



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS – CCSA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO – PPGCI

SAMYR SANTOS DELFINO

INTEROPERABILIDADE DE DADOS EM SAÚDE:
OpenFairEHR - um modelo conceitual no contexto dos princípios FAIR

JOÃO PESSOA - PB

2023

SAMYR SANTOS DELFINO

**INTEROPERABILIDADE DE DADOS EM SAÚDE:
OpenFairEHR - um modelo conceitual no contexto dos princípios FAIR**

Tese apresentada ao Programa de Pós-Graduação em
Ciência da Informação da Universidade Federal da
Paraíba, como requisito para obtenção do título de
Doutor em Ciência da Informação.

Orientador: Prof. Dr. Marckson Roberto Ferreira de
Sousa.

Linha de Pesquisa: Organização, Acesso e Uso da Informação.

JOÃO PESSOA - PB

2023

Catálogo na publicação
Seção de Catalogação e Classificação

D349i Delfino, Samyr Santos.

Interoperabilidade de dados em saúde : OpenFairEHR -
um modelo conceitual no contexto dos princípios FAIR /
Samyr Santos Delfino. - João Pessoa, 2023.
157 f. : il.

Orientação: Marckson Roberto Ferreira de Sousa.
Tese (Doutorado) - UFPB/CCSA.

1. Informações sobre saúde - Interoperabilidade. 2.
Saúde - Padrões de dados. 3. Reutilização de dados. 4.
Princípios FAIR. I. Sousa, Marckson Roberto Ferreira
de. II. Título.

UFPB/BC

CDU 007:613(043)



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

ATA DE DEFESA DE TESE

Defesa nº 078

Ata da Sessão Pública de Defesa de Tese do Doutorando **SAMYR SANTOS DELFINO** como requisito para obtenção do grau de Doutor em Ciência da Informação, Área de Concentração em Informação, Conhecimento e Sociedade e com Linha de Pesquisa em Organização, Acesso e Uso da Informação.

Aos três dias do mês de março de dois mil e vinte e três (03/03/2023), das quinze horas às dezoito horas, na sala virtual do Google Meet, conectaram-se via videoconferência a banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação para avaliar o candidato ao Grau de Doutor em Ciência da Informação na Área de Concentração Informação, Conhecimento e Sociedade, o doutorando **SAMYR SANTOS DELFINO**. A defesa ocorreu de forma remota, com acesso por meio do link: <https://meet.google.com/nrw-wwfb-njh>. A banca examinadora foi composta pelos (as) professores(as): Dr. Marckson Roberto Ferreira de Sousa – PPGCI/UFPB (Presidente/Orientador), Dr. Guilherme Ataíde Dias – PPGCI/UFPB (Examinador interno), Dr. Júlio Afonso Sá de Pinho Neto – PPGCI/UFPB (Examinador interno), Dra. Sandra de Albuquerque Siebra – PPGCI/UFPE (Examinadora externa), Dr. Edberto Ferneda – PPGCI/UNESP (Examinador externo), Dra. Marynice de Medeiros Matos Autran – PPGCI/UFPB (Suplente interna) e Dra. Lucilene Klenia Rodrigues Bandeira – UFPB (Suplente externa). Dando início aos trabalhos, o Professor Dr. Marckson Roberto Ferreira de Sousa, Presidente da Banca Examinadora, explicou aos presentes a finalidade da sessão e passou a palavra ao discente para que fizesse oralmente a apresentação do trabalho de tese intitulado: **INTEROPERABILIDADE DE DADOS EM SAÚDE: OpenFairEHR - um modelo conceitual no contexto dos princípios FAIR**. Após a apresentação, o doutorando foi arguido na forma regimental pelos examinadores. Respondidas todas as arguições, o Professor Dr. Marckson Roberto Ferreira de Sousa, Presidente da Banca Examinadora, acatou todas as observações da banca e procedeu para o julgamento do trabalho, concluindo por atribuir-lhe o conceito:

Aprovado Indeterminado Reprovado.

Proclamados os resultados e encerrados os trabalhos, eu, Professor Dr. Marckson Roberto Ferreira de Sousa, Presidente da Banca Examinadora, lavrei a presente ata que segue assinada digitalmente por mim e pelos demais membros, juntamente com os pareceres de avaliação da Tese e defesa de tese do doutorando, devidamente assinados por seus respectivos avaliadores e em formato digital.

João Pessoa, 03 de março de 2023.

Documento assinado digitalmente
 MARCKSON ROBERTO FERREIRA DE SOUSA
Data: 03/03/2023 18:15:57-0300
Verifique em <https://verificador.iti.br>

Prof. Dr. Marckson Roberto F. de Sousa
Orientador/Presidente – PPGCI/UFPB

Documento assinado digitalmente
 GUILHERME ATAIDE DIAS
Data: 13/03/2023 14:45:17-0300
Verifique em <https://verificador.iti.br>

Prof. Dr. Guilherme Ataíde Dias
Examinador Interno – PPGCI/UFPB

Documento assinado digitalmente
 SANDRA DE ALBUQUERQUE SIEBRA
Data: 08/03/2023 01:17:02-0300
Verifique em <https://verificador.iti.br>

Prof. Dr. Júlio Afonso Sá de Pinho Neto
Examinador Interno – PPGCI/UFPB

Profa. Dra. Sandra de Albuquerque Siebra
Examinadora Externa –PPGCI/UFPE

Documento assinado digitalmente
 EDBERTO FERNEDA
Data: 03/03/2023 18:33:55-0300
Verifique em <https://verificador.iti.br>

Prof. Dr. Edberto Ferneda
Examinador Externo – PPGCI/UNESP

Documento assinado digitalmente
 MARYNICE DE MEDEIROS MATOS AUTRAN
Data: 13/03/2023 09:41:06-0300
Verifique em <https://verificador.iti.br>

Profa. Dra. Marynice de M. Matos Autran
Suplente interna – PPGCI/UFPB

Documento assinado digitalmente
 SAMYR SANTOS DELFINO
Data: 14/03/2023 17:10:15-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Lucilene Klenia R. Bandeira
Suplente externa –UFPB

Samyr Santos Delfino
Doutorando

A Deus e a minha família, em especial minha esposa e filhos que estiveram juntos comigo nessa jornada e que foram, por vezes, a motivação para continuar.
Dedico.

AGRADECIMENTOS

A Deus por ter guiado os meus passos e me concedido a força e sabedoria necessária para conseguir alcançar esse objetivo.

Ao meu pai, João Bosco Delfino (in memoriam) por sempre priorizar bons estudos e à minha mãe, Darcy Santos Delfino, por estar do meu lado quando precisei e ter me mostrado, desde cedo, os valores necessários para construção de um caráter digno. Essa conquista é a prova de que todo investimento e dedicação valeram a pena.

À minha esposa Elluênia Lucena C. Delfino, pelo apoio incondicional às minhas decisões e companheirismo fundamental para suportar todos os desafios enfrentados ao longo dessa jornada. Aos meus filhos, Láiza e Enzo, por serem a personificação da minha motivação para avançar e chegar onde cheguei. Gratidão a todos vocês pela compreensão à minha ausência em alguns momentos, tempo esse necessário para conseguir, seguir e focar na pesquisa e com isso finalizar essa tese.

Ao amigo, Prof. Dr. Lourival Ferreira Cavalcante, (in memoriam), pesquisador por vocação, um grande exemplo de ser humano íntegro e ético. Gratidão pelos valorosos conselhos.

Ao meu orientador Prof. Dr. Marckson Roberto Ferreira de Sousa, pela calma, empatia disponibilidade e competência, além de um profundo conhecimento que, gentilmente, fez questão de compartilhar, viabilizando a realização dessa pesquisa.

Aos professores Prof. Dr. Guilherme Ataíde Dias, Prof. Dr. Júlio Afonso Sá de Pinho Neto, Prof. Dr. Edberto Fereda e Prof. Dra. Sandra de Albuquerque Siebra, membros da banca de qualificação, pelas significativas contribuições que enriqueceram esta pesquisa.

A todos os professores do programa de doutorado do PPGCI/UFPB que, de forma direta ou indireta, contribuíram seja com uma disciplina cursada, com uma dúvida esclarecida, com uma palestra disponibilizada ou mesmo com um exemplo a ser seguido.

Aos amigos discentes e docentes integrantes do grupo de pesquisa Renovation que sempre contribuíram através do compartilhamento de críticas construtivas, essenciais para o refinamento da pesquisa resultando em uma tese de melhor qualidade.

“A nova fonte de poder não é o dinheiro nas mãos de poucos, mas informação nas mãos de muitos”.

John Naisbitt.

RESUMO

Questões relacionadas com a busca pela melhor forma de acesso e uso dos dados, tornam-se essenciais para toda a sociedade contemporânea. Em áreas como a saúde, a necessidade do uso eficiente dos dados é latente, porém, os avanços tecnológicos que ao mesmo tempo proporcionam diversos benefícios também impõem desafios que impactam diretamente no compartilhamento desses dados. Como exemplo desses desafios, destacam-se a diversidade de bases de dados descentralizadas e quantidade excessiva de padrões de interoperabilidade de dados em saúde disponíveis. Neste contexto, o objetivo principal da pesquisa consiste na análise da interoperabilidade de dados, incluindo o desenvolvimento de um modelo conceitual de interoperabilidade em dados em saúde. O percurso metodológico teve início com a realização de uma pesquisa exploratória abordando temas relacionados a bases de dados digitais, metadados, interoperabilidade, banco de dados e princípios FAIR; com base na etapa inicial foram definidos os padrões OpenEHR e FHIR como padrões de referência; em seguida foi escolhida a ferramenta FairDataBR como sendo a ferramenta de avaliação do nível de adaptabilidade dos padrões de referência com os princípios FAIR; por fim, foi desenvolvido o modelo conceitual de interoperabilidade OpenFairEHR. Os principais resultados da pesquisa envolvem a identificação de pontos de melhorias nos padrões OpenEHR e FHIR no que se refere à interoperabilidade de dados em saúde, o desenvolvimento de um modelo conceitual de interoperabilidade denominado de OpenFairEHR. Por fim, foi possível identificar que a utilização do conjunto de princípios de interoperabilidade do FAIR, de forma independente dos demais elementos FAIR, não é recomendada devido a existência de requisitos dependentes entre esses elementos; e a utilização conjunta dos padrões OpenEHR e FHIR apresentam benefícios no que se refere à interoperabilidade de dados em saúde.

Palavras-chave: interoperabilidade; padrões de dados em saúde; reutilização de dados; princípios FAIR.

ABSTRACT

Issues related to the search for the best way to access and use data have become essential for contemporary society as a whole. In areas such as healthcare, the need for efficient use of data is evident, but technological advances that provide various benefits at the same time impose challenges that directly impact the sharing of this data. Examples of these challenges include the diversity of decentralized databases and the excessive amount of health data interoperability standards available. In this context, the main objective of the research consists of analyzing data interoperability, including the development of a conceptual model of data interoperability in healthcare. The methodological approach began with conducting an exploratory research on topics related to digital databases, metadata, interoperability, databases, and FAIR principles. Based on the initial stage, OpenEHR and FHIR standards were defined as reference standards. The FairDataBR tool was then chosen as the evaluation tool for the level of adaptability of reference standards with FAIR principles. Finally, the OpenFairEHR conceptual interoperability model was developed. The main results of the research involve identifying areas for improvement in the OpenEHR and FHIR standards regarding healthcare data interoperability, developing a conceptual interoperability model called OpenFairEHR. Finally, it was possible to identify that the use of the FAIR interoperability principles set independently of other FAIR elements is not recommended due to the existence of dependent requirements among these elements. Also, the joint use of OpenEHR and FHIR standards provides benefits in terms of healthcare data interoperability.

Keywords: interoperability; health data standards; data reuse; FAIR principles.

LISTA DE FIGURAS

Figura 1 - Percurso Metodológico _____	28
Figura 2 - Recurso Paciente do padrão FHIR _____	55
Figura 3 - Estrutura de camadas do recurso paciente do padrão de interoperabilidade FHIR _____	56
Figura 4 - Diagrama UML da estrutura de Metadados de um recurso FHIR _____	58
Figura 5 - Exemplo do arquétipo Blood Pressure (Pressão Arterial) _____	63
Figura 6 - Exemplo do modelo OpenEHR _____	64
Figura 7 - Itens analisados pela ferramenta FairDataBR _____	65
Figura 8 - Ferramenta FairDataBR: Princípio Findable (Localizável) _____	67
Figura 9 - Ferramenta FairDataBR: Princípio Accessible (Acessível) _____	69
Figura 10 - Ferramenta FairDataBR: Princípio Interoperable (Interoperável) _____	71
Figura 11 - Ferramenta FairDataBR: Princípio Reusable (Reutilizável) _____	73
Figura 12 - Ferramenta FairDataBR: Tela exemplo de nível de adaptabilidade com os princípios FAIR _____	75
Figura 13 - Ferramenta FairDataBR: Nível de adaptabilidade do FHIR aos princípios FAIR _____	81
Figura 14 - Ferramenta FairDataBR: Nível de adaptabilidade do OpenEHR aos princípios FAIR _____	89
Figura 15 - Elementos que compõem o modelo de interoperabilidade proposto _____	91
Figura 16 - Camada de Metadados do modelo proposto _____	93
Figura 17 - Diagrama de Metadados do modelo proposto _____	94
Figura 18 - Elementos de metadados e conexão entre metadados e dados _____	101
Figura 19 - Arquitetura Semântica do OpenEHR _____	104
Figura 20 - Arquitetura Semântica do OpenFairEHR _____	105
Figura 21 - Funcionalidades do OpenFairEHR _____	108
Figura 22 - Ferramenta FairDataBR: Nível de adaptabilidade do OpenFairEHR aos princípios FAIR _____	120

LISTA DE QUADROS

Quadro 1 - Tipo de informação em saúde e padrão utilizado _____	43
Quadro 2 - Descrição dos princípios FAIR _____	49
Quadro 3 - Conjunto de metadados para elementos do padrão FHIR _____	59
Quadro 4 - Conjunto de elementos básicos para os Metadados _____	98
Quadro 5 - Análise de adaptabilidade entre os padrões FHIR, OpenEHR e o modelo OpenFairEHR com os Princípios Fair _____	121
Quadro 6 - Ações desenvolvidas no modelo OpenFairEHR para alcançar o nível de adaptabilidade recomendado pelo FairDataBR _____	123

SUMÁRIO

1 INTRODUÇÃO	14
1.1 PROBLEMA DE PESQUISA	16
1.2 JUSTIFICATIVA, HIPÓTESES E TESE	18
1.3 OBJETIVOS	21
1.3.1 Objetivo Geral	21
1.3.2 Objetivos Específicos	22
1.4 ESTRUTURA DAS SEÇÕES	22
2 PERCURSO METODOLÓGICO	23
2.1 ESTRATÉGIA PARA CONDUÇÃO DA PESQUISA	23
2.2 DELIMITAÇÃO DA PESQUISA	25
2.3 DESCRIÇÃO DO ESTUDO DE CASO	26
3 INTEROPERABILIDADE NA ÓTICA DE DADOS EM SAÚDE	29
3.1 CONCEITOS E DEFINIÇÕES	29
3.1.1 Dados e Informação	29
3.1.2 Metadados	32
3.1.3 Interoperabilidade	36
3.2 NORMATIZAÇÃO DE DADOS EM SAÚDE	38
3.3 PADRÕES DE DADOS EM SAÚDE	40
4 PRINCÍPIOS FAIR	45
4.1 VISÃO GERAL	45
4.2 INICIATIVA GO FAIR E GO FAIR BRASIL	47
4.3 PRINCÍPIO DE INTEROPERABILIDADE DO FAIR	48
5 UMA PROPOSTA DE MODELO CONCEITUAL DE INTEROPERABILIDADE DE DADOS EM SAÚDE COM BASE NOS PRINCÍPIOS FAIR	53
5.1 TIPOS DE DADOS	53
5.2 PADRÕES UTILIZADOS PARA CONSTRUÇÃO DO MODELO OPENFAIREHR	53
5.2.1 <i>Fast Healthcare Interoperability Resources</i> – FHIR	54
5.2.1.1 Recursos do padrão FHIR	57
5.2.2 Open Electronic Health Records – OpenEHR	60
5.2.2.1 Arquétipos e modelos do padrão OpenEHR	62
5.3 ADAPTABILIDADE DOS PADRÕES FHIR E OPENEHR COM OS PRINCÍPIOS FAIR	64
5.3.1 Alinhamento do FHIR com os princípios FAIR	76
5.3.2 Alinhamento do OpenEHR com os princípios FAIR	82

5.4 PROPOSTA DE MODELO OPENFAIREHR DE INTEROPERABILIDADE DE DADOS EM SAÚDE	90
5.4.1 Proposta de Arquitetura.....	90
5.4.2 Camada Metadados	91
5.4.2.1 Padrão de Metadados.....	96
5.4.3 Camada Dados	102
5.4.4 Camada Tecnologia.....	106
5.4.5 Alinhamento entre o modelo proposto OpenFairEHR e os princípios FAIR	112
6 CONSIDERAÇÕES FINAIS.....	125
REFERÊNCIAS.....	129
ANEXOS.....	136
ANEXO A - Alinhamento do FHIR com os princípios FAIR.....	137
ANEXO B - Alinhamento do OpenEHR com os princípios FAIR	144
ANEXO C - Alinhamento entre o modelo proposto OpenFairEHR e os princípios FAIR.....	151

1 INTRODUÇÃO

Os avanços da tecnologia trazem para a sociedade inúmeros benefícios e desafios, os dados e as informações passam a ser o cerne de toda sociedade e seu consumo passa a ser uma prioridade. A ansiedade informacional associada a uma necessidade de imediatismo faz com que o consumo, disseminação e produção dessas informações necessite que seja dedicada uma atenção especial a questões como interoperabilidade, integridade, qualidade e até privacidade dos dados e informações disponíveis.

O uso dos dados como elemento estratégico para o embasamento de tomadas de decisão se torna uma realidade cada vez mais comum para diversas áreas da sociedade. Em especial na área da saúde, é possível reconhecer os avanços que a tecnologia tem alcançado e conseqüentemente as mudanças que esses avanços têm proporcionado para toda a sociedade contemporânea.

O volume de dados produzidos por sistemas de informação em saúde (SIS), também conhecidos como dados em saúde, vem crescendo constantemente, porém, de forma ainda não adequada às necessidades dos usuários que fazem uso desses dados. Questões relacionadas à organização, recuperação e representação desses dados ainda carecem de melhorias para que sejam consideradas ideais.

Existem milhões de dados distribuídos entre instituições de saúde, sejam elas públicas ou privadas e pequenas ou grandes, armazenados de forma isolada, dessa forma, impossibilitando o uso coletivo por profissionais de saúde e até mesmo por gestores públicos, que têm a finalidade de desenvolver políticas públicas benéficas para a sociedade em geral.

Essa realidade, que se refere à descentralização e o grande volume de dados em saúde, está presente para esses profissionais e gestores como sendo um grande desafio, para questões relacionadas à recuperação, acesso e uso desses dados, que têm sua complexidade potencializada a cada vez que esse volume de dados aumenta dentro de uma diversidade de fontes de informação individualizadas (OLIVEIRA *et al.*, 2013).

Fazendo uma reflexão em que coloca em um ponto de destaque a questão do grande volume de dados em saúde produzidos e consumidos diariamente de forma individualizada, e em paralelo a isso a compreensão que esse aumento é diretamente proporcional aos avanços tecnológicos disponibilizados para a área da saúde, é possível inferir que o desenvolvimento de ações que englobam esses dois conceitos, dados e tecnologia, se apresentam como uma alternativa de solução para esse desafio.

Nesse sentido o desenvolvimento de uma pesquisa que tenha como cerne a criação de um modelo de interoperabilidade de dados em saúde se vê como extremamente relevante para a sociedade e para a academia.

A interoperabilidade entre sistemas pode ser compreendida como sendo a capacidade de dois ou mais sistemas trocarem dados entre si, de forma a proporcionar inúmeros benefícios em um contexto globalizado, onde a necessidade de compartilhamento de dados e informações entre usuários, profissionais e gestores se torna requisito básico para a realização de atividades cotidianas outrora realizadas de forma manual e isolada.

O desejo pela interoperabilidade plena necessita de uma análise bem mais profunda do que simplesmente possibilitar a comunicação entre sistemas de informação. Para essa finalidade, faz-se necessário a utilização de diversos instrumentos sintáticos e semânticos, destacando-se entre eles os padrões de interoperabilidade de dados, que proporcionem um processo de comunicação entre esses sistemas sem ruídos e sem dubiedade na interpretação desses dados (MIRANDA, 2015).

O uso de padrões nacionais e internacionais para viabilização da interoperabilidade de dados em saúde não se caracteriza como uma alternativa simples, essa afirmação se fundamenta no grande número de padrões existentes no mercado. Essa diversidade justifica-se pela heterogeneidade dos dados de saúde existentes, que necessitam de interação entre si para viabilizar o processo de tomada de decisões na área da saúde.

Nesse sentido, busca-se com essa pesquisa o desenvolvimento de um modelo conceitual de interoperabilidade de dados em saúde que supra as necessidades de compartilhamento, compreensão e recuperação desses dados de forma a possibilitar o desenvolvimento de uma estrutura sintática e semântica adaptável às principais estruturas de dados de saúde disponíveis no mercado.

Como alternativa para o desenvolvimento do modelo mencionado, esta pesquisa acredita que o uso dos princípios FAIR, acrônimo que significa, segundo tradução nossa, Localizável, Acessível, Interoperável e Reutilizável, e que consiste em princípios criados com o propósito principal de apresentarem um conjunto de elementos esperados para os dados a fim de torná-los mais suscetíveis à atuação de sistemas computacionais, ofereça recursos didáticos relevantes para esse objetivo.

Corroborando com o que foi mencionado até o momento, destaca-se ainda que, desenvolver uma pesquisa que envolva temas como interoperabilidade, princípios FAIR e dados em saúde, além de ser um grande desafio, proporcionará diversos benefícios para a sociedade, a exemplo da redução de custos através da reutilização de dados clínicos, celeridade

e assertividade em processos de atendimento à saúde da população em geral, além do planejamento de políticas públicas com maior assertividade, devido a possibilidade de planejamento com base dados históricos que reflitam a realidade da população contemporânea.

No que se refere aos benefícios existentes para a academia, destaca-se a possibilidade de realizar uma pesquisa inédita que pode ser base para o desenvolvimento de diversas outras pesquisas.

Diante do exposto até o presente momento, com o intuito de criar uma identidade que possa a vir se tornar um padrão de referência em interoperabilidade de dados em saúde com os princípios FAIR, o modelo a ser desenvolvido como resultado dessa pesquisa passará a ser denominado de **OpenFairEHR**.

A definição do termo OpenFairEHR levou em consideração a escolha por desenvolver um padrão de livre acesso e utilização (Open), alinhado com os princípios FAIR (FAIR) e que contemple os dados necessários para o desenvolvimento de um Registro Eletrônico de Saúde – RES ou *Electronic Health Record* (EHR).

1.1 PROBLEMA DE PESQUISA

O uso de recursos de Tecnologia da Informação e Comunicação (TIC) estão cada vez mais inseridos no cotidiano da sociedade contemporânea. Essa realidade, segundo Castells (2013), traz para essa sociedade o desafio de ter que lidar com a tecnologia e aceitar a existência de um processo de transformação que não ocorre de forma homogênea para toda a sociedade, porém, influenciará toda a construção da sua história e do seu destino.

A necessidade de produzir, consumir e compartilhar dados em formato digital, exige mudanças de hábitos e o desenvolvimento de novas habilidades para lidar com um mundo digital ainda desconhecido para muitos.

Ao mesmo tempo que os avanços tecnológicos e o grande volume de sistemas de informação produzidos proporcionam benefícios à área da saúde, eles também evidenciam diversos desafios, dentre eles a necessidade de interoperabilidade de dados em saúde se destaca por possibilitar inúmeros benefícios para a sociedade.

O uso cada vez mais frequente de sistemas de informação na área da saúde, desenvolvidos de forma isolada, favorece ao surgimento de bases de dados em formato digital que operam em sua grande maioria descentralizada, proporcionando o surgimento de estruturas informacionais heterogêneas.

A heterogeneidade das bases de dados não reflete a necessidade da sociedade, que por muitas vezes precisa de respostas imediatas e que não se consegue sem que para isso seja necessário proporcionar alguma forma de integração dessas diversas estruturas de dados, para uma melhor compreensão e tomada de decisão.

Na área da saúde é perceptível a necessidade de integração de dados entre diferentes bases digitais, seja para reutilização de dados em pesquisas científicas, tomada de decisão dos profissionais da saúde ou, até mesmo, para o correto desenvolvimento de políticas públicas por parte do governo.

Nesse sentido, Freire, Meirelles e Cunha (2019), afirmam que o desenvolvimento da informação na área da saúde, através de sistemas de informação, enfraquece as ações caso a informação seja desenvolvida de modo isolado.

Contribuindo com essa afirmação, BRASIL (2016) destaca que a falta de padronização dos procedimentos para produção, armazenamento e recuperação dos dados em saúde e o elevado número de sistemas de informação e sua heterogeneidade são grandes desafios para interoperabilidade desses dados.

Ao mencionar a falta de padronização para dados em saúde, vale destacar que atualmente existem diversos padrões, nacionais e internacionais, disponíveis como alternativa para viabilizar a interoperabilidade desses dados. Essa realidade evidencia a complexidade do tema, pois mesmo com essa oferta de padrões disponíveis a busca pela interoperabilidade de dados em saúde ainda é um desejo a ser alcançado.

Como alternativa de solução para essa questão, se vê como oportuno o desenvolvimento do OpenFairEHR que define-se como sendo um modelo conceitual de interoperabilidade de dados em saúde, proporcionando um maior compartilhamento e uso desses dados de maneira mais consistente e padronizada.

A interoperabilidade de dados é um processo que envolve uma reflexão profunda sobre a estrutura dos dados. Conhecer e entender os padrões de interoperabilidade de dados em saúde disponíveis é necessário para, a partir desse ponto, buscar alternativas que façam uso das melhores práticas, ao mesmo tempo, identifiquem pontos de melhorias nos padrões existentes oferecendo alternativas de solução.

Nesse sentido, se vê como necessário a utilização de conceitos que proporcionem uma visão dos dados de forma mais ampla, possibilitando a inclusão de novos conceitos e práticas não exploradas no contexto de dados em saúde, a exemplo do padrão proposto pelos princípios FAIR.

Os princípios FAIR serão melhor detalhados em seção específica para esse tema, porém, os conceitos presentes no FAIR serão de fundamental importância para o desenvolvimento desta pesquisa.

Diante do exposto, surge o seguinte questionamento: **De que forma a aplicação dos princípios FAIR, no contexto dos dados em saúde, contribui para a interoperabilidade de dados?**

1.2 JUSTIFICATIVA, HIPÓTESES E TESE

A cada dia, novos sistemas de informação são criados e inseridos na sociedade com o objetivo de automatizar uma ou mais atividades outrora realizadas de forma manual. Essa prática, cada vez mais comum em áreas como saúde, justifica-se na necessidade de imersão em um processo conhecido por transformação digital que, segundo Tolboom (2016), consiste no processo de mudança através do uso de tecnologias sociais, móveis, analíticas e tecnologias em nuvem, que impactam de forma significativa a sociedade e seus integrantes de forma individual, organizacional e social.

Na área da saúde, os avanços da tecnologia são cada vez mais perceptíveis, muitos profissionais optam pelos benefícios que a automação de processos pode proporcionar. Dentre esses benefícios, a mobilidade ocasionada através da utilização de recursos de TIC como prontuários eletrônicos, sistemas de digitalização e armazenamento de imagens entre outros recursos possibilitam a transformação de dados em formato físico para o formato digital vêm ganhando muito espaço.

Além do mencionado, a interação remota entre profissionais de saúde e seus pacientes também se apresenta como um grande ganho, viabilizando com isso a realização da telessaúde que, segundo HIMSS (2021), possui como uma de suas diretrizes a possibilidade de transpor barreiras socioeconômicas, culturais e, sobretudo, geográficas, para que os serviços e os dados em saúde cheguem a toda população.

Todo esse contexto influenciado pela presença da tecnologia no cotidiano da sociedade, impôs uma necessidade de ressignificação de conceitos, em especial, os relacionados à comunicação interpessoal, sendo esta uma das grandes mudanças vivenciadas pela população contemporânea. Para Dicio (2020, *online*), comunicação pode ser entendida como sendo “Ação ou efeito de comunicar, de transmitir ou de receber ideias, conhecimento, mensagens etc., buscando compartilhar informações”.

Referente à comunicação interpessoal, em situações anteriores a pandemia, o contato interpessoal era praticado com muita frequência tanto para uma simples conversa como também para a realização de atividades de atendimentos a pacientes pelos profissionais da saúde. Dentro desse contexto, destaca-se como grande mudança a recomendação de distanciamento social, porém, preservando a realização de uma comunicação instantânea que simule o processo de comunicação presencial.

Sendo assim, os recursos de TI utilizados para vídeo conferências, *chats* e vídeo chamadas ganharam destaque e tornaram-se ferramentas essenciais para esse novo cenário de comunicação virtual.

Essa realidade ganhou ainda mais força com a utilização desses recursos na escolha do trabalho remoto em detrimento do trabalho presencial. Esse fato também se evidencia na área da saúde, como fomento à realização de atendimentos virtuais e a opção por arquivos digitais, possibilitando uma maior mobilidade e agilidade para os profissionais da saúde.

Toda essa mudança na área da saúde, possibilitou ganhos para sociedade e para os profissionais dessa área, porém, também surgiram muitos desafios. Dentre esses, a necessidade de interoperabilidade dos dados produzidos nesses diversos sistemas de informação em saúde (SIS), evidencia-se como sendo essencial para o bom funcionamento de processo de prestação de serviços em saúde de forma não presencial.

O desejo pela interoperabilidade de dados em saúde pode ser entendido como sendo uma necessidade fundamental pela busca da excelência no uso desses dados. Da mesma forma que a heterogeneidade dos diversos SIS existentes pode ser vista como um grande desafio para esse mesmo objetivo. Todos esses sistemas produzem um gigantesco volume de dados de forma individualizada inviabilizando seu uso de forma compartilhada e acessível.

O objetivo da interoperabilidade sobrepõe a simples função de troca de dados entre equipamentos de tecnologia, para Pan *et al.* (2021, p.8) “a interoperabilidade digital é a capacidade de alcançar a troca de dados e informações rápida, perfeita, segura e confiável”. Essa definição se alinha ao entendimento que a busca pelo padrão de interoperabilidade em dados de saúde possui um nível de complexidade elevado.

Os diversos padrões utilizados para identificação de dados em saúde também exercem influência quando se pensa no desenvolvimento de um modelo de interoperabilidade de dados. Essa diversidade de padrões eleva a dificuldade de identificação de alternativas na busca pelo uso compartilhado desses dados, não apenas pelos profissionais da saúde para realização da garantia à saúde ao paciente, mas também por entidades públicas que precisam desses dados

para o desenvolvimento de políticas públicas que objetivem melhor qualidade de vida para toda a população.

O elevado número de padrões na área da saúde destaca uma diversidade de informações igualmente elevada, e que precisam ter sua estrutura analisada quando se pensa em padrões de interoperabilidade de dados em saúde.

Ainda no que se refere aos padrões de interoperabilidade disponíveis, destacam-se os padrões *Fast Healthcare Interoperability Resources* (FHIR) e *Open Electronic Health Records* (OpenEHR). É oportuno destacar que os padrões mencionados são utilizados em larga escala de forma global, o que atesta sua eficiência em seu campo de atuação. Os padrões FHIR e OpenEHR serão explorados em tópico específico.

A fim de viabilizar a execução dessa pesquisa, os dados que serão considerados para o desenvolvimento de um modelo de interoperabilidade serão os dados que constituem o Registro Eletrônico de Saúde (RES), são compostos por informações relacionadas a dados demográficos do paciente, notas de evolução do paciente, problemas, medicamentos, sinais vitais, histórico médico, imunizações, dados laboratoriais e relatórios de radiologia.

É oportuno destacar que o RES se apresenta, segundo Dallari e Monaco (2021), como sendo um repositório de dados referente à saúde de um ou mais indivíduos, processável por meios informáticos, transmitido com segurança para um conjunto de múltiplos usuários.

A ideia de um repositório de dados universal está alinhada com a própria definição do termo prontuário médico, que consiste em um documento único contendo um conjunto de informações, sinais e imagens registradas a partir de fatos, acontecimentos e situações sobre a saúde do paciente, bem como a assistência a ele prestada. Esse documento possui caráter legal, sigiloso e científico, possibilitando a comunicação entre membros da equipe multiprofissional e a continuidade da assistência prestada ao indivíduo (CFM 2002a).

Analisando a definição do termo prontuário médico, em especial o trecho que afirma se tratar de um documento único, evidencia-se uma contradição ao aceitar a existência de vários prontuários médicos para um mesmo paciente.

Um ponto favorável na busca da interoperabilidade de dados em saúde pode ser visto no desenvolvimento do Prontuário Eletrônico do Paciente (PEP), regulamentado no Brasil em 2002 através da Resolução 1.639/2002 do CFM, possibilitando, entre outras coisas, a eliminação dos documentos em formato de papel.

É importante destacar que passados mais de 20 anos do surgimento do PEP no Brasil, ainda não existe um RES integrado nacionalmente, segundo Dallari e Monaco (2021). Essa

constatação corrobora com a afirmação de que o tema em questão é demasiadamente complexo e que carece do desenvolvimento de pesquisas específicas na área.

Refletindo sobre o contexto de utilização dos dados em RES e suas especificidades, a exemplo de existência de diversos padrões, ferramentas de acesso e compartilhamento e públicos específicos, a busca pela interoperabilidade deve analisar condições de criação de estrutura comum a todos esses dados, em nível semântico, pensando nisso a possibilidade de utilização dos metadados, seguindo as recomendações de interoperabilidade disponíveis nos princípios FAIR, se vê como uma possibilidade viável.

Os princípios FAIR, alcançaram um reconhecimento internacional assumindo o papel de referência mundial em gestão de dados, após a publicação do artigo de Wilkinson (2016), intitulado: “*The FAIR Guiding Principles for Scientific data Management and Stewardship*”. Para esse autor, os conceitos e recomendações de interoperabilidade presentes nos princípios FAIR estão completamente alinhados com as necessidades de um modelo de interoperabilidade de dados em saúde.

Nesse sentido, para se alcançar a tese de que a aplicação dos princípios FAIR no contexto dos dados em saúde contribui para a interoperabilidade de dados em saúde, surgem algumas hipóteses: a) em uma perspectiva direcionada para a interoperabilidade de dados em saúde, a aplicação do subconjunto de princípios de interoperabilidade do FAIR é satisfatória; b) A união de aspectos de interoperabilidade de dados em saúde dos padrões *Fast Healthcare Interoperability Resources* (FHIR) e *Open Electronic Health Records* (OpenEHR), apresenta resultados mais alinhados com os princípios FAIR.

Diante de exposto, é importante mencionar que essa pesquisa está em conformidade com a linha de pesquisa organização, acesso e uso da informação, considerando que serão exploradas questões referentes a produção, recuperação, uso e impactos dos dados.

