARDELLI, R. S. A. A representação temática da covid-19 na literatura de cordel. **Informação & Informação**, v. 25, n. 3, p. 402-422, 2020. DOI: 10.5433/1981-8920.2020v25n3p402 Acesso em: 13 fevereiro de 2022.

LAVILLE, C.; DIONNE, J. **A construção do saber: Manual de metodologia da Pesquisa em ciências humanas**. Belo Horizonte: Artmed, 1999. p. 342.

LECARDELLI, J.; MONTEIRO, G.; LUCAS, E. R. O. Gestão de dados científicos no contexto dos paradigmas de Capurro: algumas reflexões. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 15, n. 3, p. 3-20, set. 2019. ISSN 1980-6949. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1284>. Acesso em: 15 abril

2021.

LECARDELLI, J. **Dados Científicos Abertos em Agências de Fomento à Pesquisa: Cenário dos Planos de Gestão de Dados (PGD) e Princípios FAIR.** Dissertação (Mestrado) - Universidade do Estado de Santa Catarina, Centro de Ciências Humanas e da Educação, Programa de Mestrado Profissional em Gestão de Unidades de Informação, Florianópolis, SC, 2020.

LESSA, B. Acesso aos serviços e produtos da biblioteca em tempos de pandemia da covid-19: possibilidades de uso do facebook. **Informação & Informação**, v. 26, n. 2, p. 333-353, 2021. DOI: 10.5433/1981-8920.2021v26n2p333 Acesso em: 13 fevereiro de 2022.

LILI, Z. et al. Progress in scientific data management and sharing. **Bulletin of Chinese Academy of Sciences (Chinese Version)**, v. 33, n. 8, p. 774-782, 2018. Disponível em: <https://bulletinofcas.researchcommons.org/cgi/viewcontent.cgi?article=1518&context=journal> Acesso 11 fevereiro de 2022.

LIMA, EJ da F. *et al.* Vacinas para COVID-19: perspectivas e desafios. **Residência pediátrica—a revista do pediatra**, v. 10, n. 2, p. 1-3, 2020. Disponível em: <https://residenciapediatria.com.br/detalhes/458/vacinas%20para%20covid-19-%20perspectivas%20e%20desafios>. Acesso em: 10 de janeiro de 2022

LIMA, E. J. F.; ALMEIDA, A. M.; KFOURI, R. de Á. Vacinas para COVID-19-o estado da arte. **Revista Brasileira de Saúde Materno Infantil**, v. 21, p. 13-19, 2021. Disponível em: <https://www.scielo.br/j/rbsmi/a/hF6M6SFrhX7XqLPmBTwFfVs/abstract/?lang=pt>. Acesso em: 10 de janeiro 2022

LOTFI, M.; HAMBLIN, M. R.; REZAEI, N. COVID-19: Transmissão, prevenção e potenciais oportunidades terapêuticas. *Clínica chimica acta*, v. 508, p. 254-266, 2020. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/32474009/> Acesso em: 11 de janeiro de 2022

LUCE, R. A new value equation challenge: the emergence of e-Research and roles for research libraries. In **No Brief Candle: Reconceiving Research Libraries for the 21st Century**. Washington, D.C.: Council on Library and Information Resources (CLIR), 42-50. 2008. Disponível em: <http://www.clir.org/pubs/reports/pub142/contents.html>. Acesso em; 26 de julho de 2021.

MACARIO, C. G. do N.; DRUCKER, D. P.; FORTALEZA, J. M.; BERTIN, P. R. B; VACARI, I. Rumo à rede de implantação GO FAIR 'Agro' Brasil: a experiência de uma organização de PD&I na implantação

dos princípios FAIR. In: SALES, L. F.; VEIGA, V. S. de O.; HENNING, P.; SAYÃO, L. F. (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 53-68.

MACIEL, M. L.; APPEL, A. L.; ALBAGLI, S. A e-Science e as novas práticas de produção colaborativa de conhecimento científico. **Revista Internacional de Ciencia y Sociedad**, v. 3, n.1, 2016. Disponível em: <http://ridi.ibict.br/handle/123456789/928> . Acesso em: 19 de maio de 2021.

MACHINGAIDZE, S., WIYSONGE, C. S .Entendendo a hesitação da vacina COVID-19. **Nat Med** 27, 1338-1339 (2021). Disponível em: <https://doi.org/10.1038/s41591-021-01459-7>

Acesso em: 10 janeiro 2022

MADEIRO, A.L. F. Repositórios de dados de pesquisa: investigando sua adoção nos institutos brasileiros de pesquisa. **Informação, Dados E Tecnologia**, v. 1, p. 282-290. Disponível em: https://dadosabertos.info/enhanced_publications/idt/papers/3.pdf Acesso em: 11 de junho 2021.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 5. ed.-São Paulo: Atlas, 2003.

MÁRDERO-ARELLANO, M. N.; ANDRADE, R. S. Preservação digital e os profissionais da informação. **Datagrama Zero**, v. 7, n. 5, 2006. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/5978>. 14 de junho 2021.

MATA, M. L.; GRIGOLETO, M. C.; LOUSADA, M. Dimensões da competência em informação: reflexões frente aos movimentos de infodemia e desinformação na pandemia da covid-19. **Liinc em revista**, v. 16, 2020. DOI: 10.18617/liinc.v16i2.5340 Acesso em: 13 fevereiro de 2022.

MAY, T. Pesquisa social: questões, métodos e processos. 3. ed. Porto Alegre: **Artmed**, 2004.

MELLO, A. P. P.; MESQUITA, H.; VIEIRA, C. E. Módulo 1: introdução. **ENAP**, Brasília 2015. Disponível em: https://repositorio.enap.gov.br/bitstream/1/2399/1/M%C3%B3dulo_1_E_PING.pdf . Acesso em: 14 de junho de 2021.

MENEZES, A. K. F. *et al.* FAIR Data BR: ferramenta para a avaliação de conjuntos de dados científicos. 12º Conferência Luso-Brasileira em Ciência Aberta. 2021. Aceito na conferência da 12º conferência Luso-Brasileira em Ciência aberta (12º CONFOA).

MINAYO, M. C. S. SANCHES, O Quantitativo-qualitativo: oposição ou

complementaridade? **Cadernos de saúde pública**, v. 9, n. 3, p. 237-248, 1993. Disponível em: <https://www.scielo.br/j/csp/a/Bgpmz7T7cNv8K9Hg4J9fJDb/?lang=pt&format=pdf>. Acesso em: 19 de agosto de 2020.

MONTEIRO E. C. de S. de A.; SANT'ANA R. C. G. Repositórios de Dados Científicos na Infraestrutura de Pesquisa: adoção dos princípios FAIR. **Ciência da Informação**, v. 48, n. 3, 20 mar. 2020. Disponível em: <http://revista.ibict.br/ciinf/article/view/4878/4448>. Acesso em: 12 de fevereiro de 2021.

MONS, B. et al. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. **Information Services & Use**, v. 37, n. 1, p. 49-56, 2017. Disponível em : <https://cris.maastrichtuniversity.nl/en/publications/cloudy-increasingly-fair-revisiting-the-fair-data-guiding-princip>. Acesso em: 22 de janeiro 2022.

MORESI, E. et al. Metodologia da pesquisa. Brasília: **Universidade Católica de Brasília**, v. 108, n. 24, p. 5, 2003. Disponível em: <http://www.inf.ufes.br/~pdcosta/ensino/2010-2-metodologia-de-pesquisa/MetodologiaPesquisa-Moresi2003.pdf> . Acesso em: 04 de março 2021.

MORENS, David M. *et al.* The origin. of COVID-19 and why it matters. **The American journal of tropical medicine and hygiene**, v. 103, n. 3, p. 955, 2020. Disponível em : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7470595/> Acesso em: 11 Janeiro 2022.

MOURA JR, P. J.; ARAGÃO, M. H. F.; SOARES, B. C. Os princípios FAIR como diretrizes para maximizar o uso e (re) uso dos dados de pesquisa: traçando um paralelo com a gestão de documentos arquivísticos digitais. in : II workshop de Informação, dados e tecnologia, Santa Catarina. **Anais[...]**. Santa Catarina, 2018. Disponível em: https://dadosabertos.info/enhanced_publications/idt/enhanced_papers/39.pdf Acesso em: 7 fevereiro 2022.

Dispon

MOTA, Denysson Axel Ribeiro. **Representação e recuperação de informação em acervos digitais nos contextos da web semântica e web pragmática: um estudo crítico**. 2015. Tese de Doutorado. Universidade de São Paulo.

MUCHERONI, M. L.; SILVA, J. F. M. A interoperabilidade dos sistemas de informação sob o enfoque da análise sintática e semântica de dados na Web. **Ponto de Acesso**, v. 5, n. 1, p. 3-18, 2011. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/3622>. Acesso em: 12 de fevereiro de 2021.

NATIONAL OF HEALTH. **Nih data sharing policy and implementation guidance**. 2003. Disponível em: http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm#goals. Acesso em: 19 de agosto de 2020.

NOGUEIRA, José Vagner Delmiro. CONHECENDO A ORIGEM DO SARS-COV-2 (COVID 19). **Revista Saúde e Meio Ambiente**, v. 11, n. 2, p. 115-124, 2020. Disponível em: <https://desafioonline.ufms.br/index.php/sameamb/article/view/10321>. Acesso em: 22 de janeiro 2022.