1.3 OBJETIVOS

Os objetivos da pesquisa encontram-se descritos na sequência.

1.3.1 Objetivo Geral

Conceber um modelo conceitual de interoperabilidade em dados em saúde com base nos princípios FAIR.

1.3.2 Objetivos Específicos

- a) Identificar padrões de interoperabilidade utilizadas para dados em saúde;
- b) Investigar os princípios FAIR, com ênfase no subconjunto de princípios de interoperabilidade;
- c) Verificar o nível de adaptabilidade dos padrões de interoperabilidade identificados com base nos princípios FAIR;
- d) Avaliar possibilidades de melhorias dos padrões de interoperabilidade identificados.

1.4 ESTRUTURA DAS SEÇÕES

A referida pesquisa encontra-se distribuída em seis seções, sendo a primeira seção a Introdução, contendo definições sobre o problema da pesquisa, justificativa, hipóteses, objetivos e estrutura do trabalho. Seguindo, na seção 2 foi definido todo o percurso metodológico, descrevendo a estratégia para condução da pesquisa, delimitação do universo da pesquisa e definição do estudo de caso.

As seções 3 e 4 abordam elementos teóricos que serviram para uma melhor compreensão e embasamento da pesquisa necessária para o desenvolvimento dessa pesquisa. Respectivamente são abordados temas como interoperabilidade na ótica de dados em saúde, que traz em seu escopo temas como interoperabilidade de dados e conceitos relacionados a dados e informações. Outro tema explorado ao longo do desenvolvimento desse marco teórico da pesquisa, e que se vê como essencial para o seu desenvolvimento, consiste nos princípios FAIR, destaca-se aqui que além dos princípios FAIR propriamente dito, também são explorados temas como padrões de metadados, além do princípio da interoperabilidade do FAIR.

A seção 5 aborda elementos referente ao modelo conceitual de interoperabilidade de dados em saúde que se pretende construir ao longo do desenvolvimento dessa pesquisa. Nessa seção apresenta-se informações sobre padrões de interoperabilidade de dados em saúde existentes, relação entre os princípios FAIR e os dados em saúde e o desenvolvimento de modelo conceitual de interoperabilidade de dados em saúde com base nos princípios FAIR.

Por fim, a seção 6 destaca todos os desafios encontrados ao longo do desenvolvimento dessa pesquisa, bem como as propostas para desenvolvimentos de trabalhos futuros que possam dar continuidade ao desenvolvimento dessa pesquisa, inclusive com a possibilidade de implementação e testes do modelo proposto. Todas essas contribuições estão disponíveis nas considerações finais.

2 PERCURSO METODOLÓGICO

O percurso metodológico para o desenvolvimento dessa pesquisa possibilitou a conexão entre o problema e os objetivos definidos, uma vez que esses objetivos visavam responder à problemática da pesquisa, através da realização de uma investigação científica.

Para uma melhor compreensão referente ao percurso metodológico escolhido, é oportuno destacar que inicialmente foi definida a estratégia para execução da pesquisa, posteriormente foi definido o universo, finalizando com o desenvolvimento de um estudo de caso que objetiva a elaboração de um modelo conceitual de interoperabilidade de dados em saúde com base nos princípios FAIR.

A definição e planejamento da metodologia são essenciais para o sucesso de uma pesquisa, pois a definição da metodologia mais adequada a ser utilizada depende da sua singularidade, considerando que essa escolha possa ser complementada com o passar do tempo e de acordo com a sua necessidade. Nesse sentido, Medeiros (2019, p.35) afirma que “as peculiaridades de seu método diferenciam a ciência de muitas formas de conhecimento humano. E uma de suas particularidades é aceitar que nada é eternamente verdadeiro”.

Uma investigação científica, ainda segundo Medeiros (2019, p.18), só pode ser entendida como sendo uma pesquisa científica se “sua realização for objeto de investigação planejada, desenvolvida e redigida conforme normas metodológicas consagradas pela ciência”.

Contribuindo com o tema, Koche (2014) afirma que o resultado de uma investigação científica pode ser reconhecido através da produção de conhecimento oriunda de suas compreensões a respeito da análise detalhada de seu problema de pesquisa.

2.1 ESTRATÉGIA PARA CONDUÇÃO DA PESQUISA

Esta pesquisa pode ser considerada inicialmente como exploratória, que segundo Gil (2019) constitui a primeira etapa de uma investigação mais ampla e que exige uma revisão da literatura. Considerando as características diversificadas do tema estudado, essa pesquisa exploratória se define como sendo interdisciplinar, tendo seu foco principal a área de Ciência da Informação, na busca de contribuições teóricas para uma melhor compreensão do assunto, constituindo-se a parte inicial de uma investigação mais ampla.

A pesquisa exploratória possibilita uma maior compreensão do problema, favorecendo o surgimento de uma melhor interpretação e elucidação dos fatos a serem estudados. Nesse sentido, a pesquisa exploratória objetiva realizar um levantamento e análise de material

bibliográfico na área de CI, em relação a temas que envolvam estrutura de bases de dados digitais, metadados e interoperabilidade. Além de uma pesquisa interdisciplinar envolvendo as áreas da Ciência da Computação e Ciência de Dados, buscando informações sobre banco de dados e os princípios FAIR.

Referente ao aspecto científico, bem como a estratégia de condução, entende-se que ela é caracterizada como descritiva. De acordo com o Sampieri, Collado e Lucio (2013, p. 100), em pesquisas descritivas “o objeto do pesquisador consiste em descrever situações, acontecimentos e feitos, isto é, dizer como é e como manifesta determinado fenômeno”. Em uma pesquisa caracterizada como descritiva, o objetivo consiste em representar as características e particularidades do objeto de pesquisa, facilitando a sua identificação através de seus atributos de maior significância, sem a necessidade de manipulação de seus dados.

Outra característica importante das pesquisas descritivas consiste na descrição científica do objeto de pesquisa, através da seleção e análise de informações que fazem parte desse objeto. Dessa maneira, é possível realizar a mensuração das informações, seja de forma independente ou conjunta, a respeito dos pensamentos referentes ao objeto estudado (SAMPIERI; COLLADO; LUCIO, 2013).

No que diz respeito a sua abordagem, essa pesquisa se caracteriza como sendo quali-quantitativa, isso se justifica através da necessidade compreensão e explicação de um fenômeno, sem que ocorra subjetividade. Em pesquisas com enfoque qualitativo, o ambiente passa a ter uma importância singular para o resultado, sendo fonte renovável de informações, e tendo uma relação direta com o pesquisador sem que esse possa manipulá-lo.

O enfoque quantitativo dessa pesquisa se apresenta na necessidade de mensuração do percentual de adesão que o modelo proposto possui em relação as recomendações disponíveis nos princípios FAIR.

Diante do exposto, percebe-se que a abordagem predominante dessa pesquisa se fundamenta como sendo qualitativa, porém, diante da necessidade da realização de análises estatísticas para complementar a pesquisa, esta será fundamentalmente quali-quantitativo devido a sua abordagem mista.

Em pesquisas com enfoque qualitativo, de acordo com Prodanov e Freitas (2013, p11), “o ambiente natural é a fonte direta para coleta de dados, interpretação de fenômenos, e atribuição de significados”. No mesmo sentido, Richardson (2017, p5), afirma que “a abordagem qualitativa de um problema, além de ser uma opção do investigador, justifica-se, sobretudo, por ser uma forma adequada para entender a natureza de um fenômeno social”.

Esta pesquisa também se caracteriza como um estudo de caso, que segundo Gil (2019, p13) “o estudo de caso é caracterizado pela pesquisa profunda e exaustiva de um ou de poucos objetos, de maneira a permitir o seu conhecimento amplo e detalhado [...]”.

Através dessa interação mencionada por Gil espera-se com esse estudo de caso a coleta de elementos que subsidiem o desenvolvimento de um modelo conceitual de interoperabilidade de dados em saúde no contexto dos princípios FAIR.

2.2 DELIMITAÇÃO DA PESQUISA

A definição do universo é essencial para a realização de uma pesquisa científica, nesse sentido, Marconi e Lakatos (2022, p6), destacam que “a delimitação do universo consiste em explicitar que pessoas ou coisas, fenômenos etc. serão pesquisadas, enumerando suas características comuns [...]”.

Para o desenvolvimento desta pesquisa, o universo escolhido corresponde aos dados digitais utilizados para composição de um Registro Eletrônico de Saúde (RES), conforme definido pelo HIMSS (2021), que incluem dados demográficos do paciente, notas de evolução do paciente, problemas, medicamentos, sinais vitais, histórico médico, imunizações, dados laboratoriais e relatórios de radiologia (laudos).

A escolha por esse grupo de dados se justifica pelos benefícios existentes para a academia e para a sociedade através do uso compartilhado desses dados, obviamente proporcionando a segurança devida para essa utilização.

É possível entender o RES como sendo uma evolução do Prontuário Eletrônico do Paciente (PEP). Essa compreensão se fundamenta no entendimento de Santos (2011) quando este afirma que o PEP é direcionado apenas para um estabelecimento médico e um único software de gestão das informações médicas. Por sua vez, o RES, além de ser composto por dados relacionados à saúde do paciente tem sua abrangência maior que o PEP contemplando eventos ocorridos em múltiplas organizações de saúde.

Diante do exposto, esse autor entende que, pensando em benefícios para a sociedade, trabalhar com dados que reflitam eventos ocorridos em diversos locais possibilita um ganho maior para o paciente, titular desses dados, bem como para o desenvolvimento de políticas públicas.

2.3 DESCRIÇÃO DO ESTUDO DE CASO

Investigando sobre o conceito de estudo de caso, é possível se deparar com uma infinidade de origens e definições. De acordo com diversos autores, o estudo de caso tem sua origem nas pesquisas médicas e psicológicas, através de análises detalhadas de casos individuais utilizadas na tentativa de explicar a dinâmica e a patologia de uma doença específica (BECKER, 1999).

Outros autores, a exemplo de Gil (2022), afirmam que o estudo de caso teve sua origem vinculada à área do ensino jurídico nos Estados Unidos, tendo sua prática ligada à psicoterapêutica caracterizada pela reconstrução da história de um indivíduo.

Abstraindo a origem do conceito de estudo de caso, e direcionando o foco para definição conceitual, essa pesquisa adotará a definição apresentada por Yin (2015), que destaca o estudo de caso como sendo uma investigação empírica com um método abrangente, lógica do planejamento, da coleta e da análise de dados. Possibilitando a inclusão de estudos de caso únicos e/ou múltiplos, como também abordagens quantitativas e qualitativas de pesquisa.

Entende-se que a definição apresentada por Yin (2015), está alinhada com os objetivos e proposta dessa pesquisa, uma vez que o presente estudo objetiva a criação de um modelo de interoperabilidade de dados e que para isso não contemplará a participação de indivíduos, o que se mostra como essencial nas definições de Gil e Becker.

No que se refere à classificação, de acordo com Ventura (2007), os estudos de caso podem ser classificados como sendo intrínseco ou particular, quando se pretende investigar um caso em particular, ou instrumental, quando o objeto da investigação busca compreender melhor uma questão, orientar estudos ou ser instrumento para pesquisas posteriores.

No que se refere à classificação, entende-se que essa pesquisa, de acordo com a definição apresentada por Ventura (2007), se classifica como sendo um estudo de caso instrumental. Tal afirmação fundamenta-se na análise do objetivo que se almeja alcançar, que consiste no desenvolvimento de um modelo conceitual de interoperabilidade de dados em saúde com base nos princípios FAIR. Detalhando um pouco mais, a presente pesquisa tem como objetivo compreender melhor o processo de interoperabilidade de dados em saúde e como a imersão dos princípios FAIR pode contribuir com essa interoperabilidade.

Ainda de acordo com o Ventura (2007), mesmo sendo uma atividade complexa é possível definir um delineamento através de quatro fases básicas para o desenvolvimento de um estudo de caso, sendo elas: a) delimitação da unidade-caso; b) coleta de dados; c) seleção, análise e interpretação dos dados; d) elaboração do relatório.

No que se refere à fase “delimitação da unidade-caso” o autor define que essa fase consiste na delimitação da unidade que constitui o caso, ou seja, quais dados são suficientes para se chegar à compreensão do objeto como um todo. Considerando sua natureza exploratória e bibliográfica os dados necessários para o desenvolvimento dessa pesquisa estão centrados em materiais científicos relacionados aos dados, padrões de interoperabilidade de dados em saúde, princípios FAIR e metadados. É oportuno citar que esses tópicos estão contemplados nas seções dessa tese que compõem sua base teórica.

No que diz respeito à fase “coleta de dados”, que consiste na definição da estratégia a ser adotada para identificar e coletar os dados definidos na fase anterior, fica evidente que devido à natureza dessa pesquisa a abordagem a ser utilizada será a pesquisa bibliográfica quali-quantitativa.

Referente à “seleção, análise e interpretação dos dados” que consiste na definição de elementos que possam selecionar os dados que serão utilizados no estudo de caso considerando os objetivos e limites da investigação, busca entender quais dados são ou não relevantes para o desenvolvimento do estudo de caso.

Destaca-se aqui que pela natureza da referida pesquisa, os dados coletados devem ter sido publicados em periódicos científicos ou disponibilizados por entidades públicas ou privadas que sejam reconhecidas pelas suas contribuições na temática dessa pesquisa. Dados disponibilizados em veículos de comunicação que não estejam inseridos nessas categorias citadas, não serão considerados para o desenvolvimento dessa pesquisa.

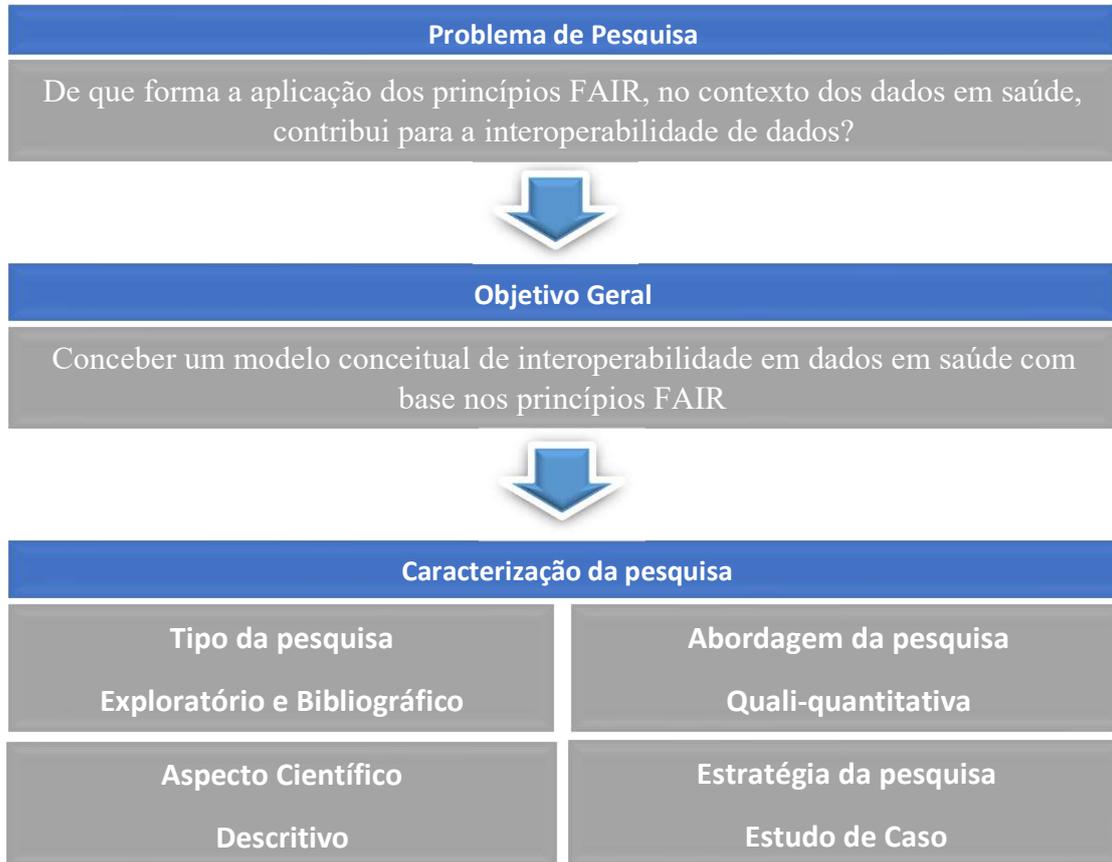
Ainda na referida fase, é oportuno destacar a utilização da ferramenta FairDataBR como instrumento de avaliação do nível de adaptabilidade dos padrões de referência OpenEHR e FHIR, bem como do modelo a ser desenvolvido nessa pesquisa, com os princípios FAIR.

O FairDataBR, que será melhor descrito em tópico específico, caracteriza-se por ser de uso simples e intuitivo, foi construído por pesquisadores vinculados ao Programa de Pós-graduação em Ciência da Informação e o Programa de Pós-graduação em Gestão nas Organizações Aprendentes da Universidade Federal da Paraíba (PPGCI/MPGOA - UFPB).

Por fim, a fase de “elaboração do relatório” consiste na definição de recomendações para apresentação dos resultados identificados ao final da realização do estudo de caso. Dentre essas recomendações, destacam-se a necessidade de apresentar transparência e confiabilidade tanto nos dados como também na execução do estudo de caso. Referente a essa fase, o resultado a ser apresentado será o modelo de interoperabilidade de dados em saúde proposto, bem como todos os artefatos que nele existirem.

De forma ilustrativa, com o objetivo de melhor visualizar o percurso metodológico dessa pesquisa, este encontra-se apresentado na Figura 1.

Figura 1 - Percurso Metodológico



Fonte: Elaborado pelo Autor (2022)

O percurso metodológico da pesquisa debruça-se a princípio sobre três conceitos, quais sejam, interoperabilidade, princípios FAIR e metadados, além dos dados em saúde em bases digitais. Os dois primeiros elementos são imprescindíveis para a elaboração de um modelo de interoperabilidade de dados em saúde com base nos conceitos de interoperabilidade disponíveis nos princípios FAIR, e o terceiro corresponde a um elemento presente no modelo proposto.

3 INTEROPERABILIDADE NA ÓTICA DE DADOS EM SAÚDE

A compreensão pelo conceito da interoperabilidade apresenta-se como elemento essencial para o desenvolvimento de um padrão de interoperável de dados em saúde. Entender conceitualmente e conhecer os padrões nacionais e internacionais de interoperabilidade de dados em saúde, caracteriza-se como sendo um dos pilares conceituais de maior relevância para o desenvolvimento desta pesquisa.

3.1 CONCEITOS E DEFINIÇÕES

A busca pela compreensão de conceitos bases para construção do conhecimento necessário para o desenvolvimento de uma pesquisa científica, é de extrema relevância para o seu sucesso. Nesse sentido, o tópico em questão discorre sobre conceitos fundamentais para a o desenvolvimento dessa pesquisa, sendo eles os conceitos sobre dados, informações e interoperabilidade.

3.1.1 Dados e Informação

O conflito conceitual entre dado e informação surgiu há bastante tempo e estende-se até a atualidade. Para muitos autores a diferença está na percepção de valor, ou seja, quando um dado possui um significado e esse agrega valor com a sua aplicação, o mesmo pode ser entendido como sendo uma informação.

Considerando exclusivamente o escopo desta pesquisa, é oportuno destacar que mesmo sendo o dado o seu objeto de interesse, vê-se como relevante a percepção de valor que a informação traz para a sociedade.

A informação é reconhecida como um bem de grande valor para a sociedade moderna. Essa afirmação alicerça-se no fato de que, segundo Neves, Santos e Guimarães (2019), a informação está presente em todas as atividades do cotidiano, independentemente de sua complexidade e/ou impacto para a sociedade, podendo, inclusive, possibilitar a criação de novas informações.

Essa definição de valor da informação teve, e continua tendo, influência direta dos avanços tecnológicos que fornecem suporte à criação, disseminação, compartilhamento, recuperação e consumo dessas informações de forma bem mais ágil e acessível a boa parte da população.

Nesse sentido, de acordo com Delfino, Sousa e Pinho Neto (2019, p8):

O grande número de meios de comunicação existentes possibilita a produção e disseminação de um abundante volume de informações, proporcionando uma sobrecarga de informações que resulta em diversas mudanças de comportamento dos integrantes da sociedade da informação. Associado a esse grande número de canais de comunicação ainda pode se mencionar o fato de a internet permitir uma integração em tempo real independente de questões geográficas que outrora eram entendidas como barreiras ao fluxo da informação.

Essa sobrecarga informacional, originada através dos diversos meios de comunicação existentes, expõem a sociedade contemporânea a uma série de novos desafios que estão conectados a necessidade de um tratamento informacional que proporcione a disseminação e consumo de informações de qualidade.

Outra abordagem para a definição de valor de uma informação, de acordo com Delfino, Sousa e Pinho (2019), está diretamente relacionada a necessidade informacional e o esforço despendido pelo usuário na tentativa de recuperar essa informação e com isso suprir as lacunas existentes. Essa definição leva em consideração o grau de dificuldade para o usuário, na tentativa de realizar a recuperação dessa informação. Porém, na atualidade, o acesso às informações estão cada vez mais facilitadas, ficando, na maioria dos casos, o desafio em identificar quais informações são de qualidade e quais devem ser descartadas.

O crescimento do volume de informações disponíveis e, conseqüentemente, a dificuldade em identificar quais informações devem ser consumidas e quais devem ser descartadas, está diretamente ligado aos avanços das tecnologias. Nesse sentido, partindo do princípio de que equipamentos produzem dados e não informações, o desenvolvimento de uma pesquisa que objetive a análise e tratamento de dados, inclusive proporcionando a redução do volume de dados através da interoperabilidade, torna-se extremamente relevante para a sociedade contemporânea, em especial, em se tratando de dados disponibilizados no contexto da saúde.

Muitos são os conceitos identificados para esses termos, conforme pode ser identificado através da definição de Chaffey e Wood (2005), as informações são dados que proporcionam valor ao entendimento de um conteúdo. Tal afirmação faz uma relação do conceito de informação com o valor da utilização do dado, para compreensão de um determinado assunto. Percebe-se que o nível de tratamento do dado, para ser considerado uma informação, deve ser suficiente para ser aplicado em um contexto e com isso proporcionar valor para um usuário.

Seguindo com esse tema, Bader (1993), afirma que as informações são dados que foram rigorosamente selecionados, refinados, validados e formatados para auxiliar indivíduos e grupos a desempenhar suas responsabilidades específicas. Para Chaffey e Wood (2005), corroborando com Bader (1993), a relação entre informações e dados está condicionada à possibilidade de uso desse dado para desempenhar alguma atividade, tomar decisões ou outras formas de aplicação.

Percebe-se na definição desses autores que o nível de tratamento dos dados, para que esses possam ser aceitos como informações, passa de um nível simplesmente elementar. Esse tratamento deve ser suficiente para que esse dado possa ser utilizado em alguma atividade, gerando valor.

Ainda referente à relação entre dado e informação, Semeler e Pinto (2019, p11) afirmam que “dados serão considerados como todo objeto criado em formato digital (digital-born) ou convertido para o formato digital (digitalizados) que possa ser usado para geração de insights de informação e conhecimento”.

Para essa pesquisa, que tem em seu cerne o estudo do dado em seu formato digital, o conceito a ser utilizado defende a ideia de que “máquinas produzem dados, usuários consomem informação”, ideia essa, presente na compreensão de Setzer (2015, *online*) referente ao relacionamento conceitual entre dado e informação onde ele destaca que “não é possível processar informação diretamente em um computador, para isso é necessário reduzi-la a dados”.

Nesse sentido, o conceito apresentado por Semeler e Pinto, que está alinhando com a compreensão de Setzer, é o que se mostrou mais adequado.

Ainda na temática referente a conceitos de dados e informação, após analisar os conceitos aqui expostos, entende-se que uma informação se refere a um dado, ou conjunto de dados, que, na posse de um usuário, gere valor na compreensão de um conteúdo ou que responda a um questionamento específico de forma clara e objetiva.

Diante do exposto, fica evidente a importância que o dado tem em um contexto informacional, pois ele é a base para a criação de informações além de ser de elemento essencial para fluxo informacional digital, realizado entre recursos de tecnologia, a exemplo do que ocorre na internet.

Referente ao tema dados e informações, para o desenvolvimento dessa pesquisa, vê-se como relevante compreender o conceito do dado aplicado no contexto da saúde.

3.1.2 Metadados

Na Ciência da Informação existem diversos conceitos sobre metadados e padrões de metadados, a exemplo do apresentado por Formenton *et al.* (2017), que destacam que metadados são elementos (dados) alicerçados na biblioteconomia, cuja principal função consiste na descrição de um recurso informacional inequívoco, com múltiplas formas de acesso, garantindo sua recuperação pelo usuário final ou sistemas informacionais convencionais ou digitais.

Corroborando com essa mesma ideia Alves (2010) contribui com a seguinte definição para metadados:

[...] atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. [...] são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades [...] com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação. (ALVES, 2010, p. 47).

Analisando esses conceitos, é possível inferir que o objetivo da utilização dos metadados implica tornar estes dados em conteúdos legíveis (interpretáveis), em contextos e momentos específicos, diferente daqueles em que foram criados, através da identificação única do dado, viabilizando a encontrabilidade deste a fim de possibilitar seu uso e/ou reuso por usuário ou sistema de informação que dele necessitar.

Analisando as definições sobre metadados e interoperabilidade, é possível identificar uma relação conceitual muito próxima entre esses termos chegando, inclusive, a identificar uma relação de dependência onde os metadados passam a ser elemento essencial para o processo de interoperabilidade, pelo menos no que se refere à interoperabilidade em nível semântico.

A interoperabilidade em nível semântico, ou apenas interoperabilidade semântica, pode ser compreendida como nível mais alto de interoperabilidade, possibilitando que todos os sistemas envolvidos sejam capazes de trocar dados, interpretá-los e fazer uso desses dados para a geração de insights para os usuários (MIYOSHI 2018).

Para esse nível de integração, onde se almeja que todo o processo seja realizado entre os sistemas envolvidos, sem a necessidade de intervenção de usuários, vê-se como necessário a utilização de padrões de metadados que possibilitem uma correta identificação e recuperação desses dados.

Ainda sobre estruturas de metadados Barbedo, Corujo e Sant'Ana (2011), destacam múltiplas categorias funcionais de metadados que podem ser contextualizadas como:

1. Metadados descritivos (identificação) – Seu foco se concentra na pesquisa, recuperação e identificação. Dentre outros elementos destacam-se: título, autor, assunto e palavras-chave.
2. Metadados estruturais – possibilitam a formação de uma hierarquia entre objetos digitais distintos (textos, imagens, áudios etc.) pertencentes de um mesmo documento ou recurso informacional.
3. Metadados administrativos – composto por elementos que viabilizam a gerência dos recursos arquivísticos eletrônicos. Possibilitam, entre outras ações, a definição da forma e em que ocasião os recursos foram produzidos, tipo de arquivo e demais informações técnicas, além de informações de segurança de acesso aos dados.
4. Metadados técnicos – disponibilizam elementos referente aos aspectos técnicos dos arquivos e dos seus formatos.
5. Metadados de preservação – contém elementos essenciais ao processo de arquivamento e guarda dos objetos digitais ao longo do tempo.

Ciente da importância que cada categoria possui, porém, levando em consideração os propósitos em destaque nessa pesquisa, a categoria de metadados de preservação será melhor explorada ao longo dessa seção.

A opção pela categoria de metadados de preservação se justifica, entre outros fatores, por essa categoria possuir elementos fundamentais para o alcance de uma gerência eficiente de dados e a sua preservação a longo prazo, além da garantia de integridade, autenticidade e confiabilidade dos dados.

Outro ponto que merece destaque, em especial no que se refere à necessidade de preservação de dados ao longo do tempo, consiste na natureza e particularidade dos dados em saúde. De acordo com a resolução 2.218/2018 do CFM o tempo de guarda de dados em saúde que compõem o prontuário dos pacientes é de 20 anos a contar da última consulta realizada pelo paciente. Diante do exposto, fica evidente a importância de uma estrutura de metadados que possibilite a guarda e recuperação desses por um longo período.

Destaca-se ainda que a escolha pela utilização da categoria de preservação não exclui a utilização dos conceitos presentes nas demais categorias de metadados de forma híbrida na

busca pela melhor opção de definição de um padrão de metadados adequado para o modelo de interoperabilidade proposto nessa pesquisa.

Referente aos padrões e esquemas de metadados, podem ser compreendidos como sendo um conjunto de elementos planejados para uma finalidade específica a exemplo da descrição de determinado recurso informacional. Vale ressaltar que na literatura, de acordo com Chan e Zeng (2006), expressões do tipo “conjunto de elementos”, “esquema”, “sistemas” e “formato” vêm sendo utilizadas indistintamente para se referir a padrões de metadados.

Nesse sentido, Zeng e Qin (2008, p25) afirmam que o esquema de metadados são:

Uma especificação processável por equipamentos que define a estrutura, a codificação de sintaxe, regras, e formatos para um conjunto de elementos de metadados em uma linguagem formal num esquema. Na literatura o termo “metadado *schema*” usualmente refere-se ao conjunto de elementos na sua totalidade, assim como a codificação dos elementos e a estrutura com uma linguagem de marcação.

De acordo com a definição apresentada, é possível compreender que o padrão de metadados se refere a uma entidade contendo itens semânticos e de conteúdo. Nessa estrutura, além dos itens mencionados, o esquema em questão também deverá disponibilizar informações sobre codificação dos itens como uma sintaxe ou linguagem de marcação (Ex. *Standard Generalized Markup Language* (SGML) e *Extensible Markup Language* (XML)) (FORMENTON *et al.*, 2017).

Dessa forma, é possível inferir que a estrutura de metadados disponível no esquema/padrão de metadados possui ao menos dois componentes básicos, sendo eles, a semântica, com as definições ou significados dos elementos e seus refinamentos, e o conteúdo, este último com as declarações ou as instruções referentes aos valores que devem ser atribuídos para os elementos.

Para o desenvolvimento dessa pesquisa, na etapa de revisão e análise bibliográfica, evidenciou-se a existência de diversos padrões ou esquemas de metadados disponíveis para contextos específicos. Dentre as diversas opções de padrões identificados, optou-se por escolher aqueles que apresentam alinhamento com os objetivos dessa pesquisa. Em sua grande maioria, os padrões aqui apresentados foram identificados como sendo de uso para fins de preservação digital.

Padrão Dublin Core (DC) – Criado em Chicago, durante a 2ª Conferência Internacional World Wide Web, em 1994, ao longo de uma discussão sobre semântica e a Web com foco nos desafios da descoberta de dados. Posteriormente, em um evento em Dublin, Ohio em 1995,

mais de 50 profissionais discutiram sobre os benefícios de um conjunto semântico para a busca e a recuperação de dados baseados na Web. Como resultado foi criado um conjunto básico de elementos de metadados denominado de “*Dublin Core*” (DUBLIN CORE METADATA INITIATIVE, 2015).

Dentre as características presentes no padrão DC, destacam-se sua simplicidade, interoperabilidade semântica, consenso internacional, extensibilidade e flexibilidade. (GRÁCIO, 2012).

Outro ponto de destaque referente à utilização do DC, consiste no fato deste ser um dos padrões mais utilizados por comunidades internacionais para descoberta de recurso no domínio da Web, além de ser codificado em *HyperText Markup Language* (HTML), ou em *Resource Description Framework* (RDF) usando XML o que o torna muito eficiente para interoperabilidade na Web.

Padrão *Metadata Object Description Schema* (MODS) – O MODS disponibiliza informações bibliográficas significativas, que servem de apoio a outros padrões que utilizam XML. Referente à preservação digital, três características do MODS destacam-se: Informação de Origem, Condição de Acesso e Item Relacionado. Estes elementos registram informações úteis para auxílio de outros padrões de metadados de preservação no apoio a autenticidade e integridade dos dados. Tais características são extremamente importantes em se tratando, especialmente, de dados em saúde (BARBEDO *et al.* 2007).

Padrão *American National Standards Institute* (ANSI)/*National Information Standards Organization* (NISO) Z39.87 – Conhecido como padrão de metadados técnicos, que tem suas principais funcionalidades centradas no desenvolvimento, interoperabilidade e interpretação de arquivos de imagem digital, facilitando a comunicação entre sistemas, serviços e software, facilitando a interação e a utilização e reutilização de dados entre esses equipamentos, assim como apoiar a gestão ao longo do tempo e o contínuo acesso a dados de imagens digitais (BARBEDO *et al.* 2007).

Ratificando o citado anteriormente, a quantidade de padrões disponíveis na literatura é muito vasta, o que inviabiliza a discussão sobre todos esses padrões, dessa forma, após revisar a literatura e analisar as principais características, foram escolhidos esses três padrões para embasar a implantação de um elemento de metadados como parte do modelo de interoperabilidade de dados em saúde com base nos princípios FAIR que está sendo proposto.

Vê-se como oportuno destacar que não faz parte dos objetivos dessa pesquisa o desenvolvimento de um padrão de metadados para dados em saúde, ficando essa sugestão para ser contemplada em desenvolvimento de trabalhos futuros, porém, vê-se como relevante a

definição de um padrão de metadados (seja um padrão disponível no mercado em formato original ou modificado, ou até mesmo um novo padrão) como elemento que compõe o modelo de interoperabilidade de dados em saúde, esse sim cerne dessa pesquisa.

3.1.3 Interoperabilidade

O grande volume de dados produzidos e consumidos periodicamente, proporciona para a sociedade contemporânea muitos desafios, dentre eles destacam-se a necessidade de produção, compartilhamento e consumo desses dados.

Quando analisamos isso no contexto da saúde, outra característica ganha ainda mais evidência, a possibilidade de troca de dados entre sistemas de informação. A essa característica é dado o nome de interoperabilidade, que segundo HIMSS (2021), consiste na capacidade de sistemas, dispositivos e aplicativos de acessar, trocar, integrar e usar dados cooperativamente de maneira coordenada. Entre outros benefícios, a interoperabilidade proporciona a portabilidade dos dados de forma segura, em tempo hábil e em qualquer lugar.

A interoperabilidade dos dados produzidos nos sistemas da área da saúde ainda não é uma realidade. Contribuindo com essa afirmação Brasil (2016, p.36) destaca que:

Os dados produzidos no contexto da saúde ainda não dispõem de uma interoperabilidade entre seus sistemas informatizados capaz de permitir a visualização integral de dados que possibilite a transparência das ações governamentais de saúde e o acompanhamento e regulação das ações do sistema de saúde privado.

A busca pela interoperabilidade completa dos dados no contexto da saúde, é um desafio que envolve diversas variáveis, que vão desde uma completa heterogeneidade desses dados, que pode ser vista analisando os vários padrões existentes, e seguem pela necessidade de manutenção da segurança e privacidade desses dados.

Referente a essa heterogeneidade dos sistemas de informação na área da saúde, Freire, Meirelles e Cunha (2019), afirmam que o tratamento dos dados na área da saúde, através de sistemas de informação, enfraquece as ações caso o dado seja tratado de modo isolado.