OCDE- Organização para Cooperação e Desenvolvimento Econômico . **Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics**. 2007. Disponível em : <http://www.oecd.org/fr/sti/inno/38500823.pdf>. Acesso em 20 de novembro de 2021.

OID- Open Data Institute. The Data Spectrum. Disponível em: <https://theodi.org/about-the-odi/>. Acesso em:4 fevereiro 2022.

OLIVEIRA, ACS de; SILVA, EM da. Ciência aberta: dimensões para um novo fazer científico. **Informação & Informação**, v. 21, n. 2, p. 5-39, 2016. Disponível em:https://www.brapci.inf.br/repositorio/2017/01/pdf_14bc9e55d1_000022211.pdf. Acesso Em:13 abril 2021.

OLIVEIRA, V. S. **Buscando interoperabilidade entre diferentes bases de dados: o caso da Biblioteca do Instituto Fernandes Figueira**. 2005. 109 f. Dissertação (Mestrado Profissional em Gestão da Informação e Comunicação em Saúde) - Escola Nacional de Saúde Pública Sérgio Arouca, Fundação Oswaldo Cruz, Rio de Janeiro, 2005.

OPENACCESS. **Sobre Repositórios OA**. Braga. 2021. Disponível em://[openaccess.sdum.uminho.pt /?page_id=348](http://openaccess.sdum.uminho.pt/?page_id=348) > Acesso Em: 20 de julho 2021.

PAGANINE L.N. *et al.* **Padrões de interoperabilidade para repositórios de dados de pesquisa**. Brasília, DF: IBICT , RNP; Rio de Janeiro: CNEN; Enschede: University of Twente, 2020. Disponível em: <http://livroaberto.ibict.br/handle/123456789/1085> Acesso em : 12 janeiro de 2022.

POKHREL, S.; CHHETRI, R. A literature review on impact of COVID-19 pandemic on teaching and learning. **Higher Education for the Future**, v. 8, n. 1, p. 133-141, 2021. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/2347631120983481> Acesso em: 12 janeiro de 2022.

QUINTELLA, C. M. *et al.* Vacinas para Coronavírus (COVID-19; SARS-COV-2): mapeamento preliminar de artigos, patentes, testes clínicos e mercado. **Cadernos de Prospecção**, v. 13, n. 1, p. 3, 2020. Disponível em: https://www.noticias.unb.br/images/Artigos/Artigo_Covid_Gracepdf.pdf. Acesso em: 09 de julho de 2021.

RDC-DRC. Research data Canada- Dinnées de Recherche Canadá. **Le glossaire original de DRC**. Disponível em <https://www.rdc-drc.ca/fr/glossaire/glossaire-original/>. Acesso em 20 de dezembro de 2021.

RIBEIRO, C. J. S.; SANTOS, L. O. B. S. ; MOREIRA, J. L. R. Princípios FAIR e interoperabilidade: reprodutibilidade com o uso de modelagem conceitual orientada a ontologias. **Informação & Informação**, v. 25, n. 3, p. 28-51, 2020.

RIBEIRO, C. J. S. ; PEREIRA, D. V. Em busca de uma interoperabilidade semântica na publicação de dados abertos governamentais: a experiência sobre os dados da Previdência Social. (XIV ENANCIB); v. 18 , 2013 , Florianópolis. **Anais** [...] Florianópolis, 2014. Disponível em: https://www.brapci.inf.br/repositorio/2015/12/pdf_a9ddc06174_0000013791.pdf Acesso em : 13 fevereiro de 2022.

ROCHA, J. A. P. Projeto febre amarela: ciclo de vida e tipologia de dados. **Liinc em Revista**, v. 15, n. 2, 2019.

RODRIGUES, E. *et al.* Os repositórios de dados científicos: estado da arte. **RCAAP: Braga**. 2010. Disponível em: <https://hdl.handle.net/10216/23806>. Acesso em: 11 de junho de 2021.

ROSA, S. S.; SILVA, F. C. C.; PAVÃO, C. M. G. Iniciativas de acesso aberto no combate à pandemia: dados abertos e propriedade intelectual na disseminação da informação e conhecimento. **Revista Digital de Biblioteconomia & Ciência da Informação**, v. 19, 2021. DOI: 10.20396/rdbci.v19i00.8666880 Acesso em: 13 fevereiro de 2022.

ROYAL SOCIETY. **Science as an open enterprise**. London: The Royal Society, 2012. Disponível em: <https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/>. Acesso em: 7 fevereiro de 2022.