Contribuindo com essa afirmação, Brasil (2016) destaca que a falta de padronização dos procedimentos para obtenção e tratamento dos dados em saúde e o elevado número de sistemas de informação e sua heterogeneidade são grandes desafios para o desenvolvimento de informações de qualidade para as organizações.

O desenvolvimento de um projeto de interoperabilidade de dados necessita, entre outros conhecimentos, de uma compreensão mais aprofundada do conceito de interoperabilidade. Nesse sentido, é oportuno destacar que, segundo Miyoshi (2018), a interoperabilidade pode ser categorizada em três níveis:

Interoperabilidade Fundamental – Caracterizada pela troca de dados entre sistemas de informação, porém, sem que esses sistemas tenham a capacidade de interpretar os dados compartilhados entre si. Nessa situação ocorre apenas o envio e recebimento dos dados entre os sistemas envolvidos, sem que seja aplicado qualquer tipo de tratamento a esse dado. Um exemplo da interoperabilidade fundamental pode ser visto na simples troca de exames de imagens ou em formato PDF.

Interoperabilidade Estrutural – Considerada como sendo um nível intermediário de interoperabilidade, os dados trocados entre os sistemas envolvidos já possuem um formato estruturado com uma sintaxe bem definida. Nessa estrutura, o objetivo operacional ou clínico para a troca de dados é preservado possibilitando que os sistemas envolvidos possam compreender o conteúdo desse dado e com isso realizar atividades como leitura e quantificação.

Interoperabilidade Semântica – Nível mais alto de interoperabilidade, possibilitando que todos os sistemas envolvidos sejam capazes de trocar dados, interpretá-los e fazer uso desses dados para a geração de *insights* para os usuários. Na interoperabilidade semântica, além dos dados trocados seguirem uma sintaxe bem definida, o seu conteúdo é codificado, garantindo assim a segurança, utilizando vocabulário controlado, terminologia ou ontologia.

O nível de interoperabilidade almejado para o resultado dessa pesquisa alinha-se com o conceito de interoperabilidade semântica, uma vez que elementos relacionados à segurança dos dados, vocabulário controlado e terminologias estão alinhados com os conceitos presentes nos princípios FAIR, base conceitual do padrão de interoperabilidade proposto.

Vale ressaltar que, a partir de 2009, diversos trabalhos científicos passaram a adotar o termo HIE (*Health Information Exchange*) como sendo sinônimo do termo interoperabilidade de dados em saúde. Essa opção surgiu inicialmente no contexto americano incentivada pela lei *Health Information Technology for Economic and Clinical Health* (HITECH) (AKHLAQ *et al*, 2016).

Ainda segundo o autor, o termo HIE pode ser definido formalmente como:

A mobilização eletrônica de informações clínicas e administrativas dentro ou entre organizações, e entre vários sistemas de informação, baseados em padrões reconhecidos, preservando a autenticidade dos dados trocados possibilitando com isso a tomada de decisões para uma melhor qualidade da assistência à saúde de

um paciente e da população de forma geral (AKHLAQ *et al.*, 2016, p.13) [tradução nossa].

Outra variável relevante para a compreensão da interoperabilidade de dados em saúde está diretamente relacionada aos avanços da tecnologia, a exemplo da nova onda de dispositivos que tornaram pervasiva a coleta de dados, em especial, dados de bem-estar e saúde, relacionados à internet das coisas.

3.2 NORMATIZAÇÃO DE DADOS EM SAÚDE

O uso de dados em saúde envolve uma série de fatores que demandam atenção dos órgãos reguladores em especial no que se refere à qualidade dos dados utilizados principalmente por se tratar de dados relacionados à saúde individual e coletiva. Além da qualidade, a segurança e a privacidade também se mostram fatores de muita relevância para o compartilhamento de dados em saúde.

Nesse sentido, muitas legislações foram criadas para regulamentar a garantia da saúde para toda a sociedade brasileira bem como o uso correto e adequado dos dados em saúde para as finalidades específicas relacionadas ao bem-estar individual e coletivo através do desenvolvimento de políticas públicas.

Ao falar em legislação, é imprescindível analisar cronologicamente o desenvolvimento dessas leis, a fim de compreender a sua contextualização ao longo do tempo e levando em consideração as evoluções impostas a sociedade. Nesse sentido, um dos primeiros relatos sobre regulamentação da saúde como sendo um direito base da sociedade brasileira pode ser vista no texto da Constituição da República Federativa do Brasil de 1988, que traz em seu artigo 196 a seguinte redação:

Art. 196. A saúde é direito de todos e dever do Estado, garantido mediante políticas sociais e econômicas que visem à redução do risco de doença e de outros agravos e ao acesso universal e igualitário às ações e serviços para sua promoção, proteção e recuperação [BRASIL, 1988].

Além da Constituição Federal, a lei número 8.080 de 19 de setembro de 1990, destaca pela regulamentação da criação do Sistema Único de Saúde (SUS), conforme redação do seu artigo 4 que descreve:

Art. 4º O conjunto de ações e serviços de saúde, prestados por órgãos e instituições públicas federais, estaduais e municipais, da Administração direta e indireta e das fundações mantidas pelo Poder Público, constitui o Sistema Único de Saúde (SUS) [BRASIL, 1990].

O SUS consolida-se como sendo um grande avanço para o desenvolvimento de ações em prol do desenvolvimento da saúde no Brasil. Seus princípios fundamentam-se nas diretrizes previstas no artigo 198 da Constituição Federal.

Dentre esses princípios, disponíveis no artigo 7 da lei 8.080, alinhado com os objetivos dessa pesquisa, destacam-se o V – direito à informação, às pessoas assistidas, sobre sua saúde e VI – divulgação de informações quanto ao potencial dos serviços de saúde e a sua utilização pelo usuário.

Seguindo com essa temática, referente ao direito a informação e divulgação dessas informações de saúde vê-se como oportuno analisar a estrutura de organização desses dados e/ou informações com o objetivo de proporcionar um compartilhamento priorizando a eficácia e segurança. Nesse sentido destaca-se a lei 8.159 de 8 de janeiro de 1991 que dispõe sobre a política nacional de arquivos públicos e privados.

É possível afirmar que a grande contribuição da lei 8.159 de 8 de janeiro de 1991, constitui a criação do Conselho Nacional de Arquivos (CONARQ), conforme redação disponível no artigo 26 da referida lei.

Art. 26 - Fica criado o Conselho Nacional de Arquivos (CONARQ), órgão vinculado ao Arquivo Nacional, que definirá a política nacional de arquivos, como órgão central de um Sistema Nacional de Arquivos (SINAR) [BRASIL, 1991].

Entrando mais detalhadamente no contexto de relevância dessa pesquisa, que consiste na pesquisa da interoperabilidade de dados em saúde, após citar a lei de criação do CONARQ, destaca-se a criação da portaria número 2.073, de 31 de agosto de 2011 do Ministério da Saúde que regulamenta o uso de padrões de interoperabilidade e informações em saúde para sistema de informação no âmbito do Sistema Único de Saúde - SUS, sistemas privados e do setor de saúde suplementar.

A referida portaria foi criada com o intuito de normatizar padrões de interoperabilidade de informações em saúde através do desenvolvimento de um conjunto mínimo de premissas, políticas e especificações técnicas que disciplinam o intercâmbio de informações entre os sistemas de saúde Municipais, Distrital e Estaduais proporcionando a interação entre esses e a sociedade (BRASIL, 2011).

Dentre os objetivos presentes no artigo 2 da portaria 2.073, destacam-se:

- I - definir a representação de conceitos a partir da utilização de ontologias, terminologias e classificações em saúde comuns, e padrões padronizados de representação da informação em saúde, criar e padronizar formatos e esquemas de codificação de dados, de forma a tornar célere o acesso a informações relevantes, fidedignas e oportunas sobre o usuário dos serviços de saúde;
- II - promover a utilização de uma arquitetura da informação em saúde que contemple a representação de conceitos, conforme mencionado no inciso I, para permitir o compartilhamento de informações em saúde e a cooperação de todos os profissionais, estabelecimentos de saúde e demais envolvidos na atenção à saúde prestada ao usuário do SUS, em meio seguro e com respeito ao direito de privacidade;
- III - contribuir para melhorar a qualidade e eficiência do Sistema Único de Saúde e da saúde da população em geral;
- IV - fundamentar a definição de uma arquitetura de informação nacional, independente de plataforma tecnológica de software ou hardware, para orientar o desenvolvimento de sistemas de informação em saúde;
- V - permitir interoperabilidade funcional, sintática e semântica entre os diversos sistemas de informações em saúde, existentes e futuros;
- VI - estruturar as informações referentes a identificação do usuário do SUS, o profissional e o estabelecimento de saúde responsáveis pela realização do atendimento;
- VII - estruturar as informações referentes aos atendimentos prestados aos usuários do SUS visando à implementação de um Registro Eletrônico de Saúde (RES) nacional e longitudinal; e
- VIII - definir o conjunto de mensagens e serviços a serem utilizados na comunicação entre os sistemas de informação em saúde

Outro ponto de destaque disponível no texto da portaria 2.073 faz referência ao Catálogo de Padrões de Interoperabilidade de Informações de Sistemas de Saúde (CPIISS), publicado pelo Departamento de Informática do SUS, disponível para a sociedade em geral.

Muitos são os padrões de interoperabilidade nacionais e internacionais disponíveis para utilização em estrutura de dados em saúde. Esse tema, basilar para o desenvolvimento dessa pesquisa, será melhor discutido na seção 3.3 dessa pesquisa, porém retornará a pauta na seção 6 que explora a construção de um modelo conceitual de interoperabilidade de dados em saúde.

3.3 PADRÕES DE DADOS EM SAÚDE

Para iniciar uma análise referente aos mais diversos padrões disponíveis para dados no contexto da saúde, vê-se como relevante a compreensão do termo “padrão”. Segundo Michaelis (2021, *online*), o termo padrão consiste em “padrão estabelecido cuja aprovação por consenso ou por autoridade oficial serve de base de comparação”. Nessa mesma linha, Lopes *et al.* (2018,

p.56) afirma que o termo padrão pode ser definido como sendo “um conjunto de regras e definições que especificam como executar um processo ou criar um produto”.

Diante do exposto, é correto afirmar que um padrão se refere a um conjunto de regras, previamente testadas e aprovadas, que orientam a execução de uma determinada atividade de forma homogênea e eficaz.

Remetendo essa reflexão ao contexto de dados em saúde, percebe-se que o desafio não está na ausência de padrões, e sim na existência de um gigantesco conjunto de padrões que funcionam muito bem de forma isolada, mais que ao longo da vida útil desses dados, eles precisam interagir para que seja possível obter o resultado esperado seja para realização de um diagnóstico de um paciente como também o desenvolvimento de políticas públicas para toda uma sociedade.

O desejo pelo desenvolvimento de um prontuário eletrônico único motivou o CTI-PRC no desenvolvimento de um conjunto mínimo de dados clínicos do paciente, necessários para a interoperabilidade entre sistemas de informação em saúde, que fornecem dados para a base nacional do Sistema Único de Saúde – SUS.

Como resultado dessa iniciativa, foram produzidas algumas recomendações, destacando um conjunto essencial de informações e *Document Type Definition* - DTDs que, posteriormente, foram utilizadas no projeto do primeiro Cartão Nacional de Saúde.

Outra iniciativa importante referente a questões de padronização e interoperabilidade de dados em saúde foram a criação da Política Nacional de Informação e Informática em Saúde - PNIIS apresentada e aprovada na 12ª Conferência Nacional de Saúde em 2003. A PNIIS foi atualizada com o passar do tempo, tendo sua última versão liberada em 2016 e trazendo como uma de suas diretrizes:

Estabelecimento de um padrão para e-Saúde que permita a construção do Registro Eletrônico de Saúde (RES) do cidadão por meio da identificação unívoca de usuários, profissionais e estabelecimentos de saúde, padrões e protocolos de interoperabilidade eletrônica e/ou digital entre os equipamentos/sistemas”. (BRASIL, 2016)

Ainda de acordo com Brasil (2016), a PNIIS tem como propósito:

Promover o uso inovador, criativo e transformador da tecnologia da informação a fim de melhorar os processos de trabalho em saúde e, assim, resultar em um Sistema Nacional de Informação em Saúde (SNIS) articulado e que produza informações para os cidadãos, a gestão, a prática profissional, a geração de conhecimento e o controle social, garantindo ganhos de eficiência e qualidade mensuráveis através da

ampliação de acesso, equidade, integralidade e humanização dos serviços de saúde, contribuindo, dessa forma, para a melhoria da situação de saúde da população.

Analisando o conteúdo da PNIIS e do CTI-PRC, fica evidente a necessidade e interesse em desenvolver alternativas de interoperabilidade de dados em saúde com o objetivo de viabilizar o compartilhamento desses dados de forma segura, eficiente e eficaz, tanto para o paciente, titular desses dados, como também para os profissionais da saúde e para o governo que poderá usufruir desses dados de forma mais assertiva para o desenvolvimento de políticas públicas.

Como mencionado anteriormente, um dos desafios para se alcançar a interoperabilidade plena se evidencia na existência de uma infinidade de padrões utilizados para troca de dados entre sistemas e dispositivos na área da saúde. Dentre esses diversos padrões disponíveis, de acordo com Lopes *et al.* (2018), podemos destacar:

Digital Imaging and Communications in Medicine – DICOM: Padrão utilizado para manipulação, transmissão, armazenamento e impressão de imagens médicas. O padrão DICOM contém uma definição de formato de arquivo e um protocolo de comunicação que permite a troca de arquivos entre dispositivos habilitados para esse padrão, a exemplo de scanners, servidores, estações de trabalho, impressoras e hardware de rede de múltiplos fabricantes em um arquivamento de imagens e sistema de comunicação – PACS. O padrão DICOM é mantido pela *National Electrical Manufacturers Association – NEMA*;

Health Level 7 – HL7: Compreendido como sendo um conjunto de padrões internacionais para transferência de dados clínicos e administrativos entre softwares da área da saúde;

IHE-PIX: Padrão utilizado para identificação de pacientes possibilitando a troca desses dados com outros sistemas conectados. Se refere a um perfil que reconhece um conjunto de normas para identificação de pacientes, elaborado pelo *Integrating the Healthcare Enterprise (IHE)*.

Logical Observation Identifiers Names and Codes – LOINC: Base de dados composta por nomes e códigos que possibilitam a identificação de observações clínicas e laboratoriais, possibilitando, inclusive, sua utilização sem a obrigatoriedade de aquisição de licenças de uso. O padrão LOINC pode ser inserido em mensagens escritas no padrão HL7 com o intuito de viabilizar a troca e análise de dados de exames, proporcionando a integração de informações clínicas e laboratoriais.

Troca de Informações na Saúde Suplementar – TISS: Definido como um padrão obrigatório para as trocas eletrônicas de dados de atenção à saúde de usuários de planos de saúde, buscando proporcionar à interoperabilidade entre os agentes da saúde suplementar. Tem como objetivo a padronização de ações administrativas referentes à avaliação e ao acompanhamento econômico, financeiro e assistencial das operadoras de planos privados de assistência à saúde.

Open electronic health records - OpenEHR: Se refere a uma especificação de padrão aberto em informática em saúde contendo a descrição dos processos de gerenciamento e armazenamento, recuperação e troca de dados de saúde entre registros eletrônicos de saúde (EHRs). Para o padrão OpenEHR, todos os dados de saúde de um paciente devem ser armazenados em um EHR por toda a vida, independente do fornecedor.

Fast Healthcare Interoperability Resources - FHIR: Padrão de interoperabilidade criado e mantido pela organização HL7, tendo como um de seus objetivos a disponibilização de dados de saúde para indivíduos e profissionais de saúde, considerando uma ampla variedade de dispositivos, a exemplo dos computadores, tablets e telefones celulares. Além do exposto o projeto do FHIR, utiliza um conjunto moderno de API baseado na Web, destacando o protocolo RESTful baseado em HTTP e uma opção de JSON, XML ou RDF para representação de dados, o que facilita a interoperabilidade entre sistemas legados.

Dentre os muitos padrões de interoperabilidade de dados disponíveis, com o intuito de orientar o desenvolvimento de iniciativas de interoperabilidade e integração de sistemas para o favorecimento do compartilhamento de dados entre os sistemas de informação, públicos e privados que compõem o SUS, o CPIISS orienta quanto ao tipo de padrão a ser utilizado para determinados tipos de informação, conforme ilustra a quadro 1.

Quadro 1 - Tipo de informação em saúde e padrão utilizado

TIPO DE INFORMAÇÃO	PADRÃO
Registro Eletrônico em Saúde (RES)	OpenEHR
Resultados e Solicitações de Exames	HL7 - Health Level 7
Codificação de Termos Clínicos e Mapeamento das Terminologias Nacionais e Internacionais	SNOMED-CT
Sistemas de Saúde Suplementar	TISS (Troca de Informações em Saúde Suplementar)
Arquitetura do Documento Clínico	HL7 CDA
Exames de Imagem	DICOM

Exames Laboratoriais	LOINC (<i>Logical Observation Identifiers Names and Codes</i>)
Dados de Identificação das Etiquetas de Produtos Relativos ao Sangue Humano, de Células, Tecidos e Produtos de Órgãos	ISBT 128
Padrões de Conhecimento, Incluindo Arquétipos, <i>Templates</i> e Metodologia de Gestão	ISO 13606-2
Cruzamento de Identificadores de Pacientes	IHE-PIX (<i>Patient Identifier Cross-Referencing</i>)

Fonte: Elaborado pelo Autor (2022)

Como mencionando anteriormente, os padrões utilizados como base no desenvolvimento do padrão proposto nessa pesquisa serão explorados com mais detalhes na seção específica sobre o desenvolvimento do modelo conceitual de interoperabilidade de dados em saúde.

Ainda sobre padrões de dados em saúde, como mencionado nessa seção, um registro eletrônico de saúde é composto por diversos tipos de dados onde cada um possui um padrão específico e isso eleva a complexidade de um projeto de interoperabilidade de dados em saúde.

Diante do exposto, essa pesquisa se propõe a analisar alternativas de interoperabilidade através da possibilidade de identificação, armazenamento e recuperação desses dados através de sua estrutura de metadados.

Como alternativa para essa estrutura de metadados, está sendo proposto a utilização dos princípios FAIR que serão descritos na sequência.

4 PRINCÍPIOS FAIR

Com o surgimento do movimento FAIR em janeiro de 2014 e da publicação de seus princípios em março de 2016, é possível identificar um crescimento significativo no interesse mundial em conhecer e aplicar esses princípios na busca de uma melhor gestão de dados, possibilitando que eles possam ser mais facilmente encontráveis, acessíveis, interoperáveis e reusáveis (SANTOS, 2021).

Os princípios foram criados com o propósito principal de apresentarem um conjunto de elementos esperados para os dados a fim de torná-los mais suscetíveis à atuação de sistemas computacionais. Destacam a capacidade de ação da máquina, ou seja, a capacidade dos sistemas computacionais de encontrar, acessar, interoperar e reutilizar dados com a mínima ou nenhuma intervenção humana.

O desejo pela busca de melhorias na recuperação desses dados para reutilização em novas pesquisas científicas ocorre principalmente pelo fato de os humanos dependerem cada vez mais do suporte computacional para lidar com os dados devido ao aumento no volume, complexidade e velocidade de criação de dados (GO FAIR, 2021).

4.1 VISÃO GERAL

As manifestações iniciais a respeito dos princípios FAIR surgiram em 2014 na conferência *Jointly Designing a Data FAIRPORT*. Nesse evento, organizado pelo grupo Elixir, *Netherlands eScience Center*, *Dutch Techcentre for Life Sciences* (DTL) e o *Centro Lorentz* foram discutidos temas relacionados aos desafios de utilização e reutilização de dados e as propostas para solução desse problema (DATA FAIRPORT, 2014).

Como resultado dessa conferência, destaca-se o consenso a respeito da necessidade de criação de uma infraestrutura global que possibilite o compartilhamento e a reutilização dos dados, bem como a discussão sobre requisitos para a disponibilização desses dados de forma aberta, como sendo um fomento ao desenvolvimento da gestão de dados.

As propostas aprovadas na conferência foram balizadores para a criação de um conjunto de princípios e práticas que fossem aplicadas na orientação da descoberta, acesso, integração e reutilização de uma ampla quantidade de dados produzida pelos pesquisadores, que ficou conhecido como princípios FAIR.

Destaca-se que o desenho inicial de criação dos princípios FAIR estava direcionado para o uso de dados científicos e, segundo FAIR Statement (2017), está sendo adotado em todo o

mundo para gerar dados de pesquisa localizáveis, acessíveis, interoperáveis e reutilizáveis. Dessa forma, é possível afirmar que a aplicação do FAIR não ficará restrita aos dados científicos, tendo sua aplicação experimentada em diversos segmentos, a exemplo da proposta dessa pesquisa.

Segundo Wilkinson (2016, p. 1), pode-se afirmar sobre os princípios FAIR:

Diferente de iniciativas de pares que focam o acadêmico humano, os Princípios FAIR enfatizam especificamente o aprimoramento da capacidade das máquinas de encontrar e usar os dados automaticamente, além de apoiar sua reutilização por indivíduos.

Corroborando com o tema, DTL (2020, *online*, tradução nossa) afirma que:

Além da coleta, anotação e arquivamento adequados, o gerenciamento de dados inclui o 'cuidado de longo prazo' dos dados de pesquisa, com o objetivo de que eles possam ser encontrados e reutilizados em estudos posteriores.

Dando continuidade, o referido autor define os conceitos de Localizável, Acessível, Interoperável e Reutilizável como sendo:

Localizável: fácil de identificar e encontrar para humanos e computadores, com metadados que facilitam a pesquisa de conjuntos de dados específicos,
Acessíveis: armazenados por longo prazo para que possam ser facilmente acessados e/ou baixados com condições de acesso bem definidas, seja no nível de metadados, seja no nível dos dados reais,
Interoperável: pronto para ser combinado com outros conjuntos de dados por humanos ou computadores, sem ambiguidades nos significados dos termos e valores,
Reutilizável: pronto para ser usado em pesquisas futuras e para ser posteriormente processado usando métodos computacionais. Isso requer informações adequadas sobre como os dados foram obtidos e processados (proveniência) e uma licença adequada (DTL, 2020, *online*, tradução nossa).

Como já citado, o FAIR constitui-se de um conjunto de princípios e práticas e que segundo SURF (2019, p.4, tradução nossa), esses princípios podem ser compreendidos não como um padrão, mas, como “[...] uma referência de preparação dos dados para a reutilização, de acordo com condições descritas por pessoas e máquinas”.

Após o surgimento do FAIR em 2014 e da publicação dos seus princípios em 2016, chegamos à consolidação dos princípios FAIR em 2019, com o primeiro encontro anual das redes de implementação da iniciativa *Global Open FAIR (GO FAIR)*, em Leiden, na Holanda,

recebendo 90 profissionais das áreas de ciência da informação e da ciência da computação de diversos países da Europa, Estados Unidos e do Brasil.

Esse evento teve como objetivo proporcionar uma aproximação da comunidade GO FAIR, afim de identificar possíveis estratégias de alinhamento e sincronização de seus projetos.

É oportuno destacar que o GO FAIR se refere a uma iniciativa dos governos da Holanda e Alemanha, posteriormente, do governo da França, que ao perceberem o crescimento na utilização do FAIR identificaram alguns desafios, dentre eles, como evitar o desperdício de recursos e a redução do risco de implementações incompatíveis?

Como alternativa de resposta para essas questões, esses governos, juntamente com o governo da França, se comprometeram a financiar a criação de uma entidade com o objetivo de apoiar e coordenar o desenvolvimento de projetos seguindo os princípios FAIR. Dessa forma, nasceu o movimento GO FAIR (SANTOS, 2021).

Os primeiros resultados desse projeto foram a criação de escritórios GO FAIR regionais no Brasil e nos Estados Unidos.

Ainda no ano de 2019, também em 2020, diversos outros eventos relacionados ao fomento da implantação dos princípios FAIR ocorreram, contribuindo com a solidificação dos princípios FAIR como sendo uma referência nacional e internacional para gestão de dados, sempre contando com um aumento significativo de participantes o que atesta a sua expansão pelo mundo.

De acordo com Sales *et al.* (2021), a participação brasileira sempre esteve presente junto à iniciativa GO FAIR, desde a época de criação do escritório GO FAIR Brasil, no final de 2018.

4.2 INICIATIVA GO FAIR E GO FAIR BRASIL

A iniciativa GO FAIR teve sua fundação em 2016, chegando ao Brasil no final de 2018, com a criação do GO FAIR Brasil, como sendo uma das primeiras instâncias desse projeto fora da Europa.

A GO FAIR tem em seu cerne o objetivo e desejo de coordenar e fomentar iniciativas de criação de redes de implementação e o desenvolvimento de projetos que promovam a gestão de dados garantindo a reutilização de dados em diferentes contextos, países e disciplinas, contribuindo para o compartilhamento e reuso de dados na geração de novos conhecimentos e para a reprodutibilidade da pesquisa (SALES *et al.* 2021).

A estrutura do GO FAIR foi idealizada considerando três bases essenciais; GO TRAIN: direcionada para atuar na realização e apoio da iniciativas de treinamentos de diferentes níveis

e naturezas, com o objetivo de fomentar a aplicação dos princípios FAIR nas práticas da ciência; GO CHANGE: com foco em ações de disseminação do conhecimento de gestão de dados através dos princípios FAIR, proporcionando mudanças culturais e políticas afim de viabilizar o reconhecimento dos FAIR como uma referência; GO BUILD: incentiva o desenvolvimento e crescimento da infraestrutura técnica e operacional necessária para suportar os dados FAIR.

Ainda sobre a GO FAIR Brasil, criando no final de 2018, com atuação em todas as áreas do conhecimento, teve seu primeiro encontro registrado no evento de comemoração de 20 anos da *Scientific Electronic Library Online (SciELO)*, momento em que o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) assumiu a coordenação geral do GO FAIR Brasil (GO FAIR, 2022).

Ainda no ano de 2018 foi realizado o lançamento da primeira rede brasileira, a GO FAIR Brasil Saúde, sob responsabilidade do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict/Fiocruz). Essa conta com a participação de várias instituições das áreas da saúde pública, vigilância sanitária, informação e comunicação em saúde, educação profissional em saúde, oncologia e enfermagem (FIOCRUZ, 2022).

A iniciativa GO FAIR Brasil, mantém-se altamente atuante com participação em diversos eventos e projetos realizados pela organização central da GO FAIR, a exemplo da participação da rede *Virus Outbreak Data Network (VODAN)*, que atua na busca de soluções ao combate do novo Corona vírus através utilização do aprendizado de máquinas e abordagens da inteligência artificial na descoberta de padrões relevantes para a gestão de dados em surtos epidêmicos.

4.3 PRINCÍPIO DE INTEROPERABILIDADE DO FAIR

Pensar em dados como item estratégico implica aceitar que esses dados devem ser localizáveis, acessíveis e reutilizáveis não apenas por indivíduos mais principalmente por máquinas.

Assim, Kelleher e Tierney (2018) destacam que cada vez mais a crescente busca pela disponibilização dos dados no lugar certo, no momento certo, no formato certo, sendo esses aderentes à necessidade informacional existente, o que coloca a integração dos dados como condição de destaque para sucesso dos processos de uso e reuso de dados.

O processo necessário para viabilizar a integração dos dados, com mencionado, é denominada de interoperabilidade que, segundo Dias *et al.* (2019) consiste na necessidade de

integrar dados a outros conjuntos de dados (considerando-se também sempre os metadados) e com as mais variadas aplicações ao longo do seu ciclo de vida.

Percebe-se, até aqui, que os atributos localizáveis, acessíveis e reutilizáveis apresentados como essenciais para o uso estratégico dos dados fazem uso da interoperabilidade como alternativa, entre outras coisas, para proporcionar a comunicação entre equipamentos em ambientes compostos por centenas ou mesmo milhares de sistemas, desenvolvidos interna ou externamente, compondo um cenário complexo e diversificado, todos requerendo integração de seus dados para que seu efetivo uso seja viabilizado (REEVE, 2013).

Diante do exposto, podemos afirmar que os princípios FAIR (Localizável, Acessível, Interoperável e Reutilizável) possuem uma relação prática e conceitual muito estreita, sendo necessário, em muitas ocasiões, mesclar conceitos para se alcançar objetivos específicos mesmo que direcionado para um dos princípios específicos.

Entendendo e aceitando esse entrelaçamento conceitual entre os princípios FAIR, vê-se como oportuno a exploração conceitual de todos os princípios mesmo que o foco mais amplo dessa pesquisa seja o princípio que abarca a interoperabilidade.

Observando o quadro 2, é possível identificar princípios orientadores do FAIR segundo Fiocruz (2022).

Quadro 2 - Descrição dos princípios FAIR

Grupo de Princípios	Princípio	Como Aplicar?
Findable (Localizável)	F.1 metadados devem ter identificadores globais, únicos e persistentes.	Adotar identificador único persistente tanto para o conjunto de dados quanto para os metadados (ex.: DOI, ARK, RRID, PID).
	F.2 dados devem ser descritos utilizando metadados ricos (impacta diretamente R1)	O conjunto de dados deve ser descrito por metadados ricos o suficiente para que, uma vez indexados em um mecanismo de busca possam ser encontrados mesmo sem o seu identificador único persistente.
	F.3 metadados devem incluir clara e explicitamente os identificadores dos dados que descrevem.	Como não podemos prever que os dados e seus metadados estejam sempre juntos, a associação entre eles deve ocorrer pela inclusão do identificador persistente dos dados nos metadados.
	F.4 Metadados devem ser registrados ou indexados em mecanismos de busca.	Para que os dados sejam encontrados, seus metadados devem ser indexados em mecanismos de busca (<i>search engine</i>), que possibilitem aos computadores e usuários encontrá-los com facilidade.
Accessible (Acessível)	A.1 Metadados devem ser recuperáveis pelos seus identificadores usando protocolo de comunicação padronizado.	Com o identificador persistente do conjunto de dados e/ou de seus metadados, o usuário deverá recuperá-los mais facilmente por meio de protocolos de comunicação padronizados. (ex.: HTTP ou FTP).

	A.1.1 O protocolo deve ser aberto, gratuito e universalmente implementável.	O protocolo de comunicação usado para dar acesso a eles deve ser aberto, gratuito e passível de ser implementado por qualquer interessado. (ex.: HTTP ou FTP).
	A.1.2 O protocolo deve permitir procedimentos de autenticação e autorização, quando necessário.	Dependendo das restrições de acesso aos dados e/ou metadados, um mecanismo de autenticação e autorização para o acesso deve ser liberado pelo protocolo de comunicação. (Ex.: os repositórios confiáveis oferecem essa opção).
	A.2 Metadados devem ser acessíveis, mesmo quando os dados não estiverem mais disponíveis.	É preciso existir um conjunto de estratégias de preservação para dados e metadados. Os metadados devem ser sempre acessíveis, possibilitando a criação de índices para o conjunto de dados atuais vigentes e aqueles não mais disponíveis.
Interoperable (Interoperável)	I.1 Metadados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento.	Para que se possa representar dados e metadados devem ser adotadas linguagens de representação do conhecimento que sejam padronizadas, acessíveis e amplamente aplicáveis. (Ex.: RDF, XML, DICOM, etc.).
	I.2 Metadados devem usar vocabulários de acordo com os princípios FAIR.	Dados e metadados devem possuir referências a vocabulários e/ou ontologias que os descrevem.
	I.3 Metadados devem incluir referências qualificadas para outros Metadados.	É necessário referenciar o conjunto de dados, possibilitando que aqueles gerados a partir de outros conjuntos, sejam interligados. Assegurando a ligação semântica entre ele.
Reusable (Reutilizável)	R.1 Metadados são descritos com uma pluralidade de atributos precisos e relevantes.	Prover metadados descritos com alto nível de detalhes que permitam ao pesquisador avaliar a possibilidade do seu reuso bem como adequação às suas necessidades.
	R.1.1 Metadados devem ser disponibilizados com licenças de uso claras e acessíveis.	É fundamental que o responsável pelos dados e metadados defina explicitamente quem pode ter acesso a eles, com que finalidade e sob quais condições. Essas informações são definidas por meio de licenças de uso.
	R.1.2 Metadados devem estar associados à sua proveniência.	Especificar a proveniência (linhagem) dos dados é importante não só para que o pesquisador possa avaliar a utilidade dos dados ou metadados, mas também para que possa atribuir o devido crédito a quem produziu, manteve ou editou esses dados.
	R.1.3 Metadados devem estar alinhados com padrões relevantes do seu domínio.	Além de atender aos padrões específicos da área de cada comunidade deve - se dar atenção as boas práticas de arquivamento e compartilhamento específicos da área de pesquisa.

Fonte: Adaptado de Fiocruz (2022)

Conforme ilustra a quadro 2, cada um dos quatro grupos de princípios FAIR apresenta princípios orientadores próprios sendo: quatro relacionados ao conceito Localizável, quatro ao

Acessível, três ao Interoperável e quatro ao Reutilizável. Destaca-se aqui que os princípios da Interoperabilidade são o foco deste tópico.

O grupo de princípios da Interoperabilidade contempla três princípios que, mesmo que apresentados de forma genérica, indicam diretrizes importantes para a viabilização da interoperabilidade de dados disponíveis em bases distintas.

Debruçando-se sobre esses princípios temos o primeiro “I.1 Metadados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento”. De forma conceitual recomendações como a necessidade de uma linguagem formal remete à necessidade de um formalismo, que por sua vez aponta para o uso de regras pré-definidas, o que favorece a mitigação de divergência de interpretações para o mesmo conteúdo.

Ainda sobre o primeiro princípio da interoperabilidade, os termos acessível e compartilhada, que por sua vez remetem as características de acessibilidade e compartilhamento, também merecem seus destaques. Tais características enfatizam que as regras utilizadas para a interoperabilidade devem estar disponíveis ao alcance de todos os envolvidos nesse processo, e mais, seu desenvolvimento deve ser fruto de uma ação estratégica que contemple todos os elementos envolvidos no processo de interoperabilidade almejado.

Por fim, ainda na análise do primeiro princípio, temos o elemento “amplamente aplicável” que remete à necessidade de viabilidade de execução de todas as regras definidas para o contexto da interoperabilidade. O critério de aplicabilidade se torna essencial quando pensamos na utilização prática desse princípio para projetos de interoperabilidade de dados.

Estendendo a análise para o segundo princípio “I.2 Metadados devem usar vocabulários de acordo com os princípios FAIR”, percebe-se que o foco se concentra nos vocabulários dos metadados. Considerando a fragmentação dos dados em saúde, necessários para compor um RES, a estrutura de vocabulários dos metadados torna-se um elemento complicador em um projeto de interoperabilidade de dados em saúde, podendo, inclusive, ser necessário o desenvolvimento de extensões desses vocabulários ou mesmo a criação de vocabulários novos (FORCE11, 2020).

Dando sequência à análise do terceiro princípio “I.3 Metadados devem incluir referências qualificadas para outros Metadados”. Evidencia-se aqui a necessidade de elementos na estrutura de metadados que possibilitem a interconexão desses dados. Esses elementos, de acordo com FORCE11 (2020), “devem ser sintaticamente analisáveis e semanticamente acessíveis por máquina”, possibilitando com isso uma interoperabilidade automatizada, sem a

necessidade de intervenção manual, possibilitando total autonomia das máquinas para realização desse processo.

Como mencionado, os princípios disponíveis da Interoperabilidade do FAIR, mesmo que genéricos, servem como base norteadora para o desenvolvimento de modelos de interoperabilidade de dados, similar ao que está sendo proposto como objetivo principal dessa pesquisa. Seguindo nessa linha, é importante destacar a importância dos metadados para os princípios FAIR, sendo esses elementos referenciados a todo momento, inclusive como elemento fundamental para o alcance dos objetivos propostos.

5 UMA PROPOSTA DE MODELO CONCEITUAL DE INTEROPERABILIDADE DE DADOS EM SAÚDE COM BASE NOS PRINCÍPIOS FAIR

A interoperabilidade de dados em saúde segue como sendo um dos grandes desafios para o pleno uso desses dados por diversos segmentos da sociedade. O desejo por uma interoperabilidade plena que possibilite, entre outros benefícios, a existência de uma base única de dados em saúde acessível, disponível, segura e confiável, configura-se como sendo um grande avanço para o desenvolvimento de diversos projetos benéficos para a sociedade atual.

Porém, ao passo que cresce o interesse pela interoperabilidade também, torna-se ainda mais evidente a que essa interoperabilidade não se configura como uma atividade simples. Questões relacionadas aos avanços da tecnologia da informação e ao crescimento exponencial no volume e diversidade de dados em saúde são alguns dos elementos que exercem influência direta no desenvolvimento de um modelo de interoperabilidade de dados em saúde.

5.1 TIPOS DE DADOS

Ao analisar a estrutura dos dados de um Registro Eletrônico de Saúde (RES) fica evidente a quantidade e diversidade de dados coletados de fontes diferentes, com estruturas diferentes e que precisam ser disponibilizadas em uma estrutura de dados centralizado.

Diante do mencionado, considerando o objeto dessa pesquisa, percebe-se como extremamente relevante o aprofundamento sobre alguns padrões de dados que possam servir de base para o desenvolvimento do modelo conceitual de interoperabilidade de dados em saúde sob luz dos princípios FAIR.

5.2 PADRÕES UTILIZADOS PARA CONSTRUÇÃO DO MODELO OPENFAIREHR

Padrões de interoperabilidade de dados de saúde, de acordo com Mukhiya e Lamo (2021), são conjuntos de representações acordadas de dados e métodos para comunicação entre Sistemas de Informação de Saúde (SIS). Nesses padrões são definidas estruturas de comunicação a serem utilizadas pelos SIS. Essas estruturas de comunicação contemplam critérios para garantir uma comunicação segura, estrutura de formato dos dados, definições de terminologias comuns, descrição e arquitetura dos protocolos de comunicação e documentação abrangente.

Um sistema que atua com dados clínicos, a exemplo de RES, agrupa diversos tipos de dados em saúde os quais possuem padrões exclusivos e que precisam de uma estrutura que permita a sua interoperabilidade viabilizando o seu acesso e uso sempre que necessário.

Nesse sentido, serão apresentados alguns padrões de dados em saúde que serviram de base para o desenvolvimento do modelo proposto nessa pesquisa. Destaca-se nesse ponto que a escolha por esses padrões em detrimento de vários outros, deu-se como base na escolha por padrões abertos, que possuam comunidades de desenvolvimento ativas, que tenham aceitação internacional e que tratem de dados de saúde que sejam utilizados em RES.

5.2.1 *Fast Healthcare Interoperability Resources* – FHIR

O HL7 FHIR é um padrão para a troca de informação em saúde, o acrônimo vem de *Fast Healthcare Interoperability Resources*.

De acordo com Mukhiya e Lamo (2021), o HL7 FHIR ou apenas FHIR, é um dos padrões populares de saúde criados com base nos padrões da Web, a exemplo do XML, JSON, HTTP e OAuth. A escolha pelos padrões Web, como elemento estruturante do FHIR, facilita o desenvolvimento de sistemas e/ou ferramentas de comunicação que adotem esse padrão como alternativa para interoperabilidade de dados em saúde.

Outro elemento relevante no que se refere à estrutura técnica do padrão FHIR, consiste no seu suporte nativo para arquiteturas RESTful e SOA, possibilitando a troca contínua de informações usando mensagens ou documentos e arquiteturas baseadas em serviços (SARIPALLE; RUNYAN; RUSSELL, 2019).

É importante destacar que questões técnicas referentes à estrutura de comunicação dos padrões aqui apresentados, mesmo sendo reconhecido por esse autor como sendo elemento de suma importância para o desenvolvimento de um modelo de interoperabilidade de dados em saúde, não serão explorados com detalhes, pois não compõem o cerne dessa pesquisa.

Destaca-se nesse ponto que o foco principal dessa pesquisa consiste no desenvolvimento de um modelo conceitual de interoperabilidade de dados em saúde sob a ótica dos princípios FAIR. Sendo assim, pretende-se explorar questões estruturais, sintáticas e semânticas dos dados que são compartilhados nos padrões de interoperabilidade escolhidos como referência para essa pesquisa.

Destacado no próprio nome, *resource*, ou recurso, é o elemento básico empregado para a interoperabilidade de sistemas de informação em saúde usando FHIR. Nesse sentido, segundo HL7 (2019), o FHIR foi projetado a partir de um conjunto de componentes modulares chamados

"Recursos". Esses recursos podem ser facilmente reunidos em sistemas de trabalho que resolvem problemas clínicos e administrativos do mundo real.

O FHIR possui recursos para conceitos administrativos como paciente, provedor, organização e dispositivo, bem como uma ampla variedade de conceitos clínicos que abrangem problemas, medicamentos, diagnósticos, planos de cuidados, preocupações financeiras entre outros.

A estrutura modular do padrão FHIR apresenta-se como uma característica muito relevante na escolha desse modelo como sendo uma das referências para o desenvolvimento do modelo conceitual de interoperabilidade proposto.

Um recurso representa algum tipo de entidade do cuidado em saúde. Por exemplo, a figura 2 ilustra partes importantes do recurso "Paciente". Neste recurso são contemplados dados demográficos ou outra informação administrativa acerca do indivíduo (paciente). Dados referentes a informações clínicas como pressão ou temperatura, por exemplo, são contemplados em outro elemento conhecido como "Observação".

Figura 2 - Recurso Paciente do padrão FHIR



Fonte: HL7 (2019)

No exemplo da figura 2, que consiste na estrutura do recurso “Paciente” do padrão FHIR, é possível identificar os elementos de forma simples devido a sua organização no arquivo HTML.

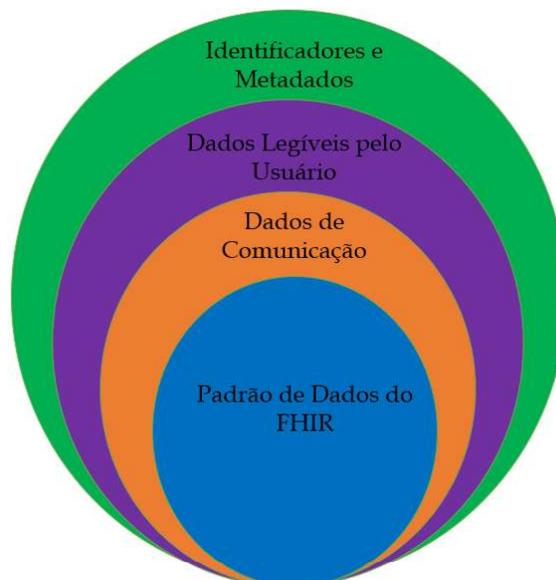
Os dados disponíveis no recurso “Paciente” estão dispostos em uma estrutura hierárquica que apresenta na parte superior os elementos de identificação da mensagem e seus metadados. Em seguida o conjunto de dados disponibilizados estão em formato legível pelos usuários, o que possibilita a interpretação desses dados até mesmo por sistemas que não tenham sido desenvolvidos para atuar diretamente com o padrão FHIR propriamente dito.

O terceiro conjunto de dados corresponde aos dados referentes à comunicação, necessários para identificação de origem/destino. Por fim, são disponibilizados os dados já no formato padrão de interoperabilidade do FHIR.

Destaca-se nesse último conjunto de dados do recurso em questão, que nele estão dispostos os dados, seguindo todas as regras definidas no padrão FHIR e que possibilitam a interoperabilidade desses dados com outros sistemas.

Com o intuito de facilitar a compreensão da estrutura hierárquica do padrão FHIR representada na figura 2 através do recurso “Paciente”, a figura 3 ilustra essa representação em forma de camadas.

Figura 3 - Estrutura de camadas do recurso paciente do padrão de interoperabilidade FHIR



Fonte: Elaborado pelo Autor (2022)

Ainda sobre o FHIR, na sua versão 4 foram contemplados elementos que possibilitem a compatibilidade com sistemas que trabalham com alguns padrões de dados em saúde amplamente disponíveis no mercado, a exemplo do SNOMED-CT, LOINC, RxNorm, família ICD 10.

Ainda segundo HL7 (2019), atualmente existem disponíveis na documentação oficial do padrão FHIR um número de 145 tipos distintos de recursos. Todos eles devidamente documentados.

5.2.1.1 Recursos do padrão FHIR

Um recurso, segundo HL7 (2019), pode ser compreendido como uma coleção de modelos de informação que definem os elementos de dados, restrições e relacionamentos para os “objetos de negócios” mais relevantes para a saúde.

A ideia de se trabalhar com uma estrutura baseada em recursos modulares estão alinhados com os princípios da reutilização de dados, escalabilidade, desempenho, usabilidade, fidelidade e implementabilidade.

Outro ponto que merece destaque no que se refere à possibilidade de trabalho com estruturas modulares pode ser identificada no fato de não existir um consenso a respeito de quais dados de saúde devem ser considerados em um processo de interoperabilidade de dados em saúde. Dessa forma, à medida que o entendimento referente ao conjunto de dados sofrer alterações, o padrão de interoperabilidade mantém, altera ou cria novos recursos, mantendo-se atualizado com as mudanças que surgirem.

No que diz respeito à estrutura de um recurso utilizado no padrão FHIR, HL7 (2019) destaca que são adotados como critérios básicos para todos os recursos os seguintes critérios:

- Ter uma identidade conhecida (uma URL) pela qual pode ser endereçada;
- Ser classificado como um dos tipos de recursos previamente existente;
- Possuir um conjunto de elementos estruturados conforme descrito na definição do tipo de recurso ao qual se classifica;
- Ser adaptável as mudanças adotadas pelo tipo de recurso que se classifica.

Além dos critérios básicos, são definidos, também, os elementos básicos contidos na estrutura dos recursos e que possuem a finalidade de definição de limites e relacionamentos. Sendo eles:

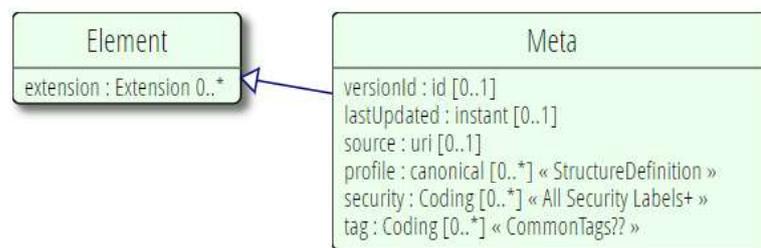
- Identidade;
- Metadados;
- Idioma Básico;
- Conjunto de Regras de Negócio.

Referente aos elementos básicos de um recurso, vê-se como relevante uma compreensão mais detalhada uma vez que essa estrutura foi escolhida como um dos elementos base para o desenvolvimento do modelo de interoperabilidade conceitual com base nos princípios FAIR que está sendo proposto.

Nesse sentido, o elemento básico “Identidade” configura-se como sendo um “id lógico” único dentro dessa estrutura, com relação a todos os recursos do mesmo tipo no mesmo servidor. Uma vez definido pelo servidor, o id lógico nunca é alterado, nem outro elemento recebe um id lógico existente.

Referente ao elemento “Metadados”, presente em cada um dos recursos existentes, identificado pela presença de elemento do tipo "Meta", que consiste em um conjunto de metadados que fornece contexto técnico e de fluxo de trabalho ao recurso. Destaca-se referente aos metadados que esses podem ser parametrizados de acordo com a necessidade do recurso. Porém, de acordo com HL7 (2019), os metadados do modelo padrão de um recurso FHIR podem ser identificados na figura 4.

Figura 4 - Diagrama UML da estrutura de Metadados de um recurso FHIR



Fonte: HL7 (2019)

Como pode ser identificado na figura 4, de forma padrão são adotados um conjunto de seis metadados para os recursos. Com o objetivo de melhor detalhar esses recursos, eles serão descritos no quadro 3.

Quadro 3 - Conjunto de metadados para elementos do padrão FHIR

Item de Metadados	Tipo	Descrição de Utilização
“<i>versionId</i>” (Versão)	Id	Sua identificação é alterada cada vez que o conteúdo do recurso seja alterado. Tem a finalidade de fornecer a informação sobre atualizações desse recurso.
“<i>lastUpdated</i>” (Última Atualização)	Instante	Indica o momento (data/hora) em que ocorreu a última atualização do recurso.
“<i>source</i>” (Fonte)	uri	Identifica a origem desse recurso. Essa origem pode ser um sistema, banco de dados, mensagem, documento etc.
“<i>profile</i>” (Perfil)	canônico	Utilizado para indicar quais versões do padrão FHIR esse recurso está em conformidade.
“<i>security</i>” (Segurança)	codificação	Utilizado através de “ <i>tags</i> ” inseridas de acordo com as regras da política de segurança definida.
“<i>tag</i>” (Etiqueta)	codificação	Utilizadas como alternativa de identificação do recurso com o fluxo do processo.

Fonte: Elaborado pelo Autor (2022)

Ainda sobre os elementos básicos presentes na estrutura de um recurso do padrão FHIR, o elemento “Idioma Básico” tem a finalidade de definir um idioma padrão para esse recurso com o objetivo de oferecer suporte à indexação e acessibilidade.

Por fim, o elemento “Conjunto de Regras de Negócio” tem o objetivo de representar uma referência de utilização. Pode ser compreendido como um acordo personalizado contendo uma descrição de utilização desse recurso. Visando a não limitação de utilização dos recursos é importante que o conhecimento referente ao conteúdo do elemento “Conjunto de Regras de Negócio” seja público, possibilitando que a utilização segura e adequada seja possível por quem tiver interesse em utilizar o referido recurso.

Não está contido no escopo dessa pesquisa um aprofundamento nos conteúdos técnicos desses padrões de interoperabilidade, sendo previsto a realização de uma discussão a nível estrutural, em especial na definição da estrutura dos dados e seus metadados.

Diante do exposto, no tópico seguinte será explorado o padrão de interoperabilidade OpenEHR.

5.2.2 Open Electronic Health Records – OpenEHR

O Registro Eletrônico de Saúde (RES) ou *Electronic Health Records* (EHR) caracteriza-se como sendo uma ferramenta para os profissionais da assistência à saúde no cuidado com qualidade e segurança com os seus pacientes.

Segundo OpenEHR Brasil (2021), o EHR trata-se de um depósito eletrônico de informações que gira em torno da saúde das pessoas, ocasionando uma análise mais completa de seus históricos clínicos sobre os atendimentos de saúde.

Ainda de acordo com Leslie (2020), apesar dos registros de EHR existirem a mais de 50 anos, continuamos na busca por alternativas de proporcionar a troca de dados entre equipamentos e sistemas de informação mais eficientes.

Nesse sentido, surge o padrão *Open Electronic Health Records* – OpenEHR que pode ser compreendido como sendo uma tecnologia para e-saúde, que consiste em especificações abertas, modelos clínicos e softwares que podem ser usados para criar padrões e construir soluções de informação e interoperabilidade para a saúde (OPENEHR 2022).

Ainda de acordo com OpenEHR (2022), o OpenEHR surgiu formalmente em 1999 como sendo uma evolução do *Good European Health Record* (GEHR), desenvolvido inicialmente por Dr. Alain Maskens e pelo Dr. Sam Heard e que tem como objetivo trabalhar na arquitetura de registros eletrônicos de saúde.

O projeto GEHR teve seu fim em 1994, sendo sucedido por outros projetos, a exemplo do projeto Synapses. Em 1998 o projeto Synapses foi concluído e com isso David Ingram propôs o estabelecimento de uma base de código aberto para levar adiante o alinhamento entre as informações do paciente e das perspectivas clínicas. Para essa iniciativa foi proposto por David Ingram o nome de OpenEHR.

Em 1999 surgiu, de acordo com o OpenEHR (2022), a *OpenEHR Foundation* que consiste em uma comunidade internacional *on-line* cujo objetivo é promover e facilitar o progresso em direção aos registros eletrônicos de saúde de alta qualidade, para atender às necessidades de pacientes e médicos em todos os lugares.

A *OpenEHR Foundation* tem como sendo um de seus pilares fundamentais o incentivo e manutenção da publicação de seus trabalhos em domínios públicos, bem como a disponibilização de seus sistemas através de uma licença *OpenSource*.

É oportuno destacar que em 2003 a *OpenEHR Foundation* foi substituída pela *OpenEHR International*, que consiste em uma organização internacional sem fins lucrativos que possui a finalidade de gerenciar os artefatos produzidos pelo projeto OpenEHR.

Ainda sobre o OpenEHR *International*, destaca-se como sendo uma de suas principais contribuições a criação de quatro programas que se concentram respectivamente em especificações, modelagem clínica, software e educação. Os três primeiros correspondem aos principais tipos de saída da comunidade OpenEHR (OPENEHR, 2022).

Seguindo na mesma de atuação do OpenEHR *Foundation*, a organização OpenEHR *International* mantém a ideia de compartilhamento de trabalhos em domínios públicos e disponibilização de sistemas com licença *OpenSource*. Todo o desenvolvimento de trabalhos e sistemas são realizados pelos membros da comunidade internacional que mantém o projeto OpenEHR e são gerenciados e organizados, de acordo com sua linha de interesse, em um dos quatro programas criados.

Dentre os programas citados, considerando os objetivos dessa pesquisa, destacam-se os programas de “Especificações” e “Modelagem Clínica”. De acordo com Tute *et al.* (2021) o padrão OpenEHR possui estrutura multinível que segmenta em estruturas específicas os dados clínicos, também conhecidos como arquétipos, dos dados de identificação e técnicos.

Como citado, mesmo entendendo que dados e informações técnicas são de extrema relevância para o desenvolvimento de modelo de interoperabilidade de dados, o cerne dessa pesquisa será a estrutura de metadados e a estrutura de dados clínicos, essa última tem uma relevância importante para viabilização da interoperabilidade a nível semântico.

A relevância atribuída aos programas de “Especificações” e “Modelagem Clínica” fundamenta-se principalmente no fato do escopo de interesse desses programas estar alinhado com o desenvolvimento de itens relacionados à estrutura de identificação das informações (metadados) e arquétipos de informações clínicas.

O programa “Especificações”, de acordo com OpenEHR (2022), define os modelos formais e plataforma técnica e inclui modelo de informação, linguagem de consulta, modelos de conteúdo OpenEHR (arquétipos e templates).

O referido autor também destaca que o programa de “Modelagem Clínica” consiste na construção de arquétipos que servem como padrões internacionais de conteúdo clínico reutilizável, dentro da estrutura OpenEHR. Esses arquétipos são construídos por profissionais clínicos e especialistas em informática em saúde que trabalham no ambiente de *Clinical Knowledge Manager* (CKM).

É importante citar que o CKM consiste em um ambiente de desenvolvimento colaborativo, gerenciamento e publicação de uma ampla gama de recursos de conhecimento clínico, incluindo arquétipos, modelos e subconjuntos de terminologia.

Nesse contexto, Leslie (2007), diferencia conceitualmente um arquétipo de um modelo quando considera que o arquétipo consiste na especificação de dados para um único conceito clínico. Referente ao modelo, o mesmo pode ser compreendido como sendo um conjunto de arquétipos que são agrupados para uma finalidade clínica.

Referente aos programas “Software” e “Educação”, eles possuem, respectivamente, as finalidades de fomentar o desenvolvimento de ferramentas de software de código aberto disponíveis para toda a comunidade OpenEHR, e compartilhar os conhecimentos e resultados obtidos com as experiências com o uso do OpenEHR em diversos locais com todos os membros da comunidade.

Fica evidente a importância desses últimos programas no que se refere à manutenção da comunidade OpenEHR, através do compartilhamento das informações e experiências, possibilitando a padronização das ações e reutilização de artefatos já desenvolvidos e testados.

A construção dos arquétipos e modelos de dados clínicos se configuram como sendo uma grande contribuição do padrão OpenEHR para o desenvolvimento dessa pesquisa. Diante do exposto, esse tópico será melhor detalhado a seguir.

5.2.2.1 Arquétipos e modelos do padrão OpenEHR

Para uma melhor compreensão dos conceitos sobre arquétipos e modelos em uma estrutura OpenEHR, vê-se como relevante enfatizar que o padrão OpenEHR foi concebido através da ideia de ser multinível, ou seja, separação completa de modelos de informação (como modelos de objeto de software, modelos de esquemas de banco de dados) de modelos de domínio.

Nessa visão de arquitetura multinível um arquétipo que, de acordo com Bale (2015), consiste em uma definição formal e reutilizável de informações de nível de domínio, definidas em termos de restrições em um modelo de informações, está completamente separada das estruturas de software e banco de dados em um contexto OpenEHR.

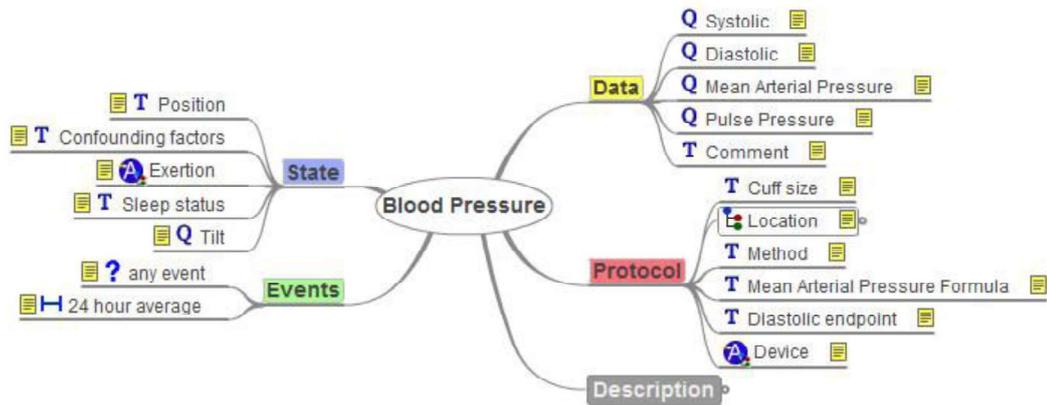
Essa segregação possibilita o desenvolvimento independente de estruturas de arquétipos e estruturas de tecnologias (softwares e banco de dados), o que é visto por esse autor como sendo uma grande contribuição de padrão OpenEHR para o desenvolvimento da interoperabilidade de dados de saúde.

Ainda de acordo com o Bale (2015), a utilização de arquétipos permite que especialistas de domínio, como médicos, criem uma biblioteca de elementos de dados e definições de grupo para os dados em seus sistemas de informação, fornecem validação em tempo de execução da

entrada de dados via interface ou qualquer processo em lote e fornece uma base para consulta semântica de dados.

A estrutura de um arquétipo, conforme ilustrado na figura 5, não define um conjunto de dados propriamente dito. Isso ocorre por esse elemento ser muito específico.

Figura 5 - Exemplo do arquétipo *Blood Pressure* (Pressão Arterial)



Fonte: Bale (2015)

Como pode ser identificado na figura 5, o arquétipo “Pressão Arterial” contém diversos elementos de dados específicos para o contexto da pressão arterial, não contemplando nesse exemplo de arquétipo demais informações clínicas do paciente necessárias para subsidiar uma tomada de decisão clínica por parte da equipe assistencial.

Para essa tomada de decisão clínica, vê-se como necessário a utilização de um conjunto de dados clínicos, que na estrutura do OpenEHR é conhecido como modelo.

De acordo com Bale (2015), um modelo consiste em um conjunto de dados que agrega pedaços de outros arquétipos para um uso específico, por exemplo, um formulário de tela ou definição de mensagem. Os conjuntos de dados quase sempre contêm pontos e grupos de dados de vários arquétipos, conforme ilustrado na figura 6.

Figura 6 - Exemplo do modelo OpenEHR

The image shows a screenshot of an OpenEHR form for a fetal examination. The form is divided into several sections, each highlighted with a red border:

- Observations: History:** Includes fields for Symptom and Clinical description.
- BP (Blood Pressure):** Fields for systolic and diastolic pressure in mm[Hg].
- Weight:** Field for weight in kg.
- Fetal movements:** Field for Presence.
- FHR (Fetal Heart Rate):** Fields for Rate in /min and a checkbox for Present.
- Examination of the uterus:** Includes Normal statements, Clinical description, Size (Fundal height in cm and Weeks), Assessment of liquor volume, and Number of fetuses.
- Examination of the fetus:** Includes Identifier, Normal statements, Clinical description, Lie of the fetus, Presentation, Position, Engagement, and Size relative to gestation.
- Assessment:** Fields for Rationale and Assessment.
- Urinalysis:** Fields for Glucose, Bilirubin, Ketones, and Specific gravity.
- Follow up:** Fields for Service, Details, and Appointment date and time (Monday, 19 March 2007).

Fonte: Bale (2015)

Como pode ser percebido na figura 6, os modelos são um meio de construção de conjuntos de dados clínicos para um caso de uso específico. Estes conjuntos de dados são elementos de um ou mais arquétipos, representados na figura pelas áreas circuladas em vermelho, e podem adicionar outras características para o uso desses arquétipos em cenários específicos.

O grande benefício da utilização de modelos consiste na possibilidade de reutilização de arquétipos existentes para diversos cenários. Como exemplo, os grupos de dados de arquétipos para pressão arterial, peso e glicemia podem ser usados no registro de uma consulta de um paciente diabético ou em uma consulta pré-natal. Dessa forma, serão criados modelos específicos para 'avaliação diabética' e 'consulta pré-natal' apenas reutilizando os arquétipos existentes e sem a necessidade de preocupação com questões de tecnologias.

5.3 ADAPTABILIDADE DOS PADRÕES FHIR E OPENEHR COM OS PRINCÍPIOS FAIR

Em consonância aos objetivos dessa pesquisa, que possui em seu cerne o desenvolvimento de uma análise da aplicação dos princípios FAIR no processo de interoperabilidade de dados em saúde, resultando no desenvolvimento de um modelo conceitual de interoperabilidade de dados em saúde, investigar o nível de alinhamento existente entre os padrões de interoperabilidade de dados em saúde escolhidos como referência para o desenvolvimento dessa pesquisa e os princípios FAIR, torna-se uma atividade essencial.

Para o desenvolvimento dessa investigação, faz-se necessário a utilização de uma ferramenta que possibilite mensurar esse alinhamento afim de indicar pontos de melhorias que serão explorados no modelo proposto por essa pesquisa.

Para essa finalidade foi escolhida a aplicação FairDataBR que consiste, de acordo com FairDataBR (2021), em uma ferramenta para a avaliação de conjuntos de dados contribuindo com a automação do processo de verificação da aderência de conjuntos de dados aos princípios FAIR.

O FairDataBR caracteriza-se por ser de uso simples e intuitivo, foi construído por pesquisadores vinculados ao Programa de Pós-graduação em Ciência da Informação e o Programa de Pós-graduação em Gestão nas Organizações Aprendentes da Universidade Federal da Paraíba (PPGCI/MPGOA - UFPB).

O processo de avaliação da ferramenta FairDataBR consiste em checagem da aderência do conjunto de dados em questão, com cada um dos princípios definidos nos princípios FAIR, conforme ilustrado da figura 7.

Figura 7 - Itens analisados pela ferramenta FairDataBR



A ferramenta para a avaliação de conjuntos de dados FairDataBR, surgiu a partir da constatação da necessidade do desenvolvimento de uma aplicação que contribuísse para a automação do processo de verificação da aderência de conjuntos de dados aos Princípios FAIR. Esta ferramenta foi desenvolvida por pesquisadores vinculados à Universidade Federal da Paraíba (PPGCI/MPGOA - UFPB). Caracteriza-se por ser de uso simples e intuitivo.

A ideia dos Princípios FAIR surgiu a partir do trabalho de um grupo variado de pessoas que constatou a necessidade premente de se melhorar a infraestrutura de apoio a prática do compartilhamento dos dados de pesquisa e de seu posterior reuso. Assim, a partir destas premissas, um conjunto de princípios foram desenvolvidos com o objetivo de servirem como diretrizes para todos que desejam compartilhar e ampliar o reuso dos seus dados de pesquisa (WILKINSON et al., 2016).

PRINCÍPIOS FAIR

Os Princípios FAIR são um acrônimo para *Findable* (localizável), *Accessible* (acessível), *Interoperable* (interoperável) e *Reusable* (reutilizável). Na publicação pioneira sobre esses princípios, Wilkinson *et al.* (2016) afirmam que, para que os dados de pesquisa estejam em conformidade com os princípios, devem seguir as respectivas orientações:

Para ser LOCALIZÁVEL (FINDABLE)	Para ser ACESSÍVEL (ACCESSIBLE)	Para ser INTEROPERÁVEL (INTEROPERABLE)	Para ser REUTILIZÁVEL (REUSABLE)
F1. Os (meta)dados são atribuídos a um identificador persistente.	A1. Os (meta)dados são recuperáveis por seu identificador, usando-se um	I1. Os (meta)dados usam uma linguagem formal, acessível, compartilhada e	R1. Os (meta)dados são ricamente descritos com uma pluralidade de

Fonte: FairDataBR (2021)

A checagem em questão ocorre de forma individualizada, considerando cada um dos princípios de cada um dos grupos definidos no FAIR. Para viabilização dessa avaliação faz-se necessário o preenchimento de questões que ao final serão pontuadas apresentando um nível de aderência do conjunto de dados analisados com os princípios FAIR.

É importante destacar que a ferramenta FairDataBR possui em seu escopo questionamentos relevantes para a compreensão do nível de adaptabilidade de uma estrutura de dados aos princípios FAIR. Nessa estrutura existem questionamentos diretamente relacionados aos princípios FAIR, como também existem outras questões que possuem relevância significativa para a adaptabilidade da base de dados analisada aos princípios FAIR.

Ainda sobre a estrutura da ferramenta FairDataBR, é importante ressaltar que essa possui uma versão padrão e uma versão estendida, disponível através de um link de acesso “FairDataBR+” na tela de acesso inicial.

A versão estendida possui um volume maior de questionamentos, porém, analisando o conteúdo disponível em ambas as versões, constatou-se que as questões presentes na versão padrão estão mais bem alinhadas com os objetivos dessa pesquisa.

Uma vez que a escolha da versão da ferramenta FairDataBR, foi a padrão é oportuno mencionar que a contagem das questões disponíveis na ferramenta pode não seguir uma sequência lógica, isso se deve ao fato de que algumas questões são exibidas apenas na versão estendida da ferramenta.

De acordo com a figura 8, é possível identificar os itens analisados no princípio *Findable* (Localizável).

Figura 8 - Ferramenta FairDataBR: Princípio Findable (Localizável)

Avaliação Fair

PRINCÍPIO FINDABLE

F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente? [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) *

Identificador Persistente

Endereço Web

Identificador Local

Sem identificador

F2. Os conjuntos de dados são descritos com metadados? [1](#) [2](#) [3](#) [4](#) [5](#) *

Metadados Ricos

Metadados Estruturados

Metadados Simples

Sem metadados

F3. O identificador está incluído em todos os registros / arquivos de metadados que descrevem os dados? [1](#) [4](#) [6](#)

Sim

Não

F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web? [1](#) [5](#)

Sim

Não

F5. O conjunto de (meta)dados está publicado em um repositório? [1](#) [2](#) *

Público Geral

Específico de [Domínio](#)

Institucional de Local

Em Nenhum Repositório

Fonte: FairDataBR (2021)

Analisando a figura 8, é possível identificar o item “F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente?”. Esse item possui uma relação direta com o princípio F1 do FAIR que, segundo Go Fair (2021), possui a seguinte descrição: “F1: (Meta) dados são atribuídos a identificadores globalmente únicos e persistentes” (tradução nossa).

Para o item F1 da ferramenta são atribuídas como opções de respostas a possibilidade de a estrutura de dados possuir “Identificador Persistente”, “Endereço Web”, “Identificador Local” e “Sem Identificador”.

É oportuno citar a existência de uma relação de dependência entre os princípios FAIR. Nesse sentido, de acordo com Go Fair (2021), o princípio F1 possui um destaque devido a constatação de que será difícil alcançar outros aspectos do FAIR sem identificadores globalmente exclusivos e persistentes.

Essa relação de dependência também foi incorporada como sendo uma das funcionalidades da ferramenta FairDataBR, pois a habilitação de opção de resposta de alguns itens do questionário, depende exclusivamente de respostas de itens anteriores.

Para o item “F2. Os conjuntos de dados são descritos com metadados?”, as possibilidades de resposta são: “Metadados Ricos”, “Metadados Estruturados”, “Metadados Simples” e “Sem Metadados”.

O princípio F2 do FAIR, requisitado na ferramenta através do item F2, evidencia a importância dos metadados para FAIR. Em um cenário ideal, espera-se encontrar uma estrutura de metadados rica o suficiente que possibilite a localização dos dados através da análise de seus metadados, e não apenas pelo seu identificador.

Para auxiliar no desenvolvimento de uma estrutura de metadados rica, deve ser considerado que essa estrutura não deve presumir que sabe quem vai fazer uso desse dado ou para qual finalidade esse dado será utilizado. Diante disso, vê-se como uma boa prática, a inclusão generosa de metadados, possibilitando a sua utilização por uma parcela maior de usuários e/ou sistemas.

Os princípios F3 e F4 do FAIR foram contemplados na ferramenta através dos itens “F3. O identificador está incluso em todos os registros / arquivos de metadados que descrevem os dados?” e “F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na *Web*?”, disponibilizando as seguintes opções de resposta: “Sim” ou “Não”.

O princípio F3 está diretamente relacionado à conexão entre os metadados e os dados propriamente dito, uma vez que na sua grande maioria os dados e os seus metadados são agrupados em arquivos distintos. Referente ao princípio F4, esse está relacionado à possibilidade de localização desses dados em uma estrutura Web através de recursos pesquisáveis.

Ainda referente ao conjunto de princípios *Findable* (Localizável), o item “F5. O conjunto de (meta)dados está publicado em um repositório?” possui como alternativas as seguintes opções: “Público Geral”, “Específico de Domínio”, “Institucional de Local”, “Em Nenhum Repositório”.