SALES, L. F.; SAYÃO, L. F. A ciência invisível: revelando os dados da cauda longa da pesquisa. In: **XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (XIX ENANCIB)**. 2018.

SALIAN, V. S. *et al.* COVID-19 transmission, current treatment, and future therapeutic strategies. **Molecular pharmaceuticals**, v. 18, n. 3, p. 754-771, 2021. Disponível em: https://pubs.acs.org/doi/abs/10.1021/acs.molpharmaceut.0c00608?casa_token=_Zlusu4fYmw_AAAAA:8V4xbMPC8AdoDRAGRr6FyMc4jicI-dwgb65dJ2MTr_VjObsaZwdyR5AEoCtOZBWcdzHT9gAjqlbtfyXC Acesso em: 22 jan. 2022

SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, P. B. **Metodología de la investigación**.

México: McGraw-Hill, 1991.

SANTOS, P. L. V. A. C.; SANTANA, R. C. G. Dado e granularidade na perspectiva da informação e tecnologia: uma interpretação pela ciência da informação. **Ciência da Informação**, v. 42, n. 2, 2013. DOI: 10.18225/ci.inf..v42i2.1382 Acesso em: 12 dezembro 2021

SANTOS, P. L. V. A. C.; SANTANA, R. C. G. Transferência da informação: análise para valoração de unidades de conhecimento. **Datagrama Zero**, v. 3, n. 2, 2002. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/5349>. Acesso em: 12 dezembro 2021

SANT'ANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, v. 21, n. 2, p. 116-142, 2016. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/27940> Acesso em: 14 janeiro 2022

SANTOS, L. R. D.; ANDRADE, E. L. M.; LIMA, E. F.; FERNANDES, J. C. C. Sala de aula

invertida no enfrentamento fake news, desinformação e infodemia em época de covid-19. *Revista ACB: Biblioteconomia em Santa Catarina*, v. 26, n. 2, p. 1-20, 2021. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/164407>. Acesso em: 13 fev. 2022

SANTA ANNA, J. Mapeamento sistemático na base de dados em Ciência da Informação: periódicos científicos em discussão. **Brazilian Journal of Information Science: research trends**, [S. l.], v. 12, n. 1, 2018. DOI: 10.36311/1981-1640.2018.v12n1.07. p68. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/7296>.

SANTOS, L. R.; ANDRADE, E. L. M.; Guimarães, M. B. H.; ISAÍAS, P. H.; SILVA, L. H.; LIMA, E. F.; FERNANDES, J. C. C.; MATOS, F. B. O papel das bibliotecas durante a pandemia da covid-19: em busca da emancipação humana. *Ciência da Informação em Revista*, v. 8, n. 1, p. 63-73, 2021. DOI: 10.28998/cirev.2021v8n1d Acesso em: 13 fev. 2022

SAYÃO, L.F.; SALES, L.F. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. Disponível em: <http://www.aben.com.br/Arquivos/420/420.pdf> Acesso em: 9 de julho 2021.

SAYÃO, L.F.; SALES, L.F. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, Londrina, v. 21, n. 2, p. 90 – 115, maio/ago 2016. Disponível em: <https://brapci.inf.br/index.php/res/download/46190>. Acesso em 14 de junho de 2021.

SEMELER, A. R.; PINTO, A. L. Os diferentes conceitos de dados de pesquisa na abordagem da biblioteconomia de dados. **Ciência da Informação**, v. 48, n. 1, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/111525>. Acesso em: 13 fevereiro 2022

SILBERSCHATZ, A.; SUNDARSHAN, S.; KORTH, H. F. **Sistema de banco de dados**. 7. ed. Rio de Janeiro: LTC, 2020.

SANTOS, C. A. C. M. A organização e representação do conhecimento na pandemia de Covid- 19. **Revista Fontes Documentais**, v. 3, p. 75-85, 2020. Disponível em: <https://aplicacoes.ifs.edu.br/periodicos/index.php/fontesdocumentais/article/view/649> Acesso em: 13. Fevereiro 2022

SILVAC. C.; DE CARVALHOC. M. O.; DE LIMAD. C.; COSTAE. S.; DE ANDRADEV.

M. B.; TENORIOB. M.; BRITTOD. B. L. DE A.; TENORIOF. C. A. M. Covid-19: Aspectos da origem, fisiopatologia, imunologia e tratamento - uma revisão narrativa. **Revista Eletrônica Acervo Saúde**, v. 13, n. 3, p. e6542, 27 mar. 2021. Disponível em: <https://doi.org/10.25248/reas.e6542.2021> Acesso em: 23 janeiro de 2022

SILVA, F. C. C. **Gestão de dados científicos**. Interciência, 2019.