O item F5 disponível na ferramenta FairDataBr não pode ser entendido como sendo um dos princípios FAIR. Porém, questionar o tipo de repositório em que se encontram os metadados é sim de grande relevância em um processo de uso e reuso de dados. Entende-se nesse ponto que quanto maior for a disponibilidade de acesso a esse repositório, igualmente maior será a possibilidade de recuperação, acesso e uso desse dado.

Destaca-se que cada uma das respostas possui um nível de aderência da base de dados com os princípios FAIR, e que o conjunto de todas as respostas possibilitará a identificação de um nível final de aderência do conjunto de dados ao princípio específico do FAIR.

Referente ao princípio *Accessible* (Acessível), a ferramenta FairDataBR realiza a checagem de adaptabilidade através dos questionamentos ilustrados na figura 9.

Figura 9 - Ferramenta FairDataBR: Princípio Accessible (Acessível)

PRINCIPIO ACCESSIBLE

A1. Quão acessíveis são os (meta)dados?²

Acessível ao Público

Acesso com [Embargo](#)

Acesso somente aos dados ou somente aos metadados

Sem acesso a dados e metadados

A2. Os (meta)dados estão acessíveis *online* sem a necessidade de intermediação de [protocolos](#) ou ferramentas especializadas a partir do momento em que o acesso é permitido?² *

API de [Web Service](#) padrão

Web Service não padrão

Download de arquivo *online*

Através de acordo individual

Sem acesso aos (meta)dados

A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?¹ [6](#)

Sim

Não

A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?¹ [3](#)

Sim

Não

Não se aplica

A5. O [protocolo](#) (ex: [HTTP](#), [SAML](#), [OAI-PMH](#)) é aberto, gratuito e universalmente implementável?¹ [6](#) *

Sim

Não

A7. É possível realizar o *download* dos conjuntos de (meta)dados?²

Sim

Não

Fonte: FairDataBR (2021)

Seguindo a mesma lógica de análise aplicada ao conjunto de itens relacionados aos princípios *Findable*, a figura 9 ilustra os itens de checagem de adaptabilidade para conjunto de princípios *Accessible* (Acessibilidade) do FAIR.

Segundo Fiocruz (2022, *online*), o conjunto de princípios da acessibilidade prevê:

Que seus dados podem ser acessados por outros. Significa o uso de protocolos de comunicação padronizados, abertos e gratuitos, que ofereçam autenticação e acesso aos metadados mesmo quando o dado não estiver mais disponível.

Nesse sentido a ferramenta FairDataBR inicia a checagem com o item “A1. Quão acessíveis são os (meta)dados?”. Como alternativas de resposta estão disponíveis as opções: “Acessível ao Público”, “Acesso com Embargo”, “Acesso somente aos dados ou somente aos metadados” e “Sem acesso a dados e metadados”.

A checagem seguinte, disponível no item “A2. Os (meta)dados estão acessíveis online sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?”, possui relação direta com os princípios A1 “(Meta)dados são recuperáveis por seu identificador usando um protocolo de comunicação padronizado” e A1.1 “O protocolo é aberto, gratuito e universalmente implementável” do FAIR.

Ainda sobre item A2 da ferramenta FairDataBR, é oportuno destacar uma parte da sua descrição em que se questiona sobre a possibilidade de acesso aos metadados a partir do momento em que o acesso é permitido. Referente a “permissão” de acesso aos dados, remete a questão de segurança de acesso. Nesse quesito, o item A2 também está alinhado ao princípio A1.2 “O protocolo permite autenticação e autorização quando necessário” do FAIR.

Destaca-se nesse quesito de segurança que a acessibilidade do FAIR não está, necessariamente, relacionada ao acesso aberto ou livre. O princípio A1.2 busca garantir condições ideais para o acesso aos dados, inclusive através da utilização de recursos de autenticação e autorização.

Como alternativas de respostas, o item A2 da ferramenta apresenta as seguintes opções: “API de Web Service padrão”, “Web Service não padrão”, “Download de arquivo online”, “Através de acordo individual” e “Sem acesso aos (meta)dados”.

O item “A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?” que apresenta como alternativas de resposta as opções “Sim” e “Não”, busca identificar uma forma de acesso aos metadados mais adequadas a sistemas ou ferramentas de Tecnologia.

O princípio A2 do FAIR “Os metadados devem ser acessíveis mesmo quando os dados não estiverem mais disponíveis” está diretamente relacionado com o item “A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?” da ferramenta FairDataBR. Como alternativa de resposta para o item A4, são disponibilizadas as opções “Sim”, “Não” e “Não se aplica”.

Ainda referente ao princípio A2 do FAIR, esse possui uma conexão direta com o princípio F4, que trata de questões relacionadas ao registro e indexação dos metadados.

Os itens A5 “O protocolo (ex.: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?” diretamente relacionado ao princípio A1.1 do FAIR e o A7 “É possível realizar o download dos conjuntos de (meta)dados?”, que possuem como alternativas as opções “Sim” e “Não” destacam outros elementos relevantes a questão da acessibilidade.

O próximo princípio que compõe o acrônimo FAIR, é o *Interoperable* (Interoperável). É relevante destacar que o princípio da interoperabilidade é sim o princípio de maior interesse dessa pesquisa, que tem como seu objetivo o desenvolvimento de um modelo de interoperabilidade de dados em saúde com base nos princípios FAIR.

A figura 10 ilustra os itens disponíveis na ferramenta FairDataBR para a realização da checagem de adaptabilidade das bases de dados com os princípios de interoperabilidade do FAIR.

Figura 10 - Ferramenta FairDataBR: Princípio Interoperable (Interoperável)

PRINCÍPIO INTEROPERABLE

11. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?^{1 2 4 5 *}

Sim

Não

12. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?^{1 5 *}

Esquema de Comunidade Padrão

Esquema Explícito

Esquema não formalizado

13. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores? ^{1 2 3 4 5}

Links Internos/Externos

Sem Links

14. Para representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico?^{1 2 5}

Uso de instrumentos de controle terminológico COM identificador global

Uso de instrumentos de controle terminológico SEM identificador global

Nenhum recurso/instrumento

Elementos de dados não descritos

Fonte: FairDataBR (2021)

A análise dos itens de interoperabilidade da ferramenta FairDataBR inicia-se com o item “I1. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?”. Como alternativas de resposta, estão disponíveis as opções “Sim” e “Não”.

O questionamento sobre os formatos dos conjuntos de metadados, mais especificamente se esses estão ou não disponíveis em formatos preferenciais, busca analisar se existe uma padronização na escolha do formato dos conjuntos de metadados.

Segundo Fankhauser *et al* (2019, p 62), o termo “formatos preferenciais”:

não apenas proporcionam uma maior certeza de que seus dados podem ser lidos no futuro, mas também ajudam a aumentar a reutilização e a interoperabilidade.

Ainda, o referido autor destaca que a escolha por formatos preferenciais favorece a leitura desses metadados por softwares, possibilitando inclusive uma maior integração com outros softwares que também façam uso dessa mesma estrutura.

Percebe-se nesse ponto que elementos do princípio FAIR I1 “(Meta)dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para representação de conhecimento” estão presentes no item I1 da ferramenta FairDataBR. Esse alinhamento fica mais evidente quando se percebe que, tanto no item I1 da ferramenta quanto no item I1 do FAIR, existe um interesse na busca pela padronização dos conjuntos de metadados com base em estruturas conhecidas e amplamente divulgadas.

O item seguinte da ferramenta “I2. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?” corrobora com os objetivos presentes no item I1.

Para o item I2, estão disponíveis as seguintes opções de respostas: “Esquema de Comunidade Padrão”, “Esquema Explícito” e “Esquema não formalizado”.

O terceiro item da ferramenta, com a descrição “I3. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores?”. Para esse item, que está em conformidade com o princípio I3 “(Meta)dados incluem referências qualificadas a outros (meta)dados” do FAIR, espera-se analisar o nível de associação entre os conjuntos de metadados da base de dados analisada.

Nesse quesito, espera-se um alto índice de vínculo entre os conjuntos de metadados através de seus indicadores globalmente únicos e persistentes, favorecendo ao enriquecimento do modelo de dados.

Para o item I3 da ferramenta FairDataBR, as opções de respostas disponíveis são: “Links Internos/Externos” e “Sem Links”.

Ainda na aba de interoperabilidade da ferramenta FairDataBR, o último item disponível é o “I4. Para representação dos dados ou conjuntos de dados, utilizaram-se recursos/instrumentos de controle terminológico?”.

As opções de resposta disponíveis são: “Uso de instrumentos de controle terminológico COM identificador global”, “Uso de instrumentos de controle terminológico SEM identificador global”, “Nenhum recurso/instrumento” e “Elementos de dados não descritos”.

Por fim, o FairDataBR segue com sua coleta de dados para os princípios da *Reusable* (Reutilizável).

A figura 11 ilustra os itens disponíveis na ferramenta FairDataBR referente ao princípio da reusabilidade.

Figura 11 - Ferramenta FairDataBR: Princípio Reusable (Reutilizável)

PRINCÍPIO REUSABLE

R1. Os conjuntos de (meta)dados estão licenciados? [1](#) [2](#) [4](#) [5](#) [6](#)*

Sim

Não

R2. Qual a licença de uso dos conjuntos de (meta)dados? [1](#) [2](#) [4](#) *

Aberta

Restrita

R4. Os conjuntos de (meta)dados possuem proveniência detalhada? [1](#) [5](#)

Sim

Não

R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados? [1](#) [5](#) [6](#) *

Rastreamento de Proveniência

Declaração de Proveniência

Sem Informações de Proveniência

R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio? [1](#) [4](#) [6](#)

Padrões de Domínio

Padrões Mínimos de Domínio

Padrões Genéricos

Nenhum Padrão Utilizado

Fonte: FairDataBR (2021)

No tópico que trata sobre o princípio da reusabilidade dos dados na ferramenta FairDataBR, a análise tem início com “R1. Os conjuntos de (meta)dados estão licenciados?” disponibilizando como alternativas de respostas as opções “Sim” e “Não”.

Nessa mesma linha que busca coletar informações referente à utilização de licenças de uso para o conjunto de metadados, destaca-se também a questão “R2. Qual a licença de uso dos conjuntos de (meta)dados?” que possuem como alternativas de respostas as opções “Aberta” e “Restrita”.

A utilização de licenças de uso para o conjunto de metadados está diretamente relacionado a critérios de autorização de acesso, inclusive já tendo sido explorado pela ferramenta FairDataBR no tópico que aborda questões sobre acessibilidade.

Revisitar esse assunto, agora na parte de reusabilidade de dados, evidencia a importância que esse tema possui para a recuperação desses dados e sua utilização em projetos futuros.

Corroborando com a importância da definição e exposição do tipo de licença definida para o conjunto de metadados, Wilkinson *et al.* (2018) destaca que as condições sob as quais é possível a reutilização, tanto dos metadados, quanto dos dados se torna aspecto central para viabilização dessa reutilização. O referido autor ainda destaca que essa informação deve ser disponibilizada de forma aberta, relativamente fácil de ser localizada e sem possibilidade de ambiguidade.

O item seguinte da ferramenta FairDataBR, “R4. Os conjuntos de (meta)dados possuem proveniência detalhada?”, cuja opções de resposta são: “Sim” e “Não”, busca analisar informações referentes às atividades, entidades e pessoas envolvidas na elaboração de um conjunto de dados. Tais informações são relevantes para aumentar a qualidade, segurança e confiabilidade dos dados.

Como um complemento ao item R4 da ferramenta FairDataBR, o item “R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados?”, com as opções de respostas: “Rastreamento de Proveniência”, “Declaração de Proveniência” e “Sem Informações de Proveniência”.

Finalizando a análise do princípio de reusabilidade, bem como de todo questionário da ferramenta FairDataBR, o item “R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio?”, com a seguintes opções de respostas: “Padrões de Domínio”, “Padrões Mínimos de Domínio”, “Padrões Genéricos” e “Nenhum Padrão Utilizado”.

O item R6 está alinhado com o princípio FAIR R1.3 “(Meta)dados atendem aos padrões da comunidade relevantes para o domínio”. Nesse princípio é recomendado a utilização de

modelos e vocabulários comuns, como alternativa para facilitar a reutilização de dados por uma parcela maior de usuários que deles necessitem.

É indicado também que a impossibilidade de adoção de modelos comuns de dados seja indicada no conjunto de metadados facilitando a compreensão de quem for reutilizar esses dados.

Após o preenchimento de todos os itens a ferramenta FairDataBR, contabiliza o nível de adesão da base de dados analisada com os princípios FAIR, através de painel segmentado por conjunto de princípios (*Findable, Accessible, Interoperable, Reusable*), além de um indicador que apresenta a média geral, conforme ilustra a figura 12.

Figura 12 - Ferramenta FairDataBR: Tela exemplo de nível de adaptabilidade com os princípios FAIR



Fonte: FairDataBR (2021)

A utilização da ferramenta FairDataBR nessa pesquisa será adaptada ao contexto que se pretende analisar. Como citado, a FairDataBR foi desenvolvida para analisar critérios de adequação dos princípios FAIR em bases de dados existentes. Uma vez que a proposta aqui consiste em utilizar a metodologia disponível no FairDataBR para analisar padrões e modelos de dados em saúde, que em nosso cenário nos referimos aos padrões FHIR e OpenEHR, bem como o modelo conceitual de interoperabilidade desenvolvido, essa aplicação pode ser entendida como sendo mais uma inovação que essa pesquisa contribui com a academia.

Nesse sentido, nas seções 5.3.1 e 5.3.2 serão exploradas questões referentes à utilização do FairDataBR para os padrões FHIR e OpenEHR.

5.3.1 Alinhamento do FHIR com os princípios FAIR

A escolha do padrão FHIR como sendo uma das bases conceituais para o desenvolvimento do modelo de interoperabilidade de dados em saúde com base nos princípios FAIR, proposto nessa pesquisa, justificou-se pelo fato de ser um modelo de livre acesso, com aceitação internacional e com uma comunidade de desenvolvimento ativa.

A estrutura modular do padrão FHIR, baseada no desenvolvimento de recursos que se adequam às necessidades dos elementos de dados em saúde também foi um fator positivo para a escolha desse padrão como sendo uma referência base para o desenvolvimento dessa pesquisa.

Nesse sentido, vê-se como relevante a realização de uma análise sobre o nível de adaptabilidade do padrão FHIR com os princípios FAIR. Essa investigação tem como objetivo a busca por elementos positivos, bem como oportunidades de melhorias entre o FHIR e FAIR.

Para o desenvolvimento dessa análise, será utilizada a ferramenta FairDataBR, que foi projetada para mensurar o nível de adaptabilidade de bases de dados com os princípios FAIR através da coleta e análise de informações.

Destaca-se nesse ponto que para a realização dessa análise não utilizaremos uma base de dados, e sim os conceitos do FHIR considerado nativo, disponíveis na página oficial do FHIR.

Iniciando a análise de adaptabilidade do FHIR aos princípios FAIR, temos o seguinte item: “F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente?”. De acordo com Atlassian (2021), o FHIR oferece nativamente duas maneiras de identificar um recurso, sendo elas:

1 – Por uma URL de “localização” que identifica onde pode ser acessado (com base no “ID Lógico”);

2 – Por algum identificador inerente (“Identificador comercial” ou “URL canônico”).

A identificação prevista no modelo nativo do FHIR define o identificador global e único, porém, não garante que ele seja persistente.

Diante do exposto a alternativa “Endereço Web” vê-se como a resposta mais adequada.

O próximo item a ser analisado “F2. Os conjuntos de dados são descritos com metadados?” busca informações sobre a possibilidade de recuperar os dados com base nas informações de seus metadados, mesmo que o dado em questão não possua um identificador definido.

De acordo com HL7 (2021), todo o recurso do FHIR possui um elemento meta, porém, em sua estrutura nativa a estrutura desse elemento não é suficiente para garantir essa

recuperação. Outro ponto que corrobora com essa afirmação, e que na estrutura nativa do FHIR o elemento meta é optativo, ou seja, seu preenchimento não é obrigatório.

Considerando as alternativas de resposta disponíveis nesse item, entende-se que a opção “Metadados Simples” se enquadra como a mais adequada.

O item seguinte “F3. O identificador está incluso em todos os registros / arquivos de metadados que descrevem os dados?” está contemplado no padrão FHIR. No Padrão FHIR pode ser encontrado metadados em estrutura distinta da estrutura dos dados, como também metadados na mesma estrutura dos dados. Para ambos os cenários existem estruturas nativas que garantem o link desse identificador com sua estrutura de dados. Nesse caso a resposta mais adequada para esse item será a opção “Sim”.

No item “F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web?” destaca-se nesse quesito que, de acordo com HL7 (2021), no padrão nativo FHIR identificadores e descrições de metadados por si só não garantirão a localização na internet. Sendo necessário, para alcançar esse objetivo, a intervenção da comunidade de desenvolvimento na definição e inclusão de metadados e dados relevantes que devem ser registrados e/ou indexados por um servidor HL7 FHIR. Como alternativa de resposta, a melhor opção escolhida por esse autor é “Não”.

Para o item “F5. O conjunto de (meta)dados está publicado em um repositório?”, considerando que a análise em questão está sendo aplicada em conceitos de um padrão de interoperabilidade de dados em saúde, e não em uma base de dados propriamente dita, percebe-se que essa questão não se aplica ao contexto dessa pesquisa. Diante do exposto, considerando que o melhor momento para avaliar essa questão não será na etapa de construção conceitual do OpenFairEHR, a opção de resposta escolhida para esse item foi “Em Nenhum Repositório”.

Finalizamos com isso a análise de adaptabilidade do padrão FHIR ao conjunto de princípios *Findable* (Localizável) do FAIR. Obtendo como resultado uma nota de **4,6** (em uma escala que vai de 0 a 10) de adaptabilidade. É oportuno citar que o resultado dessa avaliação segue no anexo A dessa pesquisa.

Seguindo com a análise em questão, e dando início aos itens referente ao princípio *Accessible* (Acessibilidade), para o item “A1. Quão acessíveis são os (meta)dados?” entende-se que na estrutura do FHIR, é possível liberar o acesso ao público. Sendo assim, diante das opções de respostas disponíveis, entende-se por ser mais adequado a adoção da alternativa “Acessível ao Público” como resposta desse item.

Para o item “A2. Os (meta)dados estão acessíveis online sem a necessidade de intermediação de protocolos ou ferramentas especializadas, a partir do momento em que o

acesso é permitido?”, considerando que FHIR prevê em sua estrutura padrão o uso da API RESTful HL7 FHIR, entende-se que esse princípio está contemplado. Diante do exposto, como opção de resposta foi escolhida a opção “API de Web Service padrão”.

Seguindo com a análise do item “A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?”, entende-se que o objetivo desse item está relacionado com o item F3 do FairDataBR. Nesse caso, como já foi mencionado no item F3, é possível sim acessar o conjunto de metadados através do identificador fornecido. Sendo assim, a opção de resposta para esse será a alternativa “Sim”.

O item “A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?”, de acordo com HL7 (2021), como requisito para contemplar esse item em sua plenitude, faz-se necessário a existência de identificadores globais, únicos e persistentes. Como visto na análise do item F1, os identificadores do modo nativo do padrão FHIR podem ser compreendidos como globais e únicos, porém, não são persistentes. Sendo assim a alternativa mais adequada para responder esse item será a alternativa “Não”.

Para o item “A5. O protocolo (ex.: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?”, considerando que o padrão FHIR foi desenvolvido com base nos padrões da Web, a exemplo do XML, JSON, HTTP e OAuth, podemos concluir que a resposta para esse item será a opção “Sim”.

A análise do item “A7. É possível realizar o download dos conjuntos de (meta)dados?” seguirá a mesma linha de raciocínio do item A1, e tendo como alternativa de resposta a opção “Sim”.

Após a conclusão dos itens referente à análise de adaptabilidade do FHIR com os princípios da seção *Accessible* (Acessibilidade), foi identificada a obtenção da pontuação **8,50**. Essa pontuação mostra que esse princípio está bem inserido no core do padrão FHIR.

Dando continuidade, iniciaremos a análise dos itens referente ao princípio da *Interoperable* (Interoperabilidade). O item “I1. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?”. A definição de formatos preferenciais consiste na escolha de padrões reconhecidamente acessíveis e que estejam em pleno uso por uma parcela significativa dos interessados.

Nesse sentido, de acordo com HL7 (2021), em sua estrutura nativa o FHIR suporta padrões legíveis por máquina, a exemplo dos padrões XML, JSON e RDF, sendo assim o FHIR torna-se legível para o questionamento presente no item I1 do FairDataBR tendo como melhor resposta a opção “Sim”.

Dando seguimento à análise dos itens referente ao princípio da interoperabilidade, o item “I2. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?” busca analisar mais a fundo a estrutura utilizada para os metadados, dado a importância dessa estrutura para a busca da interoperabilidade a nível semântico.

Mesmo considerando a utilização de padrões de comunicação amplamente conhecidos e utilizados, o padrão nativo do FHIR não define nenhuma estrutura de metadados a ser utilizada como padrão. Diante dessa constatação, entende-se que a melhor alternativa para responder esse item é “Esquema não formalizado”.

No item “I3. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores?”, é possível perceber um alinhamento bem próximo dos objetivos do item I3 com o item F3 também da ferramenta FairDataBR. Porém, esses itens se diferem no momento no qual percebemos que o item I3 busca saber se o FHIR possui em sua estrutura nativa a utilização do conceito de referência qualificada ou referência cruzada.

Destaca-se, nesse ponto, o citado por HL7 (2021), em que afirma que o padrão FHIR suporta tecnicamente diferentes tipos de referências entre recursos HL7 FHIR e para objetos não FHIR. Sendo que a utilização de referências cruzadas para metadados não é cobrada como elemento obrigatório do padrão, dessa forma não é possível mensurar, nem afirmar, que se uma referência alcança todos os metadados disponíveis.

Diante do exposto, entende-se como melhor opção de resposta a alternativa “Sem links”.

Para finalizar a análise de adaptabilidade do FHIR ao princípio da interoperabilidade do FAIR, o item da ferramenta FairDataBR “I4. Para representação dos dados ou conjuntos de dados, utilizaram-se recursos/instrumentos de controle terminológico?” não é suportado pela estrutura nativa do padrão FHIR.

Essa afirmação fundamenta-se no fato do padrão FHIR permitir que sejam vinculados elementos codificáveis a vocabulários (destinados a terminologias), porém, com poucas exceções, e seu uso não é obrigatório.

Diante do exposto, fica evidenciado que toda e qualquer ação objetivando uma padronização terminológica deve ser fruto de desenvolvimento posterior o que possibilita a incompatibilidade desses termos quando se pensa em interoperabilidade de bases de dados distintas.

Como alternativa de resposta para o item I4, a opção indicada foi “Nenhum recurso/instrumento”.

Com isso finaliza-se a análise da adaptabilidade do padrão FHIR ao princípio de interoperabilidade do FAIR, obtendo a pontuação de 4,00.

O princípio a ser analisado a seguir será o princípio *Reusable* (Reusabilidade), para essa análise a ferramenta FairDataBR faz uso de cinco itens a ser investigado.

Iniciando com o item “R1. Os conjuntos de (meta)dados estão licenciados?”, de acordo com HL7 (2021), o padrão FHIR possui diversos recursos em sua estrutura nativa que podem viabilizar critérios de segurança de acesso aos dados, a exemplo dos pacotes de segurança dos recursos FHIR, como também relacionados aos direitos autorais legíveis para humanos e em diferentes níveis de granularidade.

Porém, ainda de acordo com HL7 (2021), mesmo com os benefícios disponíveis com a utilização de recursos do FHIR, ainda assim, a adoção do FHIR não pode ser entendida como sendo suficiente para atender a esse requisito. Diante do exposto, a opção de resposta escolhida será à alternativa “Não”.

Seguindo com a análise, o item “R2. Qual a licença de uso dos conjuntos de (meta)dados?” segue a mesma linha que o item R1. Sendo assim, a melhor alternativa de resposta seria opção que indica que o conjunto de metadados não possui licença. Considerando que para essa alternativa as opções de resposta são: “Aberta” e “Restrita”, para essa pesquisa, a resposta desse item fica inconclusiva.

Com o objetivo de gerar a pontuação que reflete o nível de adaptabilidade do FHIR com FAIR, considerando que o FHIR possui em sua estrutura nativa, uma diversidade de opções de recursos que favorecem ao desenvolvimento de itens de segurança, direitos autorais e controle de acesso, a opção de resposta do item R2 será “Aberta”.

Para o item “R4. Os conjuntos de (meta)dados possuem proveniência detalhada?”, destaca-se o alinhamento desse item com a possibilidade de utilização do recurso *Provenance* do protocolo FHIR. Sendo assim, a alternativa escolhida como resposta desse item será “Sim”.

O item “R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados?” segue como sendo um detalhamento do item R4, e da mesma forma que foi apontado para o item R4, o recurso *Provenance* do protocolo FHIR supre essa demanda possibilitando inclusive o rastreamento dessas informações. Sendo assim, diante das alternativas disponíveis, a opção indicada para esse item será “Rastreamento de Proveniência”.

Para finalizar a análise de adaptabilidade do FHIR com os princípios FAIR, a ferramenta FairDataBR possui como último item o “R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio?”. Referente a esse item o HL7 (2021) afirmar que o “FHIR

é um padrão globalmente reconhecido no domínio da saúde, que pode ser usado para representar e informações de dados legíveis por humanos e máquinas”.

Por se tratar de um padrão referência para dados em saúde, entende-se que FHIR possui todos os requisitos desejáveis para o item R6. Diante do exposto, como alternativa de resposta a opção indicada será “Padrões de Domínio”.

Após concluir a análise do princípio *Reusable* (Reusabilidade), foi identificado como resultado da adaptabilidade do padrão FHIR ao referido princípio a nota **8,2**.

Ao término da análise de todos os princípios FAIR, a ferramenta FairDataBR disponibiliza, além das médias indicadas para o nível de adaptabilidade para cada conjunto de princípios, uma média geral. Para a nossa pesquisa, o nível geral de adaptabilidade do FHIR com os princípios FAIR foi de **6,33**, conforme ilustra a figura 13.

Figura 13 - Ferramenta FairDataBR: Nível de adaptabilidade do FHIR aos princípios FAIR



Fonte: FairDataBR (2021)

Analisando o resultado ilustrado na figura 13, fica evidente que os elementos de interoperabilidade do padrão FHIR não contemplam em sua estrutura nativa as recomendações disponíveis no conjunto de princípios de interoperabilidade do FAIR.

Essa constatação, que teve como base a utilização da ferramenta FairDataBR, não afirma que o uso do padrão FHIR não possibilita a realização de interoperabilidade de dados em saúde, significa sim afirmar que essa interoperabilidade não está alinhada aos princípios FAIR.

Seguindo com o desenvolvimento dessa pesquisa, vê-se como necessário a realização da investigação do nível de adaptabilidade dos princípios FAIR com o padrão OpenEHR, uma vez que esse também está sendo utilizado como padrão de referência para o desenvolvimento dessa pesquisa.

5.3.2 Alinhamento do OpenEHR com os princípios FAIR

Assim como o padrão FHIR, o OpenEHR foi escolhido como sendo um padrão base para o desenvolvimento dessa pesquisa por ser um padrão aberto, de acesso livre e por possuir uma comunidade ativa de alcance internacional.

Outro ponto de extrema relevância na escolha do OpenEHR como padrão de referência está no fato desse padrão possuir uma estrutura multinível que possibilita, entre outros benefícios, o desenvolvimento da estrutura de dados independente da estrutura técnica.

Nesse sentido, da mesma forma que foi feito com o padrão FHIR, vê-se como relevante realizar uma análise do nível de adaptabilidade desse padrão com princípios FAIR.

Assim como foi feito com o padrão FHIR, essa análise utilizará os conceitos do padrão OpenEHR nativo, ou seja, as informações disponíveis na página oficial sobre os modelos de referência de estruturas OpenEHR. Essas informações serão analisadas com o auxílio da ferramenta FairDataBR, e ao final espera-se identificar o nível de adaptabilidade do OpenEHR com os princípios FAIR.

Dando início a análise de adaptabilidade do OpenEHR aos princípios FAIR, temos como sendo o primeiro conjunto de princípios o conjunto *Findable* (Localizável) com os seguintes questionamentos:

F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente?

F2. Os conjuntos de dados são descritos com metadados?

F3. O identificador está incluso em todos os registros / arquivos de metadados que descrevem os dados?

F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web?

F5. O conjunto de (meta)dados está publicado em um repositório?

Referente aos itens do conjunto de princípios Localizável, e considerando as opções de resposta disponíveis na ferramenta FairDataBR, para o item identificado como F1 a resposta que mais se adequa é “Endereço *Web*”.

Essa escolha justifica-se pelo fato de OpenEHR ser concebido em uma estrutura multinível que, de acordo com Beale (2022), utiliza identificador único (*Uniform Resource Locator*) como identificador único de todos os elementos, porém, que não se caracteriza como sendo um identificador persistente.

Referente ao item F2, de acordo com Beale (2022), os conjuntos de dados possuem metadados ricos, porém, a estrutura OpenEHR permite a criação de conjunto de dados sem metadados, uma vez que essa estrutura é opcional.

Diante da constatação e considerando que o item F2 analisa a estrutura de metadados disponível, entende-se que a melhor opção de resposta será “Metadados Simples”.

Analisando o item F3 e considerando que na estrutura nativa do OpenEHR, é possível existir conjunto de dados sem metadados associados, a resposta que melhor se adequa é a opção “Não”.

Dando sequência na análise em questão, o item F4 da ferramenta FairDataBR, analisa a capacidade de localização dos dados e metadados através de ferramentas de pesquisa na Web. Referente a esse item, Beale (2022), destaca que na estrutura padrão do arquétipo existe uma subestrutura de rastreabilidade de viabiliza essa ação.

Ainda referente a esse item, Frexia *et al* (2021), corroboram com Beale quando afirmam que elemento ID do arquétipo pode ser utilizado como chave de pesquisa tanto para conectar os metadados do arquétipo, como também ser usado como chave de pesquisa externa.

Analisando as opções de resposta, nota-se como mais adequado a opção de resposta a alternativa “Sim”.

Finalizando os itens da ferramenta FairDataBR, correspondente aos princípios do conjunto “Localizável” do FAIR, o item F5 analisa a possibilidade de publicação desses dados em um repositório público. Seguindo a mesma linha de raciocínio que foi aplicada ao padrão FHIR, entende-se que esse item está diretamente relacionado a uma base de dados propriamente dita. Sendo que, conforme destacado no início dessa seção, a utilização da ferramenta FairDataBR contemplará a análise de uma estrutura nativa de padrões de interoperabilidade de dados em saúde.

Diante do exposto, segue o mesmo entendimento de que essa questão está fora do escopo desta pesquisa, sendo melhor analisada em uma etapa de implementação do OpenFairEHR e sua base de dados. Sendo assim, entende-se que a melhor alternativa de resposta será “Em Nenhum Repositório”.

Ao final da análise de todos os itens do FairDataBR, referente a esse conjunto de princípios FAIR, foi possível identificar que o nível de adaptabilidade do OpenEHR com o conjunto de princípios FAIR, especificamente sobre o conjunto “Localizável”, alcançou a nota 4,6.

O próximo conjunto de princípios a ser analisado será o conjunto “*Accessible*” (Acessível) do FAIR. Referente a esse conjunto, a ferramenta FairDataBR, disponibiliza os seguintes questionamentos:

A1. Quão acessíveis são os (meta)dados?

A2. Os (meta)dados estão acessíveis online sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?

A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?

A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?

A5. O protocolo (ex: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?

A7. É possível realizar o download dos conjuntos de (meta)dados?

Iniciando a análise dos itens referente aos princípios da acessibilidade, o item A1 indaga a possibilidade de acesso aos metadados. Entende-se que na estrutura nativa do OpenEHR os dados e seus metadados podem ser disponibilizados para diversos perfis de acesso, inclusive públicos. Sendo assim, é possível afirmar que o acesso aos metadados está disponível para o “Acessível ao Público”.

É oportuno destacar que esse acesso pode ser controlado através de criptografia de chaves de acesso conforme afirma Beale (2022).

O item a ser analisado na sequência corresponde ao item A2. Nesse momento se vê como oportuno a contribuição de Frexia *et al* (2021), que destaca a possibilidade de recuperação dos metadados através de consultas realizadas com auxílio da linguagem AQL (*Archetype Query Language*), que consiste em uma linguagem específica para o desenvolvimento de arquétipos OpenEHR. Diante do exposto fica evidente que a melhor opção de resposta para esse item corresponde a alternativa “*Web Service* não padrão”.

O item A3, analisa a possibilidade de acesso aos metadados através do identificador fornecido. Para esse item a melhor opção de resposta será “Sim”, uma vez que o ID faz a comunicação dos metadados com os demais itens do arquétipo e possibilita a recuperação em consultas externas. Ainda nesse item é válido destacar que essa análise considera a existência da estrutura de metadados do arquétipo, uma vez que essa estrutura não se configura como item obrigatório do padrão nativo do OpenEHR.

O próximo item a ser analisado corresponde ao item A4, que investiga a possibilidade de recuperação dos metadados mesmo após a exclusão dos dados. Para analisar esse item, será necessário compreender que a exclusão dos dados, a que se refere o item A4, significa a exclusão do arquétipo nesse caso, como os metadados fazem parte da estrutura do arquétipo, sua recuperação não é possível. Diante do exposto a opção de resposta será a alternativa “Não”.

Contradizendo essa afirmação, Frexia *et al.* (2021) destacam que a recuperação dos metadados sempre ocorrerá, pois a estrutura do OpenEHR não permite a exclusão de dados e metadados em atenção a um normativo legal. Esse autor entende que os dados clínicos devem ser mantidos por um período legal de até 20 anos, conforme define o Conselho Federal de Medicina (CFM), através da resolução CFM nº 1.821/2007.

Ainda na opinião desse autor, a exclusão dos dados e/ou metadados não deve ocorrer mesmo após o período legal, pois esses dados podem e devem ser utilizados como legados para o desenvolvimento de pesquisas científicas por parte da academia. Porém essa opinião não pode ser determinante para definição da resposta para o item A4.

Os itens A5 e A7 analisam, respectivamente, se o protocolo utilizado é aberto e gratuito e se é possível realizar o *download* dos conjuntos de metadados. Para ambos a alternativa de resposta escolhida será a alternativa “Sim”.

Essa escolha justifica-se pela adoção do padrão HTTP para comunicação com o protocolo OpenEHR REST API utilizado pelo padrão OpenEHR. Já para o item A7, considerando a possibilidade de acesso aos metadados e a possibilidade de publicação desses metadados em repositórios públicos não se identifica nenhuma objeção a realização de *download* desses metadados.

Ao final da análise dos itens correspondente aos princípios da Acessibilidade do FAIR, foi identificado a nota 8,17 como sendo o nível de adaptabilidade do OpenEHR aos princípios do conjunto de Acessibilidade do FAIR.

O próximo conjunto a ser analisado corresponde ao princípio da *Interoperable* (interoperabilidade), e para essa análise a ferramenta FairDataBR apresenta as seguintes indagações:

11. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?
12. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?
13. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores?

I4. Para representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico?