SILVA, Daniel Lins da. **Estratégia computacional para apoiar a reprodutibilidade e reuso de dados científicos baseado em metadados de proveniência**. 2017. Tese (Doutorado em Sistemas Digitais) - Escola Politécnica, Universidade de São Paulo, São Paulo, 2017. doi:10.11606/T.3.2017.tde-05092017-095907. Acesso em: 31 janeiro de 2022.

SIEBRA, S. A.; BORBA, V. R.; MIRANDA, M. K. F. O. Curadoria digital: um termo interdisciplinar. **Informação & Tecnologia**, v. 3, n. 2, p. 21-38, 2016. Disponível em:

<http://hdl.handle.net/20.500.11959/brapci/41848>. Acesso em: 23 janeiro de 2022.

SIMÕES, A.R.C. ; DÍAS G.A. ; ANJOS, R.L. Análise dos conjuntos de dados disponíveis no repositório COVID-19 Data Sharing/BR à luz dos princípios FAIR. in: SALES, Luana Farias; VEIGA, Viviane Santos de Oliveira; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**, Rio de Janeiro: IBICT, 2021. P. 97-108.

SOARES, F. B.; RECUERO, R.; VOLCAN, T.; FAGUNDES, G.; SODRÉ, G. Desinformação sobre o covid-19 no WhatsApp: a pandemia enquadrada como debate político. *Ciência da Informação em Revista*, v. 8, n. 1, p. 74-94, 2021. DOI: [10.28998/cirev.2021v8n1e](https://doi.org/10.28998/cirev.2021v8n1e) Acesso em: 13 fevereiro de 2022

SOARES, F. B.; RECUERO, R.; VOLCAN, T.; FAGUNDES, G.; SODRÉ, G. Desinformação sobre o covid-19 no WhatsApp: a pandemia enquadrada como debate político. *Ciência da Informação em Revista*, v. 8, n. 1, p. 74-94, 2021. DOI: [10.28998/cirev.2021v8n1e](https://doi.org/10.28998/cirev.2021v8n1e) Acesso em: 13 fevereiro de 2022.

SOUZA, M.; ALMEIDA, F. G. O comportamento do termo dado na ciência da informação. *Ciência da Informação em Revista*, v. 8, n. 2, p. 39-54, 2021. DOI: [10.28998/cirev.2021v8n2c](https://doi.org/10.28998/cirev.2021v8n2c) Acesso em: 10 fevereiro de 2022

SOUZA, D. L. F.; MONTEIRO, E. W. Encontrabilidade da informação e os princípios FAIR: meios alternativos de gestão de dados abertos. *Ciência da Informação*, [S. l.], v. 48, n. 3, 2020. Disponível em: <http://revista.ibict.br/ciinf/article/view/4971>. Acesso em: 12 fevereiro de 2022.

SOUZA, A. D.; JAVIER JUNIOR, G. F.; FERNANDES, M. R. Bibliotecário clínico em ação na pandemia da covid-19: recursos de informação em saúde para tomada de decisão = clinical medical librarian in action in the covid-19 pandemic: health information resources for decision making. *Revista Bibliomar*, v. 19, n. 2, p. 54-71, 2020. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/150385>. Acesso em: 13 fevereiro de 2022.

TAYLOR, J. **Presentation given at UK e-Science Meeting**, London, July. 2001.

TEIXEIRA, E. B. A Análise de Dados na Pesquisa Científica: importância e desafios em estudos organizacionais. **Desenvolvimento em Questão**, [S. l.], v. 1, n. 2, p. 177–201, 2011. DOI: [10.21527/2237-6453.2003.2.177-201](https://doi.org/10.21527/2237-6453.2003.2.177-201). Disponível

em:

<https://www.revistas.unijui.edu.br/index.php/desenvolvimentoemquestao/article/view/84>. Acesso em: 15 maio 2021.

TRIGGLE, C. R. et al. COVID-19: learning from lessons to guide treatment and prevention interventions. **Mosphere**, v. 5, n. 3, p. e00317-20, 2020. Disponível em: <https://journals.asm.org/doi/full/10.1128/mSphere.00317-20> Acesso em: 04 fev. 2022.

VALÉRIO, E. D.; SILVA, A. L. A. Não faça a “egípcia” em tempos de covid 19: LGBTI+ competentes em informação. *Revista Brasileira de Educação em Ciência da Informação*, v. 7, n. 1, p. 156-175, 2020. DOI: 10.24208/rebecin.v7iespecial.205 Acesso em: 13 fev. 2022.

VARGAS, K. B.; LAWALL, S. Reflexões Biogeográficas acerca da origem, hipóteses, dispersão e distribuição dos Sars-CoV-2 (Coronavírus). **Geografia Ensino e Pesquisa**, Santa Maria, v. 24, 19, 2020. Disponível em: <https://periodicos.ufsm.br/geografia/article/view/45302/pdf>. Acesso em: 12 janeiro 2022.

VASCONCELOS, I. G.; NASCIMENTO, D. A. Mapeamento da produção científica sobre COVID-19. **Interamerican Journal Of Medicine And Health**, v. 3, p. e202003044, 2020.

VEIGA, V. S. de O.; HENNING, P.; DIB, S.; PENEDO, E.; LIMA, J. da C.; SILVA, L. O. B. da; PIRES, L. F. Plano de gestão de dados: uma proposta para a Fiocruz. **Liinc em Revista**, [S. l.], v. 15, n. 2, 2019. DOI: 10.18617/liinc.v15i2.5030. Disponível em: <http://revista.ibict.br/liinc/article/view/5030>. Acesso em: 06 fevereiro de 2022.

XAVIER, F. ; ACOSTA, A. L.; SARAIVA, A. M. , Secondary data from Brazilian tweets about COVID-19 , **Mendeley Data**, V1, 2021. doi: 10.17632/kv52cwvskc.1. Disponível em : <https://data.mendeley.com/datasets/kv52cwvskc/1> Acesso em: 22 out. 2021.

WILKINSON, M. D.; SANSONE, S.-A.; SCHULTES, E.; DOORN, P.; SANTOS, L. O. B. S.; DUMONTIER, M. A design framework and exemplar metrics for FAIRness. *Sci. Data*, 5:180118. 2018. DOI:10.1038/sdata.2018.118. Acesso em: 06 fevereiro de 2022.

WILKINSON, Mark D. *et al.* The FAIR Guiding Principles for Scientific data management and stewardship. **Scientific data**, v. 3, n. 1, p. 1-9, 2016. Disponível em: <https://www.nature.com/articles/sdata201618>. Acesso: 12 de fevereiro de 2021.

APÊNDICE A - TÍTULOS DOS CONJUNTOS DE DADOS

Título dos dados	
1	Dados de progresso da vacinação mundial COVID-19 Conjunto de dados mostrando as taxas de vacinação da Pfizer e BioNTech.
2	Rastreador de vacinas COVID-19
3	Vacinações COVID-19 nos Estados Unidos, Condado
4	Alocações de distribuição de vacinas COVID-19 por jurisdição - Pfizer
5	Alocações de distribuição de vacinas COVID-19 por jurisdição - Moderna
6	Ensaio clínico para vacinação COVID-19
7	Alocações de distribuição de vacinas COVID-19 por jurisdição - Janssen
8	Taxa de vacinação contra o coronavírus COVID-19 do Cazaquistão
9	Taxa de vacinação do Coronavírus COVID-19 da Costa Rica
10	Pesquisas de Hesitação de Vacinas COVID-19
11	Progresso Mundial da Vacinação COVID-19
12	Coronavírus e hesitação em vacinas, Grã-Bretanha
13	Taxa de vacinação COVID-19 para a população total: Cidade de Tempe
14	Dados de replicação para: Atitudes públicas em relação à vacinação contra COVID-19: o papel dos atributos, incentivos e desinformação da vacina
15	Vacinas para COVID-19 - estado da arte

16	Vacinas para crianças: COVID-19
17	Número cumulativo de doses de vacina COVID-19 administradas no Canadá por jurisdição
18	Uma pesquisa inicial sobre a aceitação da vacina COVID-19 por habitantes urbanos e rurais no Delta do Rio Vermelho, Vietnã
19	Flash Eurobarometer 494 : Atitudes sobre a vacinação contra o Covid-19
20	Taxas de vacinação COVID-19 e odds ratio por grupo sociodemográfico
21	Taxa de vacinação COVID-19 para população elegível para vacina
22	Taxa de vacinação COVID-19 para a população total
23	Barreiras à vacinação contra a COVID-19
24	Conjunto de dados COVID-19 Indonésia (caso e vacinação)
25	Atitudes da vacina COVID-19 2021 entre pessoas que vivem sem-teto em Los Angeles, CA
26	Taxas de vacinação COVID-19 por CEP (Condado de Maricopa)
27	Vacinas contra a doença de coronavírus (COVID-19): Visão geral
28	Vacinas COVID-19 por sexo (Condado de Maricopa)
29	Doses de vacina COVID-19 dadas a São Franciscanos por Demografia
30	O desenvolvimento e validação de uma escala tridimensional de hesitação da vacina COVID-19
31	Vacinas COVID-19 por raça e etnia (Condado de Maricopa)
32	Vacinas COVID-19 por idade (Condado de Maricopa)
33	Número cumulativo e porcentagem de pessoas que receberam uma vacina COVID-19 por produto de vacina, número de doses, jurisdição e semana do relatório
34	Rússia Vacinação Covid