Para o item I1, considerando que a definição de formatos preferenciais consiste na escolha por padrões em uso por grande parte da comunidade de usuários, entende-se que o OpenEHR atende essa demanda uma vez que ele faz uso dos padrões de comunicação atualmente disponíveis para todos, a exemplo da API REST, XML e JSON. Diante do exposto, dentre as opções de resposta disponíveis na ferramenta FairDataBr, a opção escolhida foi “Sim”.

Dando seguimento na análise, o Item I2 analisa a estrutura do conjunto de metadados utilizados no padrão OpenEHR. Para esse item a opção de resposta indicada será “Esquema Explícito”. Essa escolha justifica-se quando analisamos a documentação oficial do OpenEHR disponível em Beale (2022) e que disponibiliza a estrutura do Modelo de Informação Comum, e que nesse modelo contém as informações referentes a estrutura básica dos metadados. Com base nessa documentação fica evidenciado a existência de uma estrutura básica de informações para os metadados. Destaca-se nesse ponto que essa estrutura foi criada para uso das aplicações OpenEHR.

O item I3 do FairDataBR busca investigar a existência de conectividade entre os conjuntos de metadados no padrão OpenEHR. De acordo com Frexia *et al.* (2021) a documentação do padrão OpenEHR define diversos tipos de referências cruzadas, a exemplos das classes LINK e DV_URI. Porém, a informação de relação com o link de destino não está qualificada, sendo necessário que essa informação seja disponibilizada no conjunto de metadados. Sendo assim, diante das opções de resposta disponíveis, entende-se que a melhor opção será a alternativa “*Links* Internos/Externos”.

Para finalizar a análise do conjunto de princípios correspondente a interoperabilidade da ferramenta FairDataBr, o item I4 analisa a utilização, ou não, de ferramentas de controle terminológicos.

De acordo com Beale (2022), no documento Visão Geral da Arquitetura OpenEHR, está detalhado o componente “Interface Terminológica” que possibilita a vinculação dessa estrutura terminológica, através do uso de identificadores, com um ou mais códigos de terminologias externos. Sendo assim, entende-se que a melhor opção de resposta para o item I4 será “Uso de instrumentos de controle terminológico COM identificador global”.

Ao final da análise de todos os itens correspondentes aos princípios da interoperabilidade do FAIR, através da ferramenta FairDataBR, identificou-se que o nível de

adaptabilidade do padrão OpenEHR com os princípios de interoperabilidade do FAIR possui uma nota de 8,75.

Para finalizar a análise de adaptabilidade do padrão OpenEHR com os princípios FAIR, serão analisados os itens correspondentes ao conjunto de princípios *Reusable* (Reusabilidade) do FAIR.

Para esse conjunto de princípios, a ferramenta FairDataBR disponibiliza cinco itens de investigação, sendo eles:

R1. Os conjuntos de (meta)dados estão licenciados?

R2. Qual a licença de uso dos conjuntos de (meta)dados?

R4. Os conjuntos de (meta)dados possuem proveniência detalhada?

R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados?

R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio?

Iniciando a análise desses princípios, o item R1 avalia a existência de licença de uso dos metadados. Nesse viés, destaca-se o citado por Frexia *et al.* (2021), quando afirmam que cada recurso pode ser entendido como sendo uma instância de uma classe `AUTHORED_RESOURCE` que contém, entre outros, metadados relacionados à autoria, direitos autorais, licenças, idiomas e traduções.

Para o item R1, considerando as opções existentes, entende-se que a melhor alternativa de resposta será a opção “Sim”.

O item R2 segue na mesma linha de investigação que o item R1, porém, conforme foi detectado na análise do nível de adaptabilidade do padrão FHIR com os princípios FAIR, a análise desse item se aplica melhor para bases de dados e não padrões de interoperabilidade. Diante dessa afirmação, a análise desse item se torna inconclusiva.

Da mesma forma como foi feita na análise desse item para o padrão FHIR, considerando as opções de resposta disponíveis na ferramenta FairDataBR, a opção escolhida será “Aberta”.

Como opção de resposta para o item R4, uma vez que esse item investiga a existência de controle de proveniência detalhada, a opção de resposta mais adequada será a alternativa “Sim”.

Tal resposta justifica-se pela existência de um controle de versão na estrutura do OpenEHR que possibilita a auditoria de alterações de dados do EHR e serviços demográficos através da conferência de identidade do usuário, assinatura digital, entre outros.

O item R5 configura-se como sendo um complemento ao item R4, uma vez que o questionamento disponível no item R5 faz referência à existência, ou não, de uma descrição da origem do fluxo de trabalho que produziu os metadados.

Destaca-se nesse o ponto, citado por Beale (2022), referente aos elementos de rastreabilidade existentes na estrutura do OpenEHR. Corroborando com essa afirmação Frexia *et al.* (2021), destaca a presença de elementos de segurança da informação como garantia ao não repúdio e autenticação através de assinaturas digitais.

Essas contribuições, citadas por Beale e Frexia, evidenciam a possibilidade de rastreabilidade dos metadados. Diante do exposto, considerando as alternativas disponíveis na ferramenta FairDataBR, a melhor opção de resposta será a alternativa “Rastreamento de Proveniência”.

Para finalizar a análise dos princípios de reusabilidade, bem como finalizar todos os itens disponíveis na ferramenta FairDataBR, o item R6 investiga se a estrutura de metadados está de acordo com padrões relevantes para o domínio.

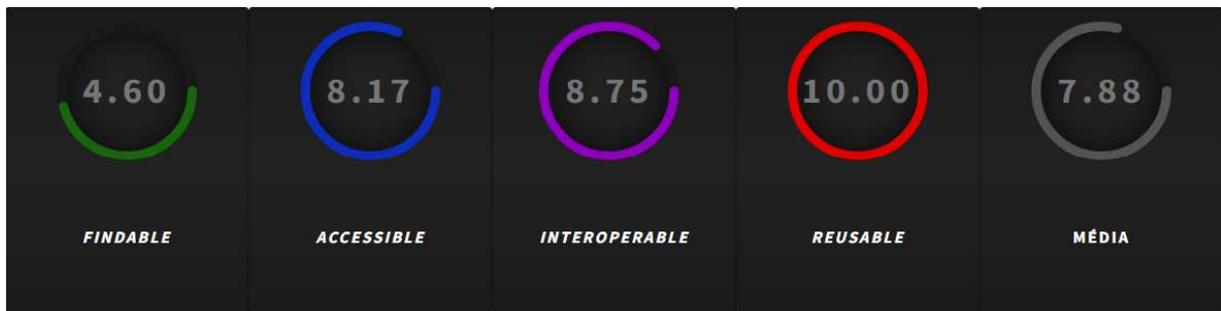
Nesse quesito, é oportuno citar que a estrutura do OpenEHR teve como referência padrões de dados conhecidos, a exemplo ISO 8601, além de possibilitar, através da sua estrutura de terminologia, a conexão com outros padrões igualmente reconhecidos no mercado, a exemplo do SNOMED-CT e LOINC.

Sendo assim, considerando as opções de resposta disponíveis, entende-se que a melhor opção de resposta para o item R6 será a alternativa “Padrões de Domínio”.

Finalizando o conjunto de itens referente aos princípios da reusabilidade, da mesma forma que foi realizado com os demais conjuntos de princípios, a ferramenta FairDataBR, destaca que o nível de adaptabilidade do padrão OpenEHR com os princípios do conjunto Reusabilidade do FAIR possui uma nota 10,00.

A nota identificada na análise dos itens referentes ao conjunto de princípios de reusabilidade corresponde a nota mais alta que a ferramenta FairDataBR considera. Como forma de ilustrar o resultado do nível de adaptabilidade do padrão OpenEHR com os princípios FAIR, a figura 14 ilustra todos os resultados, bem como a avaliação geral, que teve como resultado a nota **7.88**.

Figura 14 - Ferramenta FairDataBR: Nível de adaptabilidade do OpenEHR aos princípios FAIR



Fonte: FairDataBR (2021)

O resultado identificado na figura 14 evidencia um nível de adaptabilidade do padrão OpenEHR com os princípios FAIR. Destaca-se nesse ponto que algumas alterações são essenciais para a manutenção desse resultado, a exemplo da obrigatoriedade da existência dos metadados para cada um dos arquétipos criados no padrão OpenEHR.

Confrontando os resultados identificados nas figuras 13 e 14 que, respectivamente, apresentam os resultados das análises de adaptabilidade dos padrões FHIR e OpenEHR com os princípios FAIR, percebe-se que o OpenEHR está bem mais adaptado aos princípios FAIR.

Porém, isso não significa que o padrão FHIR não apresente elementos relevantes para o desenvolvimento de interoperabilidade que é o cerne dessa pesquisa.

A conclusão que se chega através dessas análises reflete um resultado esperado por esse autor, uma vez que o padrão OpenEHR, através do desenvolvimento de seus arquétipos e modelos, possui uma atenção maior com a questão da semântica se comparado com o padrão FHIR.

Por sua vez, o padrão FHIR possui uma estrutura tecnológica mais robusta que o padrão OpenEHR, pois esse teve na sua concepção o objetivo de possibilitar a troca de informações entre sistemas e equipamentos sem necessariamente se preocupar com o viés da semântica dos dados compartilhados.

É oportuno destacar que, conforme indicado na análise realizada com o padrão FHIR, o resultado identificado na análise do padrão OpenEHR realizada com a ferramenta FairDataBR também encontra-se disponível no anexo B dessa pesquisa.

Na seção seguinte, será iniciado a discussão referente ao desenvolvimento do padrão de interoperabilidade proposto por essa pesquisa e que terá como base elementos dos padrões FHIR e OpenEHR.

5.4 PROPOSTA DE MODELO OPENFAIREHR DE INTEROPERABILIDADE DE DADOS EM SAÚDE

A seção em questão se debruça sobre o processo de desenvolvimento de todos os elementos que constituirão o modelo de interoperabilidade de dados em saúde OpenFairEHR proposto por essa pesquisa. Destaca-se nesse momento que OpenFairEHR se configura como sendo um modelo conceitual, que objetiva viabilizar a interoperabilidade de dados em saúde com base nas orientações dos princípios FAIR, em especial o conjunto de princípios da interoperabilidade.

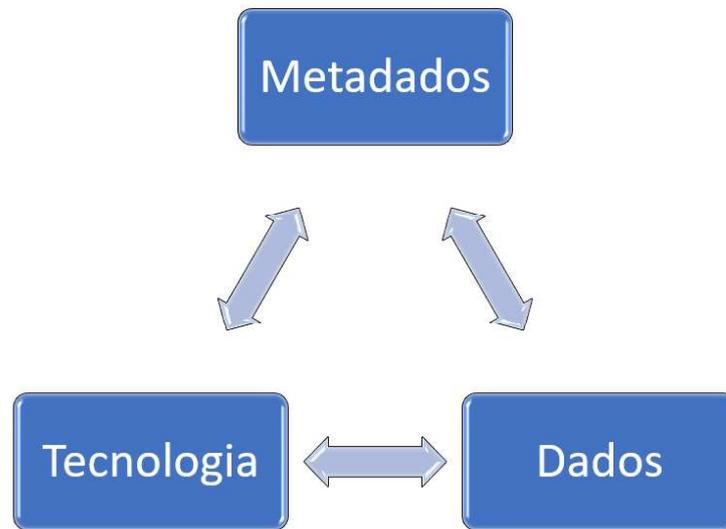
A fundamentação teórica necessária para o desenvolvimento desse modelo está presente nas seções anteriores dessa pesquisa. Nesse sentido, os tópicos que abordam os temas: princípios FAIR, Padrões FHIR e OpenEHR e a análise de adaptabilidade desses padrões com os princípios FAIR, ganham destaque por serem considerados elementos estruturantes e norteadores para esse desenvolvimento.

5.4.1 Proposta de Arquitetura

A escolha pela abordagem da arquitetura se justifica pela possibilidade de representação do modelo proposto através de estruturas independentes que estão interconectadas com o objetivo de proporcionar a interoperabilidade de dados em saúde alinhado aos princípios FAIR.

Diante do exposto, vale destacar que o modelo proposto por essa pesquisa possui em sua estrutura macro um conjunto de três camadas, conforme ilustra a figura 15, sendo elas as camadas de Metadados, Tecnologia e Dados.

Figura 15 - Elementos que compõem o modelo de interoperabilidade proposto



Fonte: Elaborado pelo Autor (2022)

A escolha por essa estrutura levou em consideração a composição arquitetural dos padrões de interoperabilidade OpenEHR e FHIR. Em sua documentação oficial esses padrões possuem estrutura de camadas bem definidas, porém, com algumas particularidades entendidas por esse autor com sendo impactantes quando se almeja o alinhamento com os princípios FAIR.

Dentre essas particularidades, destaca-se o fato desses modelos preservarem estruturas independentes de dados e tecnologia, porém, mantém uma unicidade entre as estruturas de dados e metadados. Essa realidade potencializa a complexidade de localização, acesso, uso e reuso desses metadados e, conseqüentemente, dos seus dados.

Diante do exposto, ciente de que em uma estrutura de metadados pode ser concebida no próprio recurso de dados ou separadamente como sendo um banco de dados (que nesse caso poderia ser entendido como sendo um banco de metadados), esse autor adotou a segunda opção como estratégia a ser adotada nesse modelo de metadados.

Nos tópicos seguintes serão realizadas discussões individualizadas, referente a cada umas dessas camadas, com o objetivo de apresentar sua estrutura funcional, elementos estruturantes e conexões com os outros chamados do modelo proposto.

5.4.2 Camada Metadados

Como já mencionado, os metadados podem ser compreendidos como sendo elementos identificadores dos próprios dados, popularmente conhecidos como “dados dos dados”.

Segundo Formenton *et al.* (2017), possuem como principal função a descrição de um recurso informacional inequívoco, com múltiplas formas de acesso, garantindo sua recuperação pelo usuário final ou sistemas informacionais convencionais ou digitais.

Essa compreensão sobre os metadados está bem alinhada com os objetivos dessa pesquisa, uma vez que também compreendemos que são através dos metadados que vamos conseguir alcançar o nível de interoperabilidade semântica que se almeja como resultado dessa pesquisa.

Ainda sobre a camada de metadados do modelo proposto, é oportuno resgatar o que foi descrito na seção “Justificava e Hipóteses” onde se define o escopo de abrangência que se espera alcançar com esse desenvolvimento. Nesse sentido, destacam-se que os elementos que serão contemplados nesse modelo são elementos que compõem um RES de acordo com HIMSS, sendo eles:

1. Dados demográficos do paciente,
2. Notas de evolução do paciente,
3. Problemas,
4. Medicamentos,
5. Sinais vitais,
6. Histórico médico,
7. Imunizações,
8. Dados laboratoriais e
9. Relatórios (laudos) de radiologia.

Analisando o conjunto de dados a ser trabalhado, é possível inferir que sua recuperação pode ser motivada para suprir uma infinidade de objetivos. Essa constatação evidencia a necessidade de desenvolvimento de uma estrutura de armazenamento de dados que possibilite sua recuperação total e/ou parcial, dependendo da finalidade do seu uso.

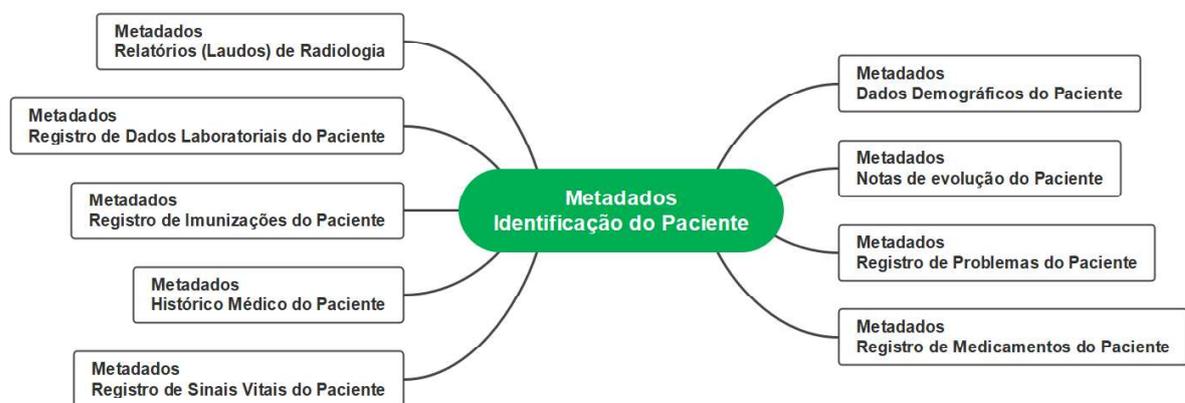
Com o objetivo de exemplificar a necessidade de utilização desses dados, podemos citar a utilização dos dados demográficos do paciente para o desenvolvimento de políticas públicas por parte de sistemas e/ou usuários vinculados aos governos municipais, estaduais e federais.

Seguindo ainda com esse exemplo, considerando a finalidade apresentada, fica evidente que esses sistemas e/ou usuários não necessitam de acesso aos dados laboratoriais ou relatórios de radiologia dos pacientes. Sendo assim, pensar em uma disponibilização que priorize o

mínimo necessário e que busque proporcionar a privacidade e proteção desses dados são essenciais.

Como alternativa para estruturação da camada de metadados do modelo proposto, objetivando a localização, acessibilidade e recuperação desses dados de forma atender as necessidades informacionais dos usuários e sistemas de informação propõe-se a adoção de uma estrutura de metadados que possibilite a sua segmentação, considerando o tipo de dado que se deseja localizar, conforme ilustra a figura 16.

Figura 16 - Camada de Metadados do modelo proposto



Fonte: Elaborado pelo Autor (2022)

Como pode ser identificado na figura 16, o elemento central dessa proposta de estrutura consiste nos metadados de identificação do Paciente. Essa escolha não foi ao acaso, ela se fundamenta no princípio de que o elemento base em qualquer EHR ou RES (Registro Eletrônico de Saúde) são os dados de identificação do paciente.

Partindo desse princípio, e considerando a utilização dos conceitos de identificadores únicos e identificadores únicos compostos, é possível, através desses metadados, localizar e recuperar qualquer conjunto de dados independente do agrupamento que seja necessário montar.

A estrutura ilustrada na figura 16 ainda possibilita a expansão desse modelo através da inclusão de outros tipos de dados, além do que já está definido, bastando para isso contemplar a inclusão de uma nova categoria de metadados e uma nova estrutura de dados a ser inserida na camada de dados.

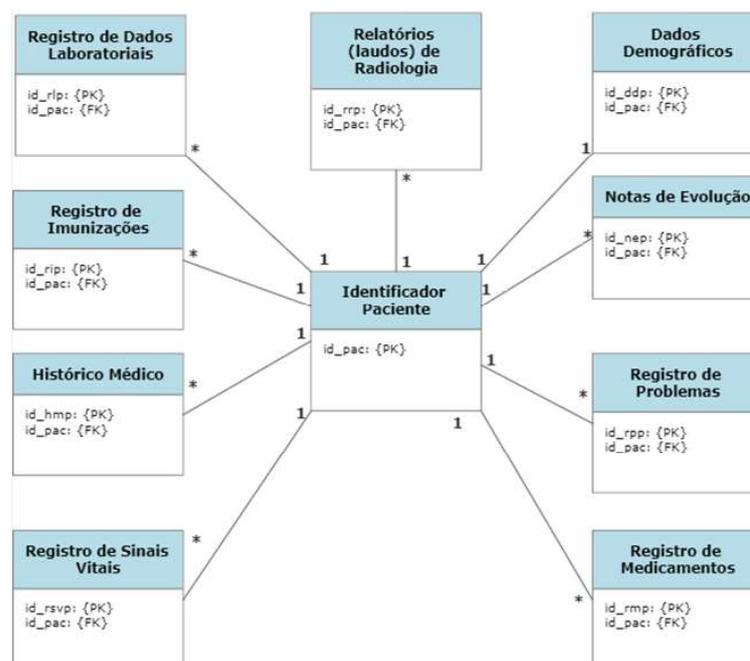
Outro benefício identificado nessa proposta consiste na possibilidade de implantação rotinas de privacidade e proteção desses dados que devem ser aplicadas a dados pessoais que,

segundo Brasil (2018), consiste em informação relacionada à pessoa natural identificada ou identificável.

Partindo dessa conceitualização, percebemos que os dados pessoais estão presentes apenas na estrutura de metadados de identificação do paciente. Essa constatação exclui a necessidade de desenvolvimento e implantação de rotinas de segurança de dados que busquem proporcionar a privacidade e proteção de dados em todos os conjuntos de metadados do modelo, bastando para isso aplicar técnicas de anonimização dos dados pessoais, possibilitando seu acesso completo apenas para profissionais da assistência que deles precisar para proceder com o fornecimento dos serviços de saúde e proteção a vida do paciente.

Com o objetivo de apresentar o processo de conexão entre as estruturas de metadados, a figura 17 ilustra o modelo de utilização de identificadores únicos e identificadores únicos compostos.

Figura 17 - Diagrama de Metadados do modelo proposto



Fonte: Elaborado pelo Autor (2022)

Como pode ser observado na figura 17, todas as estruturas de metadados estão interconectadas através de uma ligação direta com a estrutura de metadados identificação do paciente.

Também é possível perceber que essa conexão será viabilizada através da utilização do recurso de chaves primárias ou *primary key* (PK) e chaves estrangeiras ou *foreign key* (FK).

Destaca-se nesse ponto que a atenção referente à estrutura de identificadores únicos e persistentes exerce muita influência no processo de avaliação da adequação do modelo com as recomendações existentes nos princípios FAIR.

Dando sequência na análise da figura 17, fica evidenciada a necessidade obrigatória da existência de uma estrutura de metadados de identificação de paciente, para que seja possível montar um conjunto completo de dados necessários para o RES. A figura ilustra essa condição através da cardinalidade presente em cada linha de conexão entre as estruturas de metadados do modelo.

Em quase todas as conexões, com exceção da conectividade entre metadados de identificação do paciente e metadados de dados demográficos do paciente que possui uma cardinalidade de 1 para 1, a cardinalidade definida é de 1 para n, sendo n representado por * na figura 17.

A cardinalidade 1 para n, ilustrada na figura 17, implica dizer que para cada registro existente na estrutura de metadados de identificação do paciente, é aceitável a existência de relacionamento com vários registros nas demais estruturas de metadados do modelo. inclusive a ausência de elementos, do lado do n na figura 17 também é válida.

Quando analisamos a cardinalidade n para 1, fica evidenciado a obrigatoriedade da existência de um registro da estrutura de metadados de identificação do paciente para que seja possível a inclusão dos dados nas demais estruturas de metadados do modelo.

A proposta de conexão ilustrada na figura 17 contemplará uma conexão simples, utilizando apenas PK, entre metadados de identificação do paciente e as demais estruturas de metadados do modelo e uma conexão composta, utilizando a chave composta PK + FK, para conexão entre as estruturas de metadados do modelo, sem necessariamente acessar os dados disponíveis na estrutura de metadados de identificação do paciente.

Essa proposta de conexão possibilita a criação de uma infinidade de agrupamentos de dados que possam atender toda e qualquer necessidade de recuperação desses dados para suprir qualquer finalidade, desde que, obviamente, essa finalidade seja satisfeita com os dados disponíveis na base alimentada pelo modelo.

Outro ponto relevante no desenvolvimento de uma estrutura de metadados de acordo com os princípios FAIR, diz respeito à qualidade dos elementos de dados que estão disponíveis na estrutura de metadados. Essa característica melhor de que os metadados do modelo podem ser classificados como sendo metadados ricos ou com estrutura de dados simples ou limitada.

Referente a esse tópico, para o desenvolvimento do modelo em questão foi realizada uma busca na literatura com o objetivo de identificar padrões de metadados adequados aos

objetivos que o modelo se propõe e com isso foi identificado o padrão *Dublin Core* que será melhor descrito no tópico seguinte.

5.4.2.1 Padrão de Metadados

A estrutura de metadados do modelo proposto, na visão desse autor, apresenta-se como sendo elemento essencial na busca pela interoperabilidade de dados de saúde em ambientes digitais. Essa afirmação justifica-se, em especial, pela heterogeneidade semântica devido a existência de diversas fontes de dados que por sua vez possibilitam diversas interpretações da mesma informação, provocando assim conflitos de sinonímia (dados com mesmo conteúdo semântico, porém divergência em seus nomes) e homonímia (dados com nomes iguais e com conteúdo semânticos diferentes).

Como alternativa para solucionar esse e outros desafios, foi adotado para o desenvolvimento dessa pesquisa a implementação de um padrão de metadados para ser utilizado como estrutura base no desenvolvimento do modelo proposto.

A escolha do padrão de metadados a ser utilizado, assim como foi feito na escolha dos padrões de interoperabilidade FHIR e OpenEHR, levou em consideração o fato de ser um padrão de uso nacional e internacional, possuir uma comunidade de apoio atuante, ter seu conteúdo disponível de forma gratuita e ser reconhecido como referência.

Levando em consideração todos esses requisitos, dentre as mais diversas opções disponíveis o padrão de metadados *Dublin Core* (DC) foi o padrão escolhido para orientar a construção dos elementos de metadados do modelo proposto.

O padrão de metadados DC configura-se como sendo um conjunto de 15 elementos, com o objetivo de descrever um recurso eletrônico (GRÁCIO, 2012).

O padrão DC foi desenvolvido com ênfase na simplicidade semântica de seus elementos, interoperabilidade semântica obtida através da adoção de um padrão único para criação de metadados diversos, consenso internacional através de uma comunidade de apoiadores que contam com a participação de membros de mais de 20 países, extensibilidade e flexibilidade devido a possibilidade de inclusão de novos elementos e modificação dos elementos existentes.

É oportuno destacar que a proposta de utilização do padrão de metadados DC consiste no desenvolvimento de uma estrutura básica de elementos que deverão constar na camada de metadados do modelo proposto por essa pesquisa. Ainda sobre esse tópico, vê-se como relevante destacar que diante da necessidade é completamente aceitável a inclusão de outros elementos, além dos já previstos na estrutura básica do modelo proposto.

De acordo com a DCMI (2022), os 15 elementos previstos no padrão DC são:

1. **DC.título** – Corresponde ao nome dado ao recurso, pelo qual ele é formalmente conhecido;
2. **DC.autor (Criador)** – Responsável pela criação do conteúdo intelectual do recurso, podem ser pessoas ou organizações;
3. **DC.data** – Data em que o recurso ficou disponível em seu formato atual.
4. **DC.palavra-chave** – Deverá conter a essência do conteúdo do recurso, disponível através de palavras ou frases que remetam ao assunto ou ao conteúdo do recurso;
5. **DC.editor** – Responsável por disponibilizar o recurso em seu formato atual;
6. **DC.formato** – Informações sobre o formato do recurso, possibilitando a identificação de hardwares ou softwares necessários para seu acesso;
7. **DC.colaborador** – Identificação de pessoa ou organização que não tenha sido especificado como autor, que tenha contribuído com o desenvolvimento do recurso;
8. **DC.identificador** – Identificador único e inequívoco do recurso;
9. **DC.descrição** – Uma descrição do recurso. Pode ser compreendido como sendo um breve resumo do conteúdo do recurso;
10. **DC.direitos_ autorais** – Uma declaração ou referência a uma declaração para qualquer tipo de licença, a exemplo para uso do conteúdo do recurso;
11. **DC.tipo** – Define o tipo do recurso, ex. imagem, texto, mídia;
12. **DC.fonte** – Indica um segundo recurso, do qual esse seja derivado;
13. **DC.idioma** – Indica o idioma do recurso;
14. **DC.relação** – Indica relacionamentos entre recursos que têm relação formal com outros, mas que existem por si mesmos;
15. **DC.cobertura** – Características espaciais ou temporais do conteúdo do recurso.

Além dos elementos descritos no padrão DC, para o desenvolvimento do modelo proposto serão acrescentados os seguintes elementos:

16. **DC.cobertura.ciclo_de_vida** – Indica o estado do ciclo de vida do recurso, deverá refletir possíveis atualizações no recurso;
17. **DC.direitos_ autorais.licença** – Descreve o tipo de licença do artefato;

18. **DC.descrição.finalidade** – Descreve as finalidades para as quais esse recurso pode ser utilizado.

19. **DC.identificador.localizador** – Um *Archival Resource Key* (ARK) Identificador persistente multifuncional.

É oportuno destacar que, objetivando a integridade do DC, conseqüentemente favorecendo a interoperabilidade dos dados, os elementos acrescentados no modelo proposto não descaracterizam a estrutura padrão do DC, pois, esses foram inseridos como sendo extensões de elementos existentes.

A escolha pela identificação dos elementos inseridos no modelo proposto, como sendo extensões de elementos já existentes no padrão DC, se justifica pela necessidade de compreensão da sua finalidade no conjunto, uma vez que a simples repetição do elemento padrão do DC pode proporcionar ambigüidade no entendimento e com isso conduzir ao erro em sua interpretação.

Considerando os objetivos almeçados com o desenvolvimento do modelo proposto, bem como a dinâmica de inserção, manutenção e atualização de dados de saúde de um RES, os elementos que serão contemplados como base para o desenvolvimento da camada de metadados são: DC.título, DC.autor, DC.data, DC.palavra_chave, DC.identificador, DC.descrição, DC.direitos_autorais, DC.tipo, DC.fonte, DC.idioma, DC.relação, DC.cobertura.ciclo_de_vida, DC.direitos_autorais.licença, DC.descrição.finalidade e DC.identificador.localizador.

Além desses elementos cada conjunto de metadados, ilustrado na figura 16, pode fazer uso de outros elementos, inclusive sendo essa uma prática encorajada, específicos para sua finalidade. O quadro 3 ilustra o conjunto de elementos definidos para cada conjunto de metadados do modelo proposto.

Quadro 4 - Conjunto de elementos básicos para os Metadados

Metadados	Inserir identificação da categoria de Metadados	
Resumo	Inserir uma descrição da categoria de Metadados	
Elemento	Descrição	Observação
DC.título	Corresponde ao nome dado ao recurso, pelo qual ele é formalmente conhecido	Deverá ser informado o tipo de dado de saúde contemplado nessa categoria de metadados (ex.: dados laboratoriais, dados demográficos etc.)

DC.autor	Responsável pela criação do conteúdo intelectual do recurso, podem ser pessoas ou organizações	Sempre será inserido o nome do profissional ou organização responsável pelo conteúdo do dado.
DC.data	Data em que o recurso ficou disponível em seu formato atual considerando, inclusive, as atualizações.	Indicar a data de inclusão do dado na base. Destaca-se aqui que a informação referente a data de criação do dado constará no registro do dado, na camada dados do modelo
DC.palavra_chave	Deverá conter a essência do conteúdo do recurso, disponível através de palavras ou frases que remetam o assunto ou o conteúdo do recurso	Inserir elementos que remetam ao tipo de procedimento realizado, diagnóstico identificado ou dados demográficos do paciente.
DC.identificador	Identificador único e inequívoco do recurso	Inserir registros que remetam ao conteúdo do recurso (ex.: Número do Prontuário + id_registro) <i>Obs. id_registro corresponde ao id de identificação do dado na camada de dados do modelo.</i>
DC.descrição	Uma descrição do recurso. Pode ser compreendido como sendo um breve resumo do conteúdo do recurso	Texto de resumo que contextualize o dado de saúde representado por esse metadado.
DC.direitos_autorais	Uma declaração ou referência a uma declaração de direitos autorais para uso do conteúdo do recurso	Indicar a URL onde é possível acessar o registro de direitos autorais desse dado.
DC.tipo	Define o tipo do recurso, ex. imagem, texto, mídia	Imagem, texto ou mídia
DC.fonte	Indica um segundo recurso, do qual esse seja derivado	1. Sim 2. Não <i>Obs. Para resposta "Sim", indicar "Identificador" do metadado relacionado.</i>
DC.idioma	Indica o idioma do recurso	Idioma padrão de registro
DC.relação	Indica relacionamentos entre recursos que têm relação formal com outros, mas que existem por si mesmos	Indicar o "Título" dos metadados relacionados. <i>Obs. A busca ocorrerá através do "Identificador" desses metadados.</i>
DC.cobertura.ciclo_de_vida	Indica o estado do ciclo de vida do recurso, deverá refletir possíveis atualizações no recurso	Terá seu estado alterado sempre que sofrer atualização
DC.direitos_autorais.licença	Descreve o tipo de licença do recurso	Indicar a URL onde é possível acessar o registro de licença de uso desse dado.
DC.descrição.finalidade	Descreve as finalidades para qual esse recurso pode ser utilizado. A indicação da finalidade servirá como	1. Pesquisa 2. Tutela a saúde 3. Proteção a vida 4. Obrigação

	parâmetro para a definição de estratégias de segurança.	Legal 5. Políticas públicas
DC.identificador.localizador	Um <i>Archival Resource Key</i> (ARK) identificador persistente multifuncional	Indica registro utilizado para localização do dado por motores de busca da internet, sistemas ou usuários.

Fonte: Elaborado pelo Autor (2022)

No quadro 3 é possível identificar todos os elementos definidos como metadados básicos da camada de metadados do modelo proposto. Destaca-se nesse ponto que a estrutura aqui apresentada pode ser acrescida de outros elementos necessários para uma melhor utilização dos dados existentes, bem como para contemplar novas estruturas de metadados que sejam inseridas no modelo de interoperabilidade proposto.

Ainda no quadro 3, mais especificamente no elemento “Identificador”, é possível ver uma conexão da camada de metadados com a camada dados do modelo proposto. Essa conexão se configura com a presença do “id_registro”, que nesse modelo representa o identificador único do registro do dado na camada de dados do modelo.

É oportuno citar que mesmo os registros classificados como atualização no elemento “Ciclo de vida” deverão ter seu “id_registro” único e serão tratados como dados independentes. Por se tratar de um modelo de interoperabilidade de dados em saúde dinâmico, a inserção de dados classificados como atualização, que no nosso caso se referem a dados clínicos atualizados de um mesmo paciente, será constante e essa medida de utilização de “id_registro” único objetiva garantir a integridade e rastreabilidade dos dados e seus metadados.

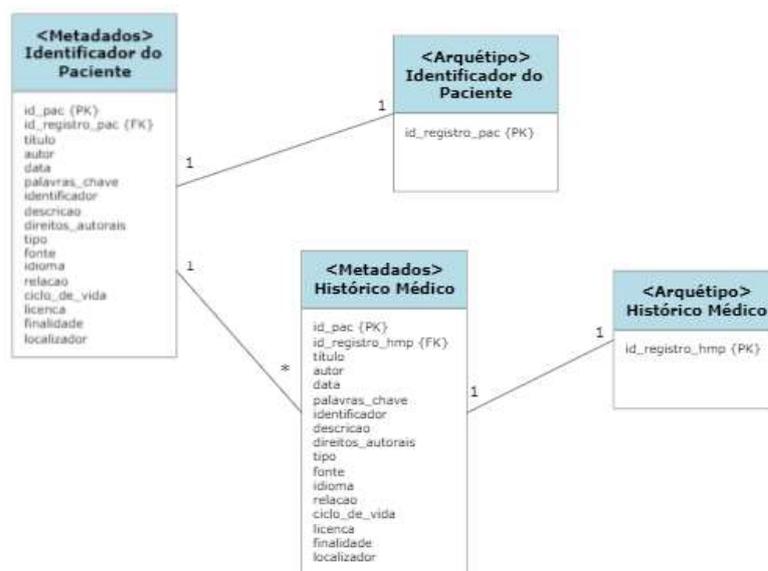
No que se refere à recuperação desses recursos de forma persistente propõe-se a utilização do elemento “Localizador”, sendo este um *Archival Resource Key* (ARK) que, segundo ARK (2021) servem como identificadores persistentes ou referências estáveis e confiáveis para objetos de informação.