35	Dados da pesquisa sobre risco percebido de COVID-19, percepção da vacina COVID-19 e intenção de vacinação COVID-19 entre os vietnamitas
36	Atitudes e aceitação da vacinação COVID-19 entre enfermeiras e parteiras em Chipre: uma pesquisa transversal.
37	Vacinas e tratamentos para COVID-19
38	Coronavírus e os impactos sociais na Grã-Bretanha: atitudes em relação às vacinas
39	Correlação entre vacinação contra SARS-Cov-2, incidência de COVID-19 e mortalidade: rastreando o efeito da vacinação na proteção da população em tempo real
40	Dados de replicação para: além da primeira dose - acompanhamento da vacina Covid-19 e medidas de proteção contínuas
41	Características psicológicas e o papel mediador do Modelo 5C na explicação da intenção de vacinação dos alunos contra a COVID-19
42	Pesquisa de vacina COVID-19 no conjunto de dados
43	Regressão logística multinomial das intenções do australiano de obter uma vacina COVID-19 se estivesse disponível hoje.
44	Dados de replicação para: A relação entre os equívocos dos adultos dos EUA sobre as vacinas COVID-19 e as preferências de vacinação
45	Dados de replicação para: Fatores associados à probabilidade de adultos dos EUA aceitarem a vacinação contra COVID-19: evidências de uma pesquisa e análise conjunta baseada em escolhas
46	Aceitação de uma vacina COVID-19: um estudo transversal na Indonésia
47	Opinião dos entrevistados sobre a aceitação da vacina COVID-19.
48	Hesitação de coronavírus e vacinas, Nível Territorial Internacional (ITL) 1
49	Vacinas COVID-19 por demografia e códigos postais de Tempe
50	Hesitação de vacinação e crenças de conspiração no Reino Unido durante a pandemia de Covid-19
51	Mapeamento atualizado do tratamento COVID-19 e desenvolvimento de vacinas (covid19-help.org data dump)
52	Dados para revisão da vacina COVID-19

53	Dados secundários de tweets brasileiros sobre COVID-19
54	Linfadenopatia após a vacina BNT162b2 Covid-19: achados preliminares de ultrassom
55	Coronavírus e os impactos sociais na Grã-Bretanha: atitudes em relação ao reforço da vacina contra o coronavírus (COVID-19) e vacinas contra a gripe de inverno
56	Taxa de vacinação COVID-19 para população elegível para a vacina: Cidade de Tempe
57	Os fatores que determinam a preferência pela vacina COVID-19
58	GISCorps - Locais de Vacinação Covid19
59	Características socioeconômicas e demográficas que influenciam a hesitação e recusa da vacina COVID-19 na África Ocidental – Gana
60	Vacinação com BNT162b2 em pessoas com mais de 80 anos induz fortes respostas imunes humorais com neutralização cruzada da variante brasileira P.1
61	Tweets da vacina COVID-19 em turco
62	Administração de vacinas DC COVID-19
63	Número acumulado e porcentagem de pessoas que receberam uma vacina COVID-19 no Canadá por número de doses, faixa etária, sexo e jurisdição
64	O labirinto de patentes das vacinas COVID 19
65	Número acumulado e porcentagem de pessoas em populações-chave que receberam uma vacina COVID-19 no Canadá, por número de doses e semana do relatório
66	Atitudes relevantes para a política em relação à vacinação contra COVID-19: associações com demografia, risco à saúde e fatores sociais e políticos.
67	COVID-19: Vigilância pós-comercialização de segurança de vacinas
68	RASTREADOR DE P&D COVID-19
69	Diferenças no uso do tempo entre bloqueios, por status de vacina e outros dados demográficos, Grã-Bretanha
70	Mensagens para impulsionar as vacinas COVID-19: enfatize os benefícios pessoais para a saúde