Ainda referente à utilização do elemento “Localizador”, considerando a possibilidade de um grande volume de atualizações de dados dos pacientes nos diversos recursos de metadados, está sendo proposto a utilização da técnica de *Suffix Passthrough in Action* que, segundo ARK (2021), consiste em definir um sufixo padrão, atribuído ao endereço Web (URL) do recurso, e todos os demais elementos que fazem uso desse sufixo serão referenciados pelo mesmo ARK.

Essa estratégia se fundamenta na necessidade de recuperação de vários registros de atualização para um mesmo elemento de metadados, uma vez que cada registro de atualização será tratado como sendo um elemento independente em nosso modelo.

A figura 18 ilustra uma conexão entre os metadados e os dados disponibilizados na camada de dados do modelo proposto.

Figura 18 - Elementos de metadados e conexão entre metadados e dados



Fonte: Elaborado pelo Autor (2022)

É possível identificar na figura 18 elementos com *tags* <Metadados> e <Arquétipo>, a escolha por essa abordagem justifica-se na facilidade de identificação dos recursos da camada metadados e da camada dados do modelo proposto. A alusão ao termo Arquétipo fundamenta-se na opção de fazer uso da estrutura de arquétipos desenvolvida pelo padrão OpenEHR. Essa escolha será melhor descrita no item Camada Dados dessa pesquisa.

Ainda referente à Camada Metadados, a estrutura proposta foi pensada com o objetivo de garantir os seguintes objetivos:

1. **Rastreabilidade** – Através do uso de chaves primárias e estrangeiras é possível obter conectividade interna entre os metadados, da mesma forma que através da utilização do elemento “Localizador” para cada registro de recurso é possível acessar os metadados através de recurso de busca disponíveis na Web;

2. **Confiabilidade** – O elemento “Identificador” possibilita a unicidade do recurso através da identificação única. Os elementos “Data” e “Ciclo de vida” permitem a identificação de alterações no recurso através do seu versionamento;
3. **Segurança** – Através do elemento “Finalidade” será possível definir níveis de acesso aos dados e metadados com o uso de *flags* definidos como regras de negócios;
4. **Privacidade e Proteção de Dados** – A estrutura de metadados e dados referentes à identificação do paciente são mantidos segmentados dos demais dados, isso possibilita a aplicação de regras de acesso e/ou a aplicação de técnicas de minimização de dados apenas nesse conjunto;
5. **Integridade** – A segmentação entre dados e metadados possibilita uma continuidade de uso dos metadados mesmo que a estrutura de dados se torne inacessível.

Como pode ser visto nos objetivos acima, muitos fazem referência à conexão entre metadados e dados. Sendo assim, no tópico seguinte será apresentado a estrutura proposta para a Camada Dados do modelo proposto.

5.4.3 Camada Dados

Para o desenvolvimento do OpenFairEHR a estrutura da Camada Dados vê-se como de extrema relevância, porém, de acordo com a proposta de desenvolvimento do OpenFairEHR seu desenvolvimento deverá ser conduzido de maneira independente do desenvolvimento das camadas Metadados e Infraestrutura.

Na proposta conceitual do OpenFairEHR, a estrutura da Camada Dados seguirá o modelo definido no padrão OpenEHR, em que é definido a criação de arquétipos e modelos de dados.

Nesse momento, com o intuito de enriquecer a discussão sobre a proposta definida para a concepção da Camada Dados, vê-se como oportuno revisitar os conceitos de arquétipos e modelos de dados do OpenEHR.

A estrutura de desenvolvimento do OpenEHR contempla o conceito de multicamadas ou multinível. Nessa estrutura os elementos são segmentados e seu desenvolvimento pode ser conduzido por equipe de especialistas no assunto. Para o OpenEHR essa segmentação ocorre em dois níveis, sendo eles o de tecnologia e o de domínio, corresponde ao nível de domínio a

definição de estruturas de dados que representem os dados clínicos de pacientes e por isso devem ser construídos por profissionais das áreas assistenciais que possuem pleno conhecimento da sua área de atuação.

Ainda referente ao nível de domínio, destacam-se as estruturas de arquétipos e modelos. Segundo Bale (2015), o arquétipo consiste em uma definição formal e reutilizável de informações de nível de domínio, definidas em termos de restrições em um modelo de informações, está completamente separada das estruturas de software e banco de dados em um contexto OpenEHR.

No que se refere à conceitualização do termo modelo, Bale (2015) destaca que um modelo consiste em um conjunto de dados que agrega pedaços de outros arquétipos para um uso específico, por exemplo, um formulário de tela ou definição de mensagem.

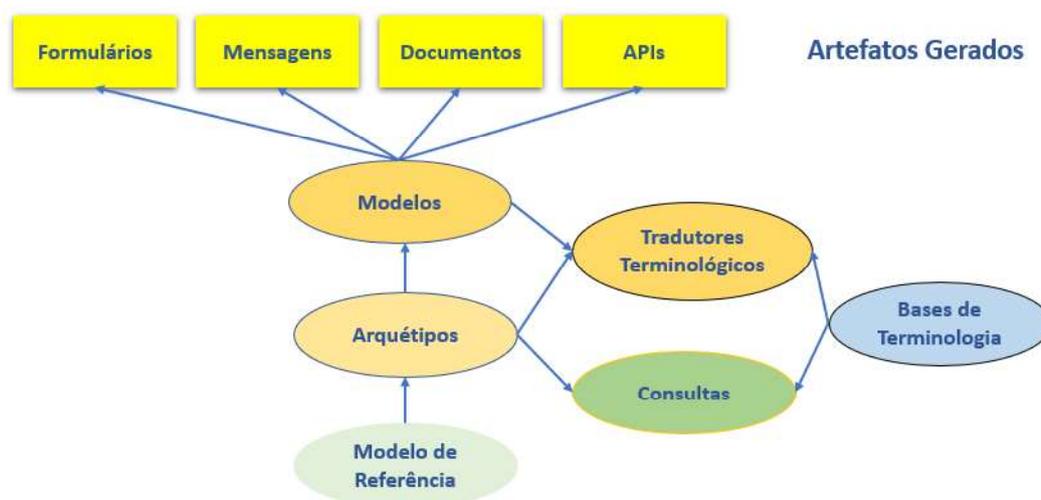
Como citado, esse autor entende que a ideia de desenvolvimento multinível apresenta-se como uma forma de desenvolvimento muito benéfica para o desenvolvimento de padrões de interoperabilidade de dados em saúde, sendo assim, essa concepção está sendo proposta para o desenvolvimento conceitual do OpenFairEHR.

Especificamente para o desenvolvimento da Camada Dados, propõe-se a utilização de arquétipos e modelos de dados disponíveis na página oficial do OpenEHR, além do desenvolvimento de outros registros seguindo a estrutura de desenvolvimento já testada e validada pelo padrão OpenEHR. De acordo com CKM (2022), atualmente existem disponíveis cerca de 1000 arquétipos e 207 modelos que podem ser utilizados para o desenvolvimento de projetos relacionados aos dados em saúde.

Considerando a estrutura definida na documentação do OpenEHR fica evidente a necessidade de um esforço para a construção de arquétipos e modelos. Destaca-se nesse ponto, de acordo com o objetivo dessa pesquisa, que consiste na criação de um modelo conceitual de interoperabilidade de dados de saúde alinhado com os princípios FAIR, a propositura da adoção da estrutura de dados do OpenEHR será validada apenas em um trabalho futuro de implementação desse modelo, aqui denominado de OpenFairEHR.

Espera-se obter com essa proposta uma estrutura de dados que seja compatível com a ilustrada na figura 19.

Figura 19 - Arquitetura Semântica do OpenEHR



Fonte: OpenEHR (2021, *online*, tradução nossa)

A estrutura ilustrada na figura 19 contempla o que se almeja alcançar, com ressalvas, na Camada Dados do OpenFairEHR, referente à semântica do modelo. Como pode ser visto a figura 19 existe uma estrutura base, aqui denominada de modelo de referência, que define as regras para o desenvolvimento da estrutura de arquétipos que serão utilizados para o desenvolvimento de modelos (templates).

Destaca-se ainda que os arquétipos e os modelos possuem uma ligação direta com uma estrutura de terminologia. Entende-se como sendo essencial essa conexão quando se almeja a interoperabilidade semântica, devido a heterogeneidade dos dados de saúde que alimentam uma estrutura EHR. Essa heterogeneidade justifica-se devido ao elevado número de padrões de dados em saúde existentes no mercado e utilizados por diversos aplicativos da área de saúde.

Outro elemento presente na figura 19 corresponde ao elemento *Queries*, presente nessa estrutura, pois permite que consultas realizadas entre arquétipos e elementos de terminologia sejam realizadas sem necessidade de verificação de estruturas de base de dados, ficando apenas a consulta em nível semântico.

Na figura 19, ainda é possível identificar que a estrutura de modelos tem ligação direta com elementos externos, a exemplos de esquemas de mensagens (XML, HTML, etc.), API's de comunicação interfaces de formulários, entre outros.

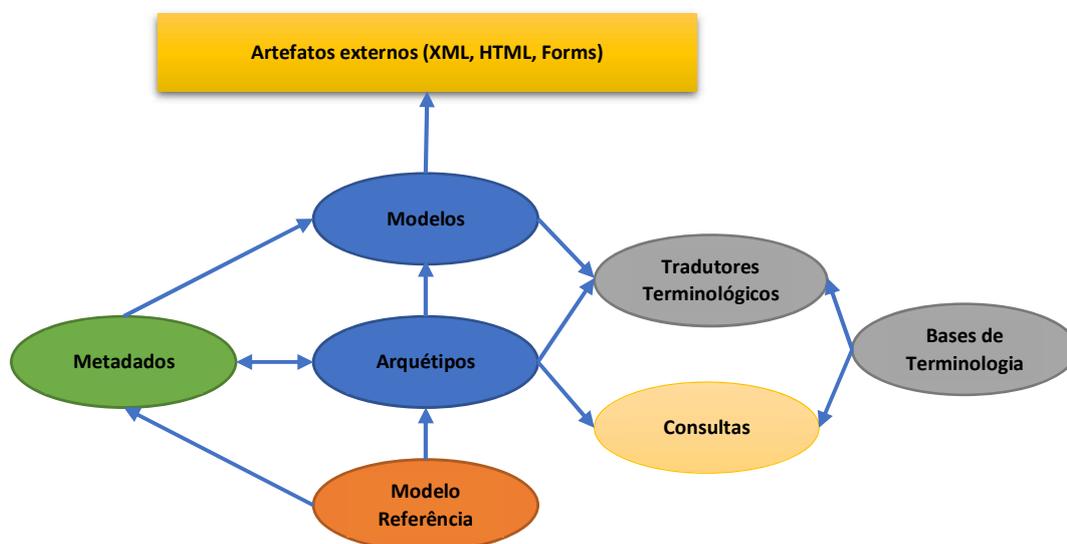
Como mencionado no início dessa discussão, a estrutura ilustrada na figura 19 corresponde ao que se almeja alcançar na concepção conceitual do modelo OpenFairEHR com ressalva ao elemento arquétipos dessa estrutura. A referida ressalva está presente na ausência

de uma estrutura de metadados independente, pois, como citado, para o OpenEHR os metadados estão na mesma estrutura dos dados, logo na mesma estrutura dos arquétipos.

No modelo conceitual proposto para o OpenFairEHR, a estrutura de metadados apresenta-se como sendo uma estrutura independente das camadas de dados e tecnologia, sendo requisito necessário para atividades de cadastro e recuperação dos dados armazenados.

Na figura 20, é possível identificar os elementos que compõem a arquitetura semântica do OpenFairEHR com base no padrão OpenEHR.

Figura 20 - Arquitetura Semântica do OpenFairEHR



Fonte: Elaborado pelo Autor (2022)

Analisando os elementos presentes na figura 20, percebe-se que o modelo de referência agora fornece subsídio para o desenvolvimento dos elementos arquétipos e metadados. Esses, por sua vez, mantêm uma comunicação constante, pois sua relação de dependência pode indicar a necessidade de alterações em sua estrutura com base em atualizações.

Fica evidenciado que a segmentação dos dados e metadados não implicam em prejuízo a ideia de separação da estrutura de aplicação com a estrutura de domínio. Dados e metadados fazem parte da estrutura de domínio, porém, o seu desenvolvimento deve ser realizado por profissionais com domínio em suas respectivas áreas.

Entende-se que o desenvolvimento dos elementos de metadados deverá ser realizado, ou pelo menos validado, por profissionais atuantes em áreas relacionadas à Ciência da Informação. No que se refere aos arquétipos e modelos, esses devem ser construídos por

profissionais da área assistencial relacionados à temática do dado, objeto desse arquétipo ou modelo.

Dando sequência na análise da figura 20, nessa proposta os modelos serão construídos com base nos arquétipos e nos metadados. Essa estrutura de dependência evidencia a obrigatoriedade dos metadados para concepção da estrutura dos dados do OpenFairEHR.

O desenvolvimento de elementos de metadados do OpenFairEHR foi discutido no tópico 5.4.2, sobre esse assunto destaca-se que a comunicação entre metadados e dados serão viabilizadas através do conceito de chaves primárias e estrangeiras.

Referente ao desenvolvimento de arquétipos e modelos para o OpenFairEHR, considerando que esse desenvolvimento ou mesmo a seleção de elementos existentes deverá ser realizado por profissional de amplo conhecimento no domínio de atuação. Por esse motivo não se pretende, nesse momento, avançar com esse desenvolvimento ou seleção de arquétipos e modelos. Fica essa atividade para o desenvolvimento de trabalhos futuros, a exemplo da implementação do modelo conceitual do OpenFairEHR.

Como ilustrado na figura 15, a estrutura do OpenFairEHR fundamenta-se nos elementos Metadados, Dados e Tecnologia. Já discutidas as camadas de Metadados e Dados, fica faltando apenas a Camada Tecnologia que será discutida a seguir.

5.4.4 Camada Tecnologia

Na arquitetura do OpenFairEHR, a Camada Tecnologia fica responsável por obter respostas para questões referentes à troca de dados entre sistemas de tecnologia da informação, bem como a disponibilização desses dados para acesso direto pelos usuários, além de questões relacionadas à segurança e privacidade dos dados.

Destaca-se nesse ponto que a referência a ser utilizada para o desenvolvimento da proposta conceitual da camada tecnologia do OpenFairEHR é o padrão HL7 FHIR. Essa escolha justifica-se pelo fato de o FHIR possuir uma documentação robusta e diversos *cases* que demonstram a efetividade de atuação nesse cenário.

Vê-se como relevante destacar que ao longo do desenvolvimento dessa pesquisa, mas especificamente na revisão da literatura sobre os padrões de interoperabilidade de dados em saúde FHIR e OpenEHR, ficou evidente que ambos possuem êxito em suas propostas de atuação, porém, eles foram concebidos com propósitos de atuação bem diversificados. O padrão OpenEHR busca trabalhar com interoperabilidade semântica, por esse motivo ele foi escolhido como sendo o padrão de referência para o desenvolvimento das camadas metadados e dados.

Em uma abordagem um pouco diferente o padrão FHIR teve sua concepção direcionada para uma interoperabilidade fundamental e estrutural, onde há o interesse em garantir a troca de informações entre sistemas, destaca-se em detrimento a interpretação desses dados como é almejado na interoperabilidade semântica. Não é correto afirmar que o FHIR não almeja a interoperabilidade a nível semântico, porém, é justo afirmar que esse objetivo é melhor executado pelo padrão OpenEHR.

Ainda sobre a Camada Tecnologia do OpenFairEHR, destaca-se que em um nível conceitual não está contemplado no escopo dessa pesquisa discutir questões de cunho técnico, bem como propor todo o conjunto de ferramenta ou tecnologias necessárias para sua implementação.

Diante do exposto, referente à Camada Tecnologia, será apresentado e discutido o que se pretende obter com sua implementação de acordo com a documentação do FHIR.

Referente ao processo de compartilhamento de dados entre sistemas de informação, o OpenFairEHR deverá contemplar o desenvolvimento de uma interface que possibilite gerenciar esse processo, desde a seleção dos dados até a configuração de itens de segurança, privacidade e auditoria.

A recomendação pelo desenvolvimento de gestão através de interface, em detrimento à utilização de linhas de comando, possibilita um ganho na usabilidade para usuários que não possuem perfis técnicos. Como mencionado, de acordo com HL7 (2022) essa prática de projeto com especificação de interface, é encorajada em projetos que trabalham com a estrutura do FHIR.

No que se refere à estrutura de comunicação, prioriza-se a utilização de recursos de TI caracterizados como softwares livres, disponíveis para utilização em qualquer finalidade e sem custo para os usuários.

Referente aos elementos de segurança dos dados armazenados e compartilhados, o modelo OpenFairEHR deverá disponibilizar recursos que possibilitem que todas as comunicações possam ser criptografadas para evitar acesso não autorizado, nenhum registro de vazamento de dados na ocorrência de erros na comunicação e possibilidade de monitoramento de tráfego com o objetivo de detectar padrões de acesso anômalos.

No que se refere a questões de privacidade, propõe-se que o modelo OpenFairEHR implemente o conceito de *Privacy by design* que, de acordo com Dallari e Monaco (2021), consiste em considerar a privacidade como elemento fundamental desde o início do processo de desenvolvimento dos artefatos.

Referente a questões de privacidade, almeja-se proporcionar a implementação de protocolos baseados em padrões que reflitam as preferências individuais dos titulares dos dados.

Ainda referente a essa questão da privacidade, a proposta de segmentação dos metadados, em que os metadados de identificação do paciente ficam separados dos demais, favorece a implementação de rotinas de anonimização e acesso controlado que potencializam as ações em busca da garantia da privacidade dos dados pessoais sensíveis.

Outro item proposto para o modelo OpenFairEHR, referente à privacidade, consiste na implementação de recursos que permitem a identificação de dados com base na sua sensibilidade ou confidencialidade (etiquetas de segurança).

Por fim, propõe-se a implementação de registros de acesso a dados e *logs* de auditoria com a finalidade de controle baseado em perfis com diferentes níveis de acesso.

Com o objetivo de ilustrar a estrutura do OpenFairEHR, a figura 21 apresenta as principais funcionalidades do modelo separadas por Camada. Destaca-se que as funcionalidades apresentadas não se isolam na camada em que se apresentam, em determinados momentos essa funcionalidade interage com elementos presentes nas outras camadas do OpenFairEHR.

Figura 21 - Funcionalidades do OpenFairEHR



Fonte: Elaborado pelo Autor (2022)

Analisando a figura 21, identificamos que na Camada Tecnologia agrupam-se as funcionalidades referente a: Regras de Negócio, Segurança, Controle de Acesso Conectividade e Auditoria.

No que se refere a regras de negócio, objetiva-se definir regras que definem formas de acesso, disponibilidade de conteúdo, definição e controle de perfis, entre outras atribuições. É na camada tecnologia que serão desenvolvidas interfaces para gestão de acesso ao conteúdo, bem como elementos de gestão de acessos realizados por outros sistemas de informação em saúde.

A funcionalidade denominada de Segurança refere-se à utilização de recursos de tecnologia que proporcionem o acesso seguro aos dados. Como citado, esse item não se restringe aos elementos da camada tecnologia, diz respeito também a regras aplicadas na base de dados que se encontram na Camada Dados.

O Controle de Acesso pode ser compreendido como sendo um item relacionado à segurança de acesso ao ambiente. Porém, ele também está vinculado conceitualmente com as funcionalidades de Auditoria e Privacidade. Objetiva-se com o controle de acesso proporcionar uma liberação de acesso a dados controlados e passíveis de auditoria.

A Conectividade pode ser compreendida como sendo a funcionalidade que possibilitará a comunicação entre OpenFairEHR e os recursos externos, sejam eles usuários ou sistemas de informação em saúde. Compete a Conectividade viabilizar a implementação das regras de negócios referente a acessos de elementos externos ao modelo OpenFairEHR através da definição de recursos de tecnologia como por exemplo interfaces para acessos por parte de usuários, ou estruturas em XML para acesso via sistemas de informação em saúde.

Por fim, na Camada Tecnologia consta a funcionalidade Auditoria. Essa tem por objetivo proporcionar visibilidade e rastreabilidade das atividades realizadas na estrutura do OpenFairEHR. A auditoria também pode ser compreendida como sendo um complemento das funcionalidades Segurança e Controle de Acesso.

Na Camada Metadados, constam as seguintes funcionalidades: Identificação dos Dados, Persistência de Acesso aos Dados, Conexão entre Tecnologia e Dados e Privacidade.

É oportuno destacar que a Camada Metadados possui uma relevância significativa para o OpenFairEHR no que se refere ao alinhamento desse modelo com os princípios FAIR. São nos metadados que se encontram vários elementos analisados pelo conjunto de princípios FAIR.

Referente às funcionalidades da Camada Metadados, iniciamos as análises pela funcionalidade Identificação dos dados. A identificação dos dados consiste em uma das formas de se conceituar os metadados, que podem ser compreendidos como sendo dados sobre dados. Almeja-se com a funcionalidade identificação dos dados que o conjunto padrão de metadados, definidos para o modelo OpenFairEHR, seja rico o suficiente para possibilitar a identificação dos dados de maneira clara e inequívoca.

Dando seguimento na análise, temos a funcionalidade Persistência de Acesso aos Dados. Essa funcionalidade se caracteriza como sendo a disponibilização de um recurso de localização dos metadados e dados por motores de busca externos (na internet) de forma persistente. Essa funcionalidade está presente no elemento “Localizador”, disponível em todos os conjuntos de metadados do OpenFairEHR.

A funcionalidade Conexão entre Tecnologia e Dados descreve a forma padrão de funcionamento do OpenFairEHR. Encontram-se na Camada Metadados os elementos necessários para interconexão entre todas as camadas do modelo. A conexão entre metadados e dados deverá ocorrer através das chaves primárias e estrangeiras formalmente criadas na estrutura de conexão da base de dados. No que se refere à conexão entre a camada metadados e tecnologia, mas especificamente com elementos externos ao modelo, essa conexão deverá ter viabilidade com auxílio do recurso “Localizador” que contém os dados do elemento persistente.

Finalizando as funcionalidades da Camada Metadados, temos a funcionalidade Privacidade. Justifica-se a inclusão dessa funcionalidade na camada metadados devido o projeto que prioriza a segmentação da estrutura do OpenFairEHR, isolando os metadados e dados de identificação do paciente e com isso possibilitando a aplicação de estratégias, ou regras de negócios, nessa estrutura com o objetivo de alcançar a proteção e privacidade desses dados categorizados como sendo dados pessoais sensíveis.

A funcionalidade de Privacidade está alinhada com as funcionalidades de Segurança, Controle de Acesso e Auditoria da Camada Tecnologia.

Referente às funcionalidades da Camada Dados, temos: Arquétipo de Dados, Modelo de Dados, Terminologia e Estrutura de Banco de Dados. Destaca-se que na Camada Dados serão definidas as regras e estruturas de base de dados, que estão diretamente relacionadas às regras de negócio e estruturas de metadados.

Iniciamos a análise das funcionalidades da camada dados com a funcionalidade Arquétipo de Dados. Como já mencionado no tópico referente à camada dados, a estratégia escolhida para o desenvolvimento da estrutura dos dados do modelo OpenFairEHR foi o desenvolvimento de arquétipos de dados, conforme realizado do padrão de interoperabilidade de dados em saúde o OpenEHR.

Essa escolha justifica-se na possibilidade de fomentar um desenvolvimento de arquétipos, que são estruturas padrões de dados reutilizáveis e adaptáveis, sem a necessidade de envolver simultaneamente profissionais das áreas assistências e de TI. Nesse modelo de arquitetura, essas duas equipes podem trabalhar de forma independente.

Uma inovação que o OpenFairEHR apresenta é a segmentação dos metadados, uma vez que nos padrões FHIR e OpenEHR os metadados estão na mesma estrutura dos dados. Essa segmentação favorece o desenvolvimento de estruturas de metadados ricas, pois podem ser desenvolvidas por profissionais das áreas relacionadas à ciência da informação, independente do desenvolvimento dos arquétipos e dos elementos de tecnologia.

Nesse cenário de desenvolvimento individualizado, é possível, no que se refere à manutenção da integridade dos dados, devido a utilização de conexão através de chaves primárias e estrangeiras.

Dando seguimento na análise com funcionalidade Modelo de Dados, destaca-se nesse ponto que o modelo consiste em uma estrutura que agrupa vários arquétipos, como também os elementos definidos na Camada Metadados.

Esse modelo, possivelmente será utilizado pela Camada Tecnologia como um formulário para coleta de dados ou mesmo com um XML, caso essa coleta seja realizada através da interoperabilidade desses dados com outros recursos de tecnologia da informação.

Os modelos também serviram como estruturas para a disponibilização de dados realizados através de consultas realizadas a base de dados OpenFairEHR. Destaca-se que nesse ponto as consultas que podem ser executadas de forma individualizada ou mesmo agrupada, nesse caso os modelos criados serão sugeridos como alternativa de agrupamento de dados.

A funcionalidade Terminologia tem como função básica servir de interpretador semântico entre os dados dos mais diversos padrões de dados em saúde atualmente em uso pelo mercado. Esse interpretador tem a funcionalidade de evitar conflitos de sinonímia (dados com mesmo conteúdo semântico, porém divergência em seus nomes) e homonímia (dados com nomes iguais e com conteúdo semânticos diferentes).

Finalizando a análise das funcionalidades da Camada Dados e dos elementos da figura 21, temos a funcionalidade Estrutura de Banco de Dados. Compete a essa funcionalidade a definição de elementos estruturais, de segurança e de controle de acesso referente à base de dados do OpenFairEHR.

Na funcionalidade de Estrutura de Banco de Dados é que se define a conectividade interna dos elementos do OpenFairEHR, bem como o tipo de base de dados a ser utilizada no desenvolvimento desse modelo. Destaca-se nesse ponto que a recomendação será a utilização de recurso disponíveis para uso sem ônus para o modelo, a exemplo do banco *MYSQL* ou *POSTGRESQL*.

Finalizando a discussão sobre a Camada Tecnologia, bem como as camadas Metadados e Dados, o tópico seguinte discorrerá sobre o alinhamento entre OpenFairEHR com os princípios FAIR, seguindo a mesma abordagem realizada para os padrões FHIR e OpenEHR.

5.4.5 Alinhamento entre o modelo proposto OpenFairEHR e os princípios FAIR

Seguindo a mesma abordagem utilizada para os padrões FHIR e OpenEHR, o nível de adaptabilidade do OpenFairEHR será analisado através das recomendações definidas em sua estrutura básica, desenvolvida ao longo dessa pesquisa, confrontadas com os questionamentos disponíveis na ferramenta FairDataBR, tendo seu resultado disponibilizado no anexo C dessa pesquisa. Espera-se ao final dessa análise identificar o nível de adaptabilidade do OpenFairEHR com os princípios FAIR.

Como início da análise de adaptabilidade do OpenFairEHR aos princípios FAIR, temos como sendo o primeiro conjunto de princípios o conjunto *Findable* (Localizável) com os seguintes questionamentos:

F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente?

F2. Os conjuntos de dados são descritos com metadados?

F3. O identificador está incluso em todos os registros / arquivos de metadados que descrevem os dados?

F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web?

F5. O conjunto de (meta)dados está publicado em um repositório?

No que se refere ao item F1, que busca analisar a existência de identificador único, global e persistente no conjunto de metadados, destaca-se nesse ponto que o modelo OpenFairEHR considerou como item essencial em sua estrutura de metadados o elemento **Localizador** que consiste em um elemento de metadados a ser utilizado para contemplar um identificador persistente na estrutura básica do OpenFairEHR.

Como proposta conceitual, sugere-se a utilização de um dado do tipo *Archival Resource Key* (ARK) que, segundo ARK (2022), classifica-se como sendo um identificador persistente multifuncional.

A escolha pelo tipo de identificador persistente ARK justifica-se por ser um identificador de ampla utilização, semelhante ao DOI, URNs e Handles, constituído há mais de 20 anos e atualmente existem aproximadamente 8,2 bilhões de ARKs (ARK, 2022).

Diante do exposto, considerando as opções de resposta disponíveis, a opção escolhida será: Identificador Persistente.

Para o item F2, que analisa a qualidade dos metadados disponíveis no modelo OpenFairEHR, vale ressaltar que o conjunto de metadados definido para o modelo, além de ser item obrigatório para o cadastro de novos dados, a estrutura de metadados foi idealizada tendo como referência o padrão Dublin Core, além da inclusão de outros elementos.

Diante do exposto, é possível afirmar que, em um estágio conceitual, os metadados estão estruturados e acredita-se que essa estrutura possibilitará, em um estágio futuro de implementação do modelo, o conjunto de metadados definido seja suficiente para viabilizar a identificação do dado sem a necessidade obrigatória do elemento identificador.

Nesse cenário, considerando as opções de resposta disponíveis na ferramenta FairDataBR, a opção indicada será a opção: Metadados Estruturados.

O item F3 busca analisar a presença obrigatória ou não dos elementos de metadados no cadastro do dado. Como citado na análise do item F2, a camada metadados foi desenvolvida segmentado da camada dados e tecnologia, porém, analisando a estrutura de chaves utilizadas na conexão desses elementos não se vê como possível a realização dessa comunicação sem a presença da estrutura metadados. Tal característica implica em afirmar que o item metadados é obrigatório para todo o dado inserido no OpenFairEHR.

Diante do exposto, considerando as opções de resposta, a alternativa selecionada será: Sim.

Dando seguimento às análises, o item F4 avalia a possibilidade de localização dos dados e seus metadados por mecanismos de busca na Web. Essa opção é atendida com êxito quando retomamos a discussão sobre o elemento Localização.

Destaca-se que esse elemento, além de ser persistente, ele pode executar funções similares a de uma URL, possibilitando com isso sua recuperação por pesquisa realizadas na Web. Como alternativa de resposta, a opção indicada será: Sim.

No item F5, que analisa questões referentes aos repositórios desses dados e metadados, será adotada a mesma lógica aplicada nas análises dos padrões FHIR e OpenEHR. Considerando que estamos analisando a adaptabilidade desses padrões aos princípios FAIR e não de estruturas de bases de dados, entende-se que essa questão não se aplica a esse contexto. Avaliando essa questão, será escolhida a opção: Em Nenhum Repositório.

Ao término da análise do item F5, conclui-se, portanto, a análise do conjunto de princípios FAIR *Findable* (Localizável) obtendo como resultado a pontuação máxima permitida na ferramenta FairDataBR: 7.60.

Dando continuidade, será iniciada a análise do conjunto de princípios “*Accessible*” (Acessível) do FAIR. Referente a esse conjunto, a ferramenta FairDataBR, disponibiliza os seguintes questionamentos:

A1. Quão acessíveis são os (meta)dados?

A2. Os (meta)dados estão acessíveis online sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?

A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?

A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?

A5. O protocolo (ex.: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?

A7. É possível realizar o download dos conjuntos de (meta)dados?

O item A1 busca analisar o nível de acessibilidade dos metadados disponíveis no modelo OpenFairEHR. É relevante citar que a estrutura do OpenFairEHR foi conceituada levando em consideração o acesso aos dados e metadados por um público diversos, composto, por exemplo, de estudantes, pesquisadores, governo (desenvolvimento de políticas públicas), bem como os profissionais da saúde. Porém, também foi previsto a proteção dos dados pessoais além de outros itens de segurança e auditoria.

Sendo assim, entende-se que o acesso aos metadados pode ser classificado como sendo Acessível ao Público, sendo que preservadas regras de privacidade, proteção e segurança desses dados, em especial por serem classificados como sendo dados sensíveis.

O item A2, que contém o seguinte questionamento: “Os (meta)dados estão acessíveis online sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?”. Como alternativa de resposta a opção indicada será: API de Web Service Padrão.

Essa resposta, justifica-se no fato do modelo OpenFairEHR ter sido conceituado utilizando padrões e ferramentas de TI de livre acesso e com grande referência de uso pelos usuários, a exemplo do XML e HTML. Além disso, destaca-se também o conceito de

desenvolvimento de interfaces para gestão desses dados, além do uso de modelos (formulários) que auxiliam nas atividades de inserção e recuperação desses dados e metadados.

No item A3, que busca analisar a possibilidade de acesso aos metadados através do identificador fornecido, destaca-se que no modelo OpenFairEHR, além do elemento **Localizador**, existem as chaves primárias e estrangeiras que possibilitam o acesso à base de dados através de consultas realizadas por APIs ou usuários. Diante do exposto, e considerando as opções de respostas disponíveis na ferramenta FairDataBR, a opção indicada será a opção **Sim**.

O item seguinte a ser analisado consiste no item A4, que traz o seguinte questionamento: Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis? Como alternativa de resposta a opção **sim** foi a escolhida.

Considerando que à arquitetura do modelo OpenFairEHR, como já citado em vários momentos, separa os metadados dos dados e da tecnologia através das camadas metadados, dados e tecnologia. Essa arquitetura garante a continuidade do acesso aos metadados mesmo que os dados não estejam mais disponíveis.

Os itens A5 e A7 que, respectivamente, analisam se os protocolos utilizados para criação do modelo OpenFairEHR são abertos, gratuitos e universalmente implementável e se existe a possibilidade de realizar *download* do conjunto de metadados possuem como resposta a opção **sim** para ambos.

Referente às indagações sobre os protocolos utilizados, item A5, destaca-se que no OpenFairEHR foram contemplados protocolos abertos, gratuitos e amplamente utilizados a exemplo dos protocolos XML e HTTP.

No que se refere aos questionamentos do item A7, que avalia a possibilidade de *download* do conjunto de metadados, a opção escolhida fundamenta-se no fato de não existir nenhum impedimento para realização de downloads. Tal afirmação ganha força quando consideramos a adoção de um padrão de banco de dados de amplo conhecimento a exemplo do MYSQL.

Com isso chegamos ao final da análise do conjunto de princípios “*Accessible*” (Acessível) do FAIR. Considerando a metodologia adotada pelo FairDataBR, o resultado identificado para esse conjunto de princípios foi a nota máximo disponível na ferramenta 10,00.

O próximo conjunto a ser analisado corresponde ao princípio da *Interoperable* (interoperabilidade), e para essa análise a ferramenta FairDataBR apresenta as seguintes indagações:

11. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?
12. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?
13. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores?
14. Para representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico?

No que se refere ao conjunto de princípios da interoperabilidade do FAIR, vê-se como oportuno revisitar o objetivo geral dessa pesquisa que apresenta a seguinte redação: “Analisar a interoperabilidade de dados de saúde de acordo com os princípios FAIR”. Diante desse objetivo fica evidente que na definição conceitual do modelo OpenFairEHR foi dedicada uma atenção especial a suprir lacunas identificadas nos padrões FHIR e OpenEHR referentes às questões de interoperabilidade.