71	Cobertura da vacina DC COVID-19 por ala
72	Aceitação da vacina Harvard EPREP COVID-19
73	A vacina Sputnik V provoca soroconversão e capacidade de neutralização para SARS CoV-2 após uma dose única - conjunto de dados
74	Covid-19 e evolução das vacinas no mundo
75	Dados sobre a hesitação da vacina COVID-19 entre a população adulta em Bangladesh
76	Dados sobre crença em saúde, comportamento planejado ou antecedentes psicológicos da hesitação da vacina COVID-19 entre os adultos de Bangladesh
77	Percepção e participação de habitantes urbanos e suburbanos na vacinação COVID-19: conjunto de dados de uma pesquisa on-line em Hanói, Vietnã
78	Paisagem e seleção de epítomos de vacina em SARS-CoV-2
79	Apêndice suplementar para "Reações cutâneas relatadas após a vacinação Moderna e Pfizer COVID-19: um estudo baseado em registro de 414 casos"
80	Monitoramento de conhecimento, percepções de risco, comportamentos preventivos e confiança para informar a resposta a surtos pandêmicos.
81	Fornecimento de vacinas DC COVID-19
82	Alocação de vacinas do conjunto de dados
83	Intenções dos alunos sobre a participação e razões da vacinação COVID-19: Um relatório de dados na cidade de Ho Chi Minh, Vietnã
84	Atitudes dos profissionais de saúde turcos em relação à vontade dos pais de imunizar crianças com vacinas COVID-19
85	Os 100 principais artigos sobre vacinas COVID-19 citados
86	Elicitação de imunidade amplamente protetora ao sarbecovírus por vacinas de nanopartículas de domínio de ligação ao receptor, Walls et al.
87	Dados de replicação para: Blocos de construção da aliança de vacinas: um experimento conjunto sobre suporte popular para formatos internacionais de cooperação COVID-19
88	Orientação de planejamento para clínicas de imunização para vacinas COVID-19

89	A variante B.1.1.7 do SARS-CoV-2 é suscetível a anticorpos neutralizantes induzidos pelas vacinas Spike ancestrais. Shen et al. Tabelas Suplementares S1, S2, S3
90	Dados globais de vacinas com mapa base do Johns Hopkins Centers for Civic Impact
91	Reações Adversas à Vacina Mundial COVID-19
92	COVID-19 Análise de tópicos de vacinas no Twitter
93	Disposição para doar para o fundo de vacinas COVID-19 de pessoas altamente educadas no Vietnã
94	Eficácia da ação percebida e confiança na vacina COVID-19.
95	Taxas de vacinação COVID-19 por cidade (Condado de Maricopa)
96	Material Suplementar: Cálculos de Comprimento de Página para Manuscrito BMJ EBM sobre Transparência da Vacina COVID-19
97	Este material suplementar é criado para apoiar os dados descritos no manuscrito: Covid-19 em Cartagena e no Departamento de Bolívar, Colômbia. Situação atual, perspectivas e desafios até a chegada da vacina
98	Número acumulado e porcentagem de pessoas que receberam uma vacina COVID-19 no Canadá por número de doses e jurisdição
99	Preditores de aceitação, intenção e hesitação da vacina COVID-19: uma revisão de escopo.
100	COVID-19: Como as vacinas são desenvolvidas
101	Aceitação da vacinação COVID-19 entre profissionais de saúde e não profissionais de saúde na China: uma pesquisa.
102	Vacina Moderna COVID-19: o que você deve saber
103	Os determinantes das percepções relacionadas ao COVID
104	As percepções relacionadas ao COVID na Índia e nos EUA: uma análise comparativa
105	Brazil Covid Vaccination
106	Atitudes em relação à vacinação contra COVID-19 entre jovens adultos em Zurique, Suíça, setembro de 2020.

107	Vacina Moderna COVID-19: monitoramento contínuo
108	O efeito dos quadros na resistência à vacina COVID-19
109	Shen_et al 2021_SARS-CoV-2 Variante B.1.1.7 Neutralization_Neutralization Titration Data
110	Abordando a hesitação da vacina no contexto do COVID-19: uma cartilha para profissionais de saúde
111	Distribuição da aceitação da vacina, hesitação e falta de vontade por infecção por COVID-19 e condições crônicas autorreferidas.
112	_Preditores de aceitação, intenção e hesitação da vacina COVID-19: uma revisão de escopo.
113	Intervalos de dose estendidos para vacinas COVID-19 para otimizar o lançamento precoce de vacinas e a proteção da população no Canadá no contexto de fornecimento limitado de vacinas
114	Plano de imunização COVID-19 do Canadá: salvando vidas e meios de subsistência
115	Orientação sobre a priorização de populações-chave para a imunização COVID-19
116	COVID-19: Organize uma sessão de engajamento sobre vacinas e medidas de saúde pública em sua comunidade
117	Infográfico COVID-19: Vida após a vacinação
118	Vacinas de mRNA COVID-19
119	Autorizações de medicamentos e vacinas para COVID-19: lista de pedidos recebidos
120	Componentes das vacinas COVID-19 autorizados pela Health Canada
121	Health Canada atualiza o rótulo da vacina Pfizer-BioNTech COVID-19 para refletir relatos muito raros de paralisia de Bell
122	Para viajantes totalmente vacinados sem sintomas de COVID-19 que chegam ao Canadá por terra, ar ou água