Ainda referente aos princípios da interoperabilidade, dando seguimento na análise de adaptabilidade do OpenFairEHR com os princípios FAIR, o item I1 da ferramenta FairDataBR busca investigar se os metadados estão disponíveis em formatos preferenciais.

Entende-se que esse quesito está contemplado na estrutura do OpenFairEHR, uma vez que foi descrito na definição da camada tecnologia a escolha pela utilização de padrões de comunicação amplamente conhecidos e utilizados por diversos sistemas de informação em saúde, a exemplo do padrão XML.

Diante do exposto, considerando as opções de respostas disponíveis, a alternativa escolhida foi a resposta: Sim.

Para o item I2, que indaga a estrutura de metadados sobre a utilização ao não de esquemas ou modelos de metadados aprovados pela comunidade, destaca-se nesse ponto que na concepção conceitual do OpenFairEHR, mas especificamente sobre a definição da camada metadados, foi utilizado como referência o padrão de metadados *Dublin Core* com algumas adaptações necessárias para uma melhor utilização desses metadados.

Considerando as opções de respostas disponíveis na ferramenta FairDataBR, a alternativa escolhida foi: “Esquema de Comunidade Padrão”.

Dando sequência nessa análise, o item I3 busca investigar se os metadados estão vinculados a outros metadados usando identificadores. Referente a esse quesito, na definição da camada metadados, foi definida a estratégia de conectividade entre os metadados e entre metadados e dados através da utilização de chaves primárias e chaves estrangeiras. Destaca-se

também a definição do elemento localizador que contempla a presença de um registro persistente que viabiliza a busca externa dos metadados e dados.

Diante das opções de resposta disponíveis na ferramenta, a opção escolhida para o item I3 foi: “Links Internos/Externos”.

Finalizando a análise do conjunto de princípios da interoperabilidade, o item I4 analisa se para a representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico.

Seguindo a mesma lógica definida no padrão OpenEHR, foi contemplado na concepção conceitual do OpenFairEHR a utilização de uma estrutura de terminologia que busca proporcionar o alinhamento semântico através do mapeamento dessa estrutura de terminologia com outras terminologias ou vocabulários existentes, a exemplo do CID, DeCS, ISO 3166-1, entre outros.

É oportuno destacar que a composição dos dados necessários para a construção de um RES é formada por dados em saúde que utilizam padrões diversificados, sendo necessária a utilização de recursos de terminologia para garantir a interpretação correta do conteúdo disponível nos dados compartilhados ao longo do processo de interoperabilidade entre sistemas ou mesmo pela operacionalização desses dados por usuários.

Diante do exposto, considerando que as estruturas de terminologias e vocabulários contemplados são de amplo reconhecimento e utilização global, como opção de resposta, a alternativa escolhida foi: “Uso de instrumentos de controle terminológico COM identificador global”.

Ao final da análise do conjunto de princípios de interoperabilidade disponível na ferramenta FairDataBR, considerando as opções de respostas aqui justificadas, a nota obtida nesse conjunto de princípios foi a nota máxima: 10,00.

Para finalizar a análise de adaptabilidade do padrão OpenFairEHR com os princípios FAIR, serão analisados os itens correspondentes ao conjunto de princípios *Reusable* (Reusabilidade) do FAIR.

Para esse conjunto de princípios, a ferramenta FairDataBR disponibiliza cinco itens de investigação, sendo eles:

R1. Os conjuntos de (meta)dados estão licenciados?

R2. Qual a licença de uso dos conjuntos de (meta)dados?

R4. Os conjuntos de (meta)dados possuem proveniência detalhada?

R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados?

R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio?

Iniciando a análise desses princípios, o item R1 avalia a existência de licença de uso dos metadados. Nesse contexto, é oportuno destacar que na camada metadados do OpenFairEHR está previsto a utilização de elementos que abordem questões relacionadas aos direitos autorais, tipo e localização da licença de uso e questões relacionadas à LGPD. Esses itens estão respectivamente contemplados nos elementos Direitos Autorais, Licença e Finalidade contemplados na camada metadados do OpenFairEHR.

Diante do exposto, considerando as opções existentes, entende-se que a melhor alternativa de resposta será a opção Sim.

O item R2 segue na mesma linha de investigação que o item R1, porém, conforme foi detectado na análise do nível de adaptabilidade dos padrões FHIR e OpenEHR com os princípios FAIR, a análise desse item aplica-se melhor para bases de dados e não padrões de interoperabilidade. Diante dessa afirmação, a análise desse item torna-se inconclusiva.

Destaca-se apenas previsibilidade de registro de elementos de metadados relacionados à licença de uso dos conjuntos de dados, através do elemento Licença disponível na camada metadados do OpenFairEHR. Referente ao tipo de licença utilizado, fica essa resposta para uma etapa de implementação e implantação do modelo aqui definido de forma conceitual.

Da mesma forma como foi feita na análise desse item para os padrões FHIR e OpenEHR, considerando as opções de resposta disponíveis na ferramenta FairDataBR, a opção escolhida será Aberta.

Dando seguimento com a análise, entende-se como opção de resposta para o item R4, uma vez que esse item investiga a existência de controle de proveniência detalhada, a opção de resposta mais adequada será a alternativa Sim.

Tal resposta fundamenta-se na existência dos elementos de metadados Autor, Data, ciclo de vida, que viabiliza um controle de versão dos dados no OpenFairEHR possibilitando, entre outros benefícios, a realização de auditoria de possíveis alterações de dados.

Ainda referente à análise do item R4, destacam-se os itens de segurança de acesso previstos na concepção da camada tecnologia que corroboram com a ideia de proveniência dos dados analisadas nesse ponto.

O item R5 pode ser compreendido como sendo um complemento ao item R4, uma vez que o questionamento disponível nesse item, faz referência a existência, ou não, de uma descrição da origem do fluxo de trabalho que produziu os metadados.

Considerando os elementos destacados como itens de rastreabilidade dos dados, bem como a obrigatoriedade de preenchimento dos metadados para a inclusão de novos dados na estrutura do OpenFairEHR, além da existência de um elemento Relação que possibilita a vinculação desse dado ou metadados com outros dados ou metadados e possível afirmar a existência de rastreabilidade dos dados e metadados no modelo OpenFairEHR. Diante do exposto, a opção de resposta escolhida foi: Rastreamento de Proveniência.

Finalizando a análise dos princípios de reusabilidade, bem como de todos os itens disponíveis na ferramenta FairDataBR, o item R6 investiga se a estrutura de metadados está de acordo com padrões relevantes para o domínio.

Para essa análise é oportuno citar que desde a concepção inicial do modelo OpenFairEHR foi priorizado a utilização de padrões de metadados e de interoperabilidade de dados em saúde, bem como ferramentas de tecnologia, que tenham licença aberta e que sejam referência em seus domínios de utilização.

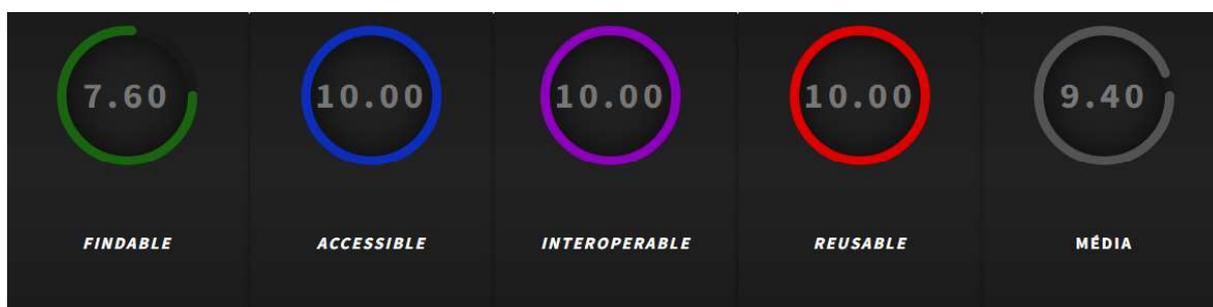
Seguindo nessa linha, o modelo OpenFairEHR teve como referência o padrão de metadados Dublin Core e os padrões de interoperabilidade de dados em saúde FHIR e OpenEHR. No que se refere a itens de tecnologia, o OpenFairEHR faz uso de recursos como XML, HTML e APIs disponíveis para viabilização dessa troca de dados entre sistemas e usuários.

Sendo assim, considerando as opções de resposta disponíveis, entende-se que a melhor opção de resposta para o item R6 será a alternativa Padrões de Domínio.

Finalizando o conjunto de itens referente aos princípios da reusabilidade, da mesma forma que foi realizado com os demais conjuntos de princípios, a ferramenta FairDataBR, destaca que o nível de adaptabilidade do modelo OpenFairEHR com os princípios do conjunto reusabilidade do FAIR possui uma nota 10,00.

Finalizando a análise do modelo OpenFairEHR com os princípios FAIR, ilustrado na figura 22, de acordo com a ferramenta FairDataBR, ficou evidente que esse modelo buscou em sua concepção conceitual fazer uso das melhores práticas disponíveis nos padrões FHIR e OpenEHR, bem como, propor modificações objetivando melhorias nos pontos que esses padrões não estão alinhados com os itens da ferramenta FairDataBR.

Figura 22 - Ferramenta FairDataBR: Nível de adaptabilidade do OpenFairEHR aos princípios FAIR



Fonte: FairDataBR (2021)

Com o objetivo de facilitar a compreensão da análise de adaptabilidade do modelo OpenFairEHR com os princípios FAIR, utilizando a ferramenta FairDataBR, no quadro 4 encontra-se, de forma agrupada, as opções de resposta selecionada para cada item do FairDataBR, separado pelos padrões FHIR, OpenEHR e o modelo OpenFairEHR.

A criação desse quadro contempla uma distribuição de elementos através de cores que representam o nível de adaptabilidade do padrão ou modelo com o respectivo item disponível no FairDataBR.

Opções em cor verde significa que o padrão ou modelo estão plenamente de acordo com o respectivo item disponível no FairDataBR. Para os itens na cor amarela, implica afirmar que o padrão está próximo de alcançar a adaptabilidade desejada na análise da ferramenta e para itens na cor vermelha vê-se como necessário a realização de ajustes maiores para alcançar o nível de adaptabilidade desejado.

Quadro 5 - Análise de adaptabilidade entre os padrões FHIR, OpenEHR e o modelo OpenFair

Ferramenta FairDataBr		Padrões de interoperabilidade saúde	
Conjunto de Princípios FAIR	Itens Analisados	FHIR	C
<i>FINDABLE</i>	F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente?	Endereço Web	Ende
	F2. Os conjuntos de dados são descritos com metadados?	Metadados Simples	Me Si
	F3. O identificador está incluso em todos os registros / arquivos de metadados que descrevem os dados?	Sim	
	F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web?	Não	
	F5. O conjunto de (meta)dados está publicado em um repositório?	Em nenhum repositório	Em rep
<i>ACCESSIBLE</i>	A1. Quão acessíveis são os (meta)dados?	Acessível ao Público	Ace P
	A2. Os (meta)dados estão acessíveis online sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?	API de Web Service padrão	Web S p
	A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?	Sim	
	A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?	Não	
	A5. O protocolo (ex.: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?	Sim	
	A7. É possível realizar o download dos conjuntos de (meta)dados?	Sim	
<i>INTEROPERABLE</i>	I1. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?	Sim	

	I2. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?	Esquema não formalizado		Es Ex
	I3. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores?	Sem links		I Intern
	I4. Para representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico?	Nenhum recurso/instrumento		U instru co term C iden g
REUSABLE	R1. Os conjuntos de (meta)dados estão licenciados?	Não		
	R2. Qual a licença de uso dos conjuntos de (meta)dados?	Aberta		A
	R4. Os conjuntos de (meta)dados possuem proveniência detalhada?	Sim		
	R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados?	Rastreamento de Proveniência		Rastre Prov
	R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio?	Padrões de Domínio		Pad Do

Fonte: Elaborado pelo Autor (2022)

Legenda:

COR	SIGNIFICADO
	Plenamente de acordo com o respectivo item disponível no FairDataBR.
	Próximo de alcançar a adaptabilidade desejada.
	Necessário a realização de ajustes maiores para alcançar o nível de adaptabilidade desejado.

Ainda referente à análise das informações disponíveis no quadro 4, é possível perceber que o padrão FHIR não atendendo plenamente 8 dos 20 itens que compõe todos os conjuntos de princípios FAIR de acordo com o FairDataBR. Já o padrão OpenEHR não atende plenamente 6 dos 20 itens do FairDataBR.

Itens que não foram atendidos por nenhum dos padrões de referência (FHIR e OpenEHR) totalizaram 4 itens, os demais ou foram atendidos pelo FHIR ou pelo OpenEHR ou até mesmo pelos dois.

Destacam-se que os itens do FairDataBR que foram atendidos pelos padrões FHIR e OpenEHR tiveram suas recomendações adotadas, sempre que possível, na íntegra pelo modelo OpenFairEHR. Itens que foram atendidos apenas por um dos padrões tiveram suas recomendações analisadas e adaptadas, sempre que necessário, para o modelo OpenFairEHR.

Para os itens não contemplados por nenhum dos padrões, foi necessário realizar uma definição conceitual completa para alcançar o nível de adaptabilidade desejado.

Nesse sentido, a nível de ampliar ainda mais a compreensão do processo de definição conceitual do modelo OpenFairEHR, o quadro 5 fará um recorte nos dados do quadro 4 contemplando apenas os itens que não foram atendidos por nenhum dos padrões e apresentada a estratégia adotada no modelo OpenFairEHR.

Quadro 6 - Ações desenvolvidas no modelo OpenFairEHR para alcançar o nível de adaptabilidade recomendado pelo FairDataBR

Conjunto de Princípios FAIR	FINDABLE
Itens Analisados	F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente?
Estratégia adotada para atender a demanda do FairDataBR	Foi contemplado na camada metadados o elemento DC.identificador.localizador que se refere a um <i>Archival Resource Key</i> (ARK), identificador persistente multifuncional. Destaque-se que esse tipo de elemento não foi localizado em nenhum padrão de interoperabilidade estudado, pelo menos quando analisamos sua estrutura padrão.
Conjunto de Princípios FAIR	FINDABLE
Itens Analisados	F2. Os conjuntos de dados são descritos com metadados?
Estratégia adotada para atender a demanda do FairDataBR	Ao contrário do que foi identificado nos padrões FHIR e OpenEHR, onde a descrição dos metadados é opcional, no modelo OpenFairEHR essa descrição é obrigatória e requisito essencial para o cadastro de novos dados. Outro ponto relevante para obter êxito nesse item consiste na qualidade dos elementos definidos como metadados que teve como forte influência o padrão Dublin Core. A opção pela indicação de metadados estruturados levou em consideração apenas o fato de se tratar de uma proposta conceitual, porém, acredita-se que na implementação desse modelo a

	compreensão dessa alternativa será revista e classificada como metadados ricos.
Conjunto de Princípios FAIR	ACCESSIBLE
Itens Analisados	A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?
Estratégia adotada para atender a demanda do FairDataBR	Analisando a estrutura de metadados dos padrões FHIR e OpenEHR foi identificado que nesses padrões dados e metadados dividem a mesma estrutura, essa arquitetura impossibilita a recuperação dos metadados na impossibilidade de acesso aos dados. No modelo OpenFairEHR dados e metadados coexistem de forma dependente, porém, mantidos em estruturas de domínio separadas. Dessa forma a recuperação dos metadados pode ocorrer mesmo com a indisponibilidade dos dados.
Conjunto de Princípios FAIR	INTEROPERABLE
Itens Analisados	I2. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?
Estratégia adotada para atender a demanda do FairDataBR	Nesse quesito foi adotado, como referência, o padrão de metadados com aceitação nacional e internacional <i>Dublin Core</i> para o desenvolvimento da estrutura da camada metadados.

Fonte: Elaborado pelo Autor (2022)

Como pode ser visto no quadro 5, a inclusão de um elemento de metadados com uma estrutura persistente, bem como a adoção de um modelo de referência para o desenvolvimento da camada metadados, além da opção de desenvolvimento de estruturas separadas para dados e metadados possibilitou implantar melhorias no modelo OpenFairEHR que até então não eram objeto de interesse dos padrões FHIR e OpenEHR.

Diante do exposto até o presente momento, vê-se como relevante mencionar que todas as definições apontadas por essa pesquisa para o desenvolvimento do modelo OpenFairEHR tiveram como base elementos teóricos obtidos através da revisão bibliográfica aqui presente e foram adaptadas para proporcionar uma construção conceitual de um modelo de interoperabilidade de dados em saúde alinhado com os princípios FAIR.

A validação e refinamento desse modelo será fruto de trabalhos futuros, de interesse desse autor, que objetivem a construção do modelo OpenFairEHR, bem como a aplicação prática desse modelo para testes de sua eficiência e eficácia.

Encerra-se aqui o tópico que abordou o nível de adaptabilidade do OpenFairEHR com os Princípios FAIR, através da análise com auxílio da ferramenta FairDataBR constatando que o modelo proposto alcançou nota máxima em todos os conjuntos de princípios analisados.

No tópico seguinte, serão apresentadas as considerações finais dessa pesquisa, destacando o trabalho realizado e os objetivos alcançados.

6 CONSIDERAÇÕES FINAIS

Os dados e as informações, como objetos centrais desta pesquisa e objetos de pesquisa da área de Ciência da Informação, exercem grande influência no cotidiano da sociedade contemporânea. Essa influência torna-se ainda mais evidente quando atrelamos a sua produção, uso e recuperação a necessidade de utilização de recursos de tecnologia.

Os avanços da tecnologia trazem para a sociedade inúmeros benefícios e desafios, os dados e as informações passam a ser o cerne de toda sociedade e seu consumo passa a ser uma prioridade. O consumo da informação deixa de ser apenas um desejo da sociedade contemporânea e passa a ser uma necessidade, e com isso o volume de dados e informações produzidos e consumidos, em especial, através dos meios digitais cresce exponencialmente.

O cenário aqui descrito evidencia-se em diversos contextos da vida cotidiana, e na área da saúde não seria diferente, em especial pela influência da Pandemia da COVID-19, iniciada em 11 de março de 2020, que obrigou a todos a evitarem o contato presencial e com isso potencializando a utilização de meios digitais para viabilização desse contato social.

O aumento significativo no volume de dados de saúde produzidos e consumidos, durante o período na Pandemia, ressaltou a necessidade de desenvolvimento de pesquisas com objetivo de buscar um melhor uso desses dados. Seguindo essa linha, essa pesquisa, que tem como foco central conceber um modelo conceitual de interoperabilidade em dados em saúde com base nos princípios FAIR, que pode ser vista como sendo uma contribuição relevante para a sociedade e para academia.

A referida pesquisa buscou investigar abstrações relacionadas à interoperabilidade de dados em saúde, por meio da elaboração de um modelo conceitual de interoperabilidade, desenvolvido como fruto da análise de modelos disponíveis relacionados ao tema, além de conceitos sobre metadados e princípios FAIR.

O desenvolvimento de uma pesquisa multidisciplinar que demanda uma imersão em áreas como a Ciência da Informação, Ciência da Computação e Saúde, apresenta-se como sendo desafiador devido a sua heterogeneidade conceitual. A necessidade de entrelaçar essas áreas na

busca pelo desenvolvimento de um modelo conceitual de interoperabilidade de dados em saúde, alinhado com os princípios FAIR, evidenciou a complexidade existente em estudos dessa natureza.

O percurso utilizado para o desenvolvimento dessa pesquisa teve sua origem na identificação de uma lacuna existente na melhor utilização desses dados que, por vezes, não são concebidos com ênfase no compartilhamento automatizado de forma eficiente possibilitando sua recuperação em tempo hábil. Essa lacuna ganhou evidência no período da Pandemia, momento em que os dados em saúde foram produzidos em larga escala no formato digital e seu compartilhamento não realizado de forma automatizada. Esse não compartilhamento impactou diretamente a recuperação da informação e consequentemente o desenvolvimento de pesquisas científicas, que necessitam de acesso ao maior volume possível de dados para proporcionar resultados mais consistentes e confiáveis.

Entendendo e aceitando que a alternativa mais eficiente para proporcionar o compartilhamento de dados de saúde entre sistemas seria através da interoperabilidade, e considerando os benefícios que os princípios FAIR proporcionam para desenvolvimento de bases de dados de pesquisa, essa pesquisa debruçou-se sobre um vasto conteúdo conceitual para alcançar a tese de que a aplicação dos princípios FAIR no contexto dos dados em saúde contribui para a interoperabilidade de dados em saúde.

No desejo de alcançar a tese mencionada algumas hipóteses foram estabelecidas, nesse sentido, pode-se afirmar que essas hipóteses foram parcialmente ratificadas. No que se refere à aplicação do subconjunto de interoperabilidade dos princípios FAIR, como alternativa para proporcionar a interoperabilidade de dados em saúde, evidenciou-se que essa alternativa não é válida devido a interdependência existente entre os princípios FAIR pertencentes a subconjuntos distintos.

Essa constatação fundamenta-se quando analisamos os resultados da análise da ferramenta FairDataBR, que foi a ferramenta adotada para avaliar o nível de adaptabilidade dos padrões em modelos de interoperabilidade aqui estudados, em consonância com os princípios FAIR. O primeiro subconjunto analisado pela ferramenta corresponde ao subconjunto de princípios *Findable* (localizável) e as respostas disponibilizadas para os questionamentos desse subconjunto exercem influência direta em questões analisadas em outros subconjuntos, a exemplo do subconjunto da interoperabilidade.

Além do exposto, através das análises realizadas nos padrões de interoperabilidade estudados, foi possível identificar que os elementos necessários para o desenvolvimento de um modelo de interoperabilidade vão além, pelo menos em parte, dos elementos contemplados nos

princípios do subconjunto de interoperabilidade do FAIR. Como forma de exemplificação destacam-se elementos relativos à segurança, conectividade, privacidade e controle de acesso dos dados e seus metadados.

Retomando a questão das hipóteses estabelecidas, no que se refere à questão de identificação de um melhor resultado para o alinhamento de um modelo de interoperabilidade de dados em saúde com os princípios FAIR, através da união de aspectos de interoperabilidade de dados em saúde dos padrões FHIR e o padrão OpenEHR, essa foi ratificada.

Considerando as estruturas nativas desses padrões, foi possível identificar pontos de destaque em cada um deles, de forma individualizada, como também de forma conjunta. Essa constatação serviu como referência básica para o desenvolvimento do modelo de interoperabilidade proposto nessa pesquisa e que foi nomeado de OpenFairEHR.

Destaca-se nesse quesito que também foram identificados elementos analisados pelo FairDataBR que não estavam contemplados em nenhum dos modelos analisados, sendo necessário, para contemplar essas lacunas, o desenvolvimento de inovações na estrutura do OpenFairEHR.

Destaca-se também que, ao longo do desenvolvimento dessa pesquisa, os objetivos específicos foram sendo alcançados. Sendo os dois primeiros, identificar padrões de interoperabilidade utilizadas para dados em saúde e investigar os princípios FAIR, com ênfase no subconjunto de princípios de interoperabilidade, contemplados na etapa de identificação de padrões utilizados para dados em saúde, seleção dos padrões FHIR e OpenEHR como sendo padrões de referência em interoperabilidade de dados em saúde e na pesquisa sobre os princípios FAIR, mas especificamente o subconjunto de princípios de interoperabilidade.

O terceiro e o quarto objetivos específicos, sendo eles, respectivamente, verificar o nível de adaptabilidade dos padrões de interoperabilidade identificados com base nos princípios FAIR e avaliar possibilidades de melhorias dos padrões de interoperabilidade identificados, foram contemplados com o desenvolvimento do modelo de interoperabilidade de dados em saúde denominado de OpenFairEHR.

A partir dessa pesquisa conceitual inédita, podem surgir novas possibilidades de trabalhos futuros. Dentre essas oportunidades destaca-se a necessidade de desenvolvimento e implantação do OpenFairEHR, aqui proposto de forma conceitual.

Como limitação da pesquisa realizada, ressalta-se a existência de uma lacuna a ser considerada entre o desenvolvimento conceitual de um modelo de interoperabilidade de dados em saúde e sua concepção prática, e que ajustes podem ser necessários para que se consiga materializar o que foi proposto no modelo conceitual.

Destacam-se como desafios entre o modelo conceitual e sua implementação, os avanços tecnológicos presentes tanto nas ferramentas necessárias para o desenvolvimento do modelo quanto nas estruturas de dados de saúde utilizadas pelos mais diversos sistemas de informação em saúde existentes e que ainda vão surgir.

Propõe-se como trabalho futuro a construção de uma ferramenta, com a consequente implementação do modelo OpenFairEHR, e com isso validar a eficácia do modelo, hoje conceitual, mas que poderá vir a se tornar um padrão de referência.

REFERÊNCIAS

AKHLAQ A, SHEIKH A, PAGLIARI C. Defining health information exchange: scoping review of published definitions. **J Innov Health Inform.** 2016. Disponível em: <https://hijournal.bcs.org/index.php/jhi/article/view/838/960>. Acesso em: 17 jun. 2021

ALVES, R. C. V. **Metadados como elementos do processo de catalogação.** 2010. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

ARK. ARK Alliance, 2021. Disponível em: <https://arks.org/>. Acesso em: 22 jun. 2022

ARK. ARK Alliance, 2022. Disponível em: <https://arks.org/about/>. Acesso em: 26 jun. 2022

ATLASSIAN. FHIR identifiers and FAIR principles on IDs. **Confluence-hl7.** 2021. Disponível em: https://confluence-hl7-org.translate.goog/display/SOA/FHIR+identifiers+and+FAIR+principles+on+IDs?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt-BR&_x_tr_pto=op,sc. Acesso em: 10 de jun. 2022.

BADER, B.S. CQI progress reports: The dashboard approach provides a better way to keep board informed about quality. **Healthcare executive**, Chicago, v. 8, n. 5, p. 8-11, Set 1993.

BALE, T. **What is an archetype?** 2015. Disponível em: <https://openehr.atlassian.net/wiki/spaces/resources/pages/4554753/Archetype+FAQs> Acesso em: 23 de set. 2022.

BARBEDO, F; CORUJO, L; SANT'ANA, M. Recomendações para a produção de planos de preservação digital. Lisboa: **Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB)**, nov. 2011. 111 p. Disponível em: http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2014/02/Recomend_producao_PPD_V2.1.pdf. Acesso em: 9 jul. 2021.

BARBEDO, F. et al. RODA: Repositório de objectos digitais autênticos. **Actas dos Congressos Nacionais de Bibliotecários, Arquivistas e Documentalistas**, Lisboa, n. 9, 2007. Disponível em: <http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/535/320>. Acesso em: 9 jul. 2021.

BEALE, T. **OpenEHR International Board**, Ars Semântica, Reino Unido. Disponível em: <https://specifications.openehr.org/releases/AM/latest/Overview.html>. Acesso em: 21 jul. 2022.

BECKER, H. S. **Métodos de pesquisa em ciências sociais.** 4ª ed. São Paulo: HUCITEC; 1999.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil.** Brasília, DF: Presidência da República, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 23 mar. 2022.

BRASIL. **LEI Nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm Acesso em: 02 abr. 2022.

BRASIL. **Política Nacional de Informação e Informática em Saúde - PNIIS** / Ministério da Saúde, Secretaria-Executiva, Departamento de Monitoramento e Avaliação do SUS. – Brasília: Ministério da Saúde, 2016.

BRASIL. Ministério da Saúde. Secretaria-Executiva. Departamento de Monitoramento e Avaliação do SUS. **Política Nacional de Informação e Informática em Saúde - PNIIS** / Ministério da Saúde, Secretaria-Executiva, Departamento de Monitoramento e Avaliação do SUS. – Brasília: Ministério da Saúde, 2016.

BRASIL. **Portaria nº 2.073**, de 31 de agosto de 2011. Regulamenta o uso de padrões de informação em saúde e de interoperabilidade entre os sistemas de informação do SUS, nos níveis Municipal, Distrital, Estadual e Federal, e para os sistemas privados e de saúde suplementar. Publicada no Diário Oficial da União em 31 de agosto de 2011. Disponível em: https://bvsmms.saude.gov.br/bvs/saudelegis/gm/2011/prt2073_31_08_2011.html. Acesso em: 15 abr. 2021.

BRASIL. **Lei 8080 de 19 de setembro de 1990**, dispõe sobre a organização do SUS. Publicada no Diário Oficial da União em 19 de setembro de 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8080.htm. Acesso em: 8 abr. 2021.

BRASIL. **Lei 8.159, de 8 de janeiro de 1991.** Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Publicada no Diário Oficial da União em 8 de janeiro de 1991. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8159.htm. Acesso em: 8 abr. 2021.

CASTELLS, M. **A sociedade em rede.** São Paulo: Paz e Terra, 2013

CKM OpenEHR. 2022 Disponível em: <https://ckm.openehr.org/ckm/>. Acesso em: 22 de jul. 2022.

CHAFFEY, D; WOOD, S. 2005. **Business Information Management: Improving Performance Using Information Systems**, Harlow, FT Prentice Hall. 2005.

CHAN, L. M.; ZENG, M. L. Metadata interoperability and standardization: a study of methodology part i: achieving interoperability at the schema level. **D-Lib Magazine**, v. 12, n. 6, jun 2006. Disponível em: <http://www.dlib.org/dlib/june06/chan/06chan.html>. Acesso em: 9 jan. 2022.

DALLARI, B. A; MONACO, C. G. F; **LGPD na Saúde.** Thomson Reuters 1 ed. São Paulo, 2021.

DATA FAIRPORT, 2014. Disponível em: <https://www.datafairport.org/> Acesso em: 02 fev. 2022.

DUBLIN CORE METADATA INITIATIVE (DCMI). **Dublin Core™ Metadata Element Set, Version 1.1: Reference Description**. Disponível em: <https://www.dublincore.org/specifications/dublin-core/dces/> Acesso em: 12 dez 2022.

DUBLIN CORE METADATA INITIATIVE (DCMI). **History of the Dublin Core Metadata Initiative**. 2015. Disponível em: <http://dublincore.org/about/history/>. Acesso em: 9 jul. 2021.

DELFINO, S. S.; SOUSA, R. F. M.; PINHO NETO, J. A. S. Desafios da Sociedade da Informação na Recuperação e Uso de Informações em Ambientes Digitais. **Revista Digital Biblioteconomia e Ciência da Informação**. Campinas, SP v.17, p.1-16, 2019.

DIAS, G. A; ANJOS, L. R; RODRIGUES, A. A. Os princípios FAIR: viabilizando o reuso de dados científicos. **Dados científicos: perspectivas e desafios** - João Pessoa, Editora UFPB, 2019. 216 p.

DICIO, **Dicionário Online de Português**. Porto: 7Graus, 2020. Disponível em: <https://www.dicio.com.br/trabalho/>. Acesso em: 04 ago. 2021.

DTL. **Dutch Techcentre for Life Sciences**. 2020. Disponível em: <https://www.dtls.nl/fair-data/fair-data/>. Acesso em: 2 ago. 2020.

FAIR STATEMENT. **Policy Statement on F.A.I.R. Access to Australia's Research Outputs**. 2017. Disponível em: <https://www.fair-access.net.au/fair-statement> Acesso em: 02 fev. 2022

FAIRDATABR. **FairDataBR: uma ferramenta para a avaliação de conjuntos de dados**. 2021. Disponível em: <https://wrco.ufpb.br/fair/index.html>. Acesso em: 13 out de 2022.

FANKHAUSER, E; VRIES, J; WESTZAAN, N; ÅKERMAN, V. **SATIFYD: Self-Assessment Tool to Improve the FAIRness of Your Dataset**. 2019 disponível em: <https://satifyd.dans.knaw.nl/> Acesso em 03 de maio de 2022.

FIOCRUZ, 2022. **Ciência Aberta**. Disponível em: <https://portal.fiocruz.br/go-fair-brasil-saude>. Acesso em: 21 dez. 2021.

FORCE11. **Guiding Principles for Findable, Accessible, Interoperable and Reusable Data Publishing**. Disponível em: <https://www.force11.org/fairprinciples>. Acesso em: 01 mar. 2022.

FORMENTON, D; CASTRO, F. F; GRACIOSO, L. S; FURNIVAL, A. C M; SIMÕES, M G. M. Os padrões de metadados como recursos tecnológicos para garantia da preservação digital. **Biblios**, nº 68. 2017.

FREIRE, M; MEIRELLES, F. R; CUNHA, J. A. P. R. Convergências de Padrões de Interoperabilidade para o Fluxo de Informações entre as Redes de Atenção à Saúde no Portal do Datasus. **Ponto de Acesso**, Salvador, v.13, n.1, p. 87-101, abr. 2019.

FREXIA, F et al. OpenEHR s FAIR-Enabling by Design. **Public Health and Informatics**. European Federation for Medical Informatics (EFMI) and IOS Press. 2021.

GIL, A. C. **Como elaborar projetos e pesquisa**. 7a ed. São Paulo: Atlas; 2022

GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. 7. ed. São Paulo: Atlas, 2019.

GO FAIR, 2021. **Iniciativa Go Fair**. Disponível em: <https://www.go-fair.org/how-to-go-fair/>
Acesso em: 25 nov. 2021

GO FAIR, 2022. **Escritório GO FAIR Brasil**. Disponível em: <https://www.go-fair.org/go-fair-initiative/go-fair-offices/go-fair-brazil-office/>. Acesso em: 05 fev. 2022

GRÁCIO, J. C. A. **Preservação digital na gestão da informação: um modelo processual para as instituições de ensino superior**. São Paulo, SP: Cultura Acadêmica, 2012. 214 p.

Disponível em:

http://www.culturaacademica.com.br/img/arquivos/Preservacao_digital_na_gestao_da_informacao-WEB_v2.pdf. Acesso em: 9 jul. 2021.

HIMSS, 2021. **Interoperabilidade HIMSS e Comunidade HIE**. Disponível em: <https://www.himss.org/membership-participation/himss-interoperability-hie-community>.
Acesso em: 05 de jan. 2022.

HL7. **Introducing HL7 FHIR**. 2019. Disponível em: <https://www.hl7.org/fhir/summary.html>
Acesso em: 25 de set de 2022.

HL7. **Implementing the FAIR principles**. 2021. Disponível em: https://build-fhir-org.translate.google.com/ig/HL7/fhir-for-fair/FHIRandFAIR.html?x_tr_sl=en&x_tr_tl=pt&x_tr_hl=pt-BR&x_tr_pto=sc&x_tr_sch=http. Acesso em: 26 de set de 2022.

KELLEHER, J. D; TIERNEY, B. Data Science. **The MIT Press Essential Knowledge Series**. Cambridge: The MIT Press, 2018.

KOCHE, J. C. **Fundamentos de Metodologia Científica: teoria da ciência e iniciação à pesquisa**. 34. Ed. Petrópolis, RJ: Vozes, 2014.

LESLIE, H. **An archetype or a template?** 2007. Disponível em:

<https://openehr.atlassian.net/wiki/spaces/healthmod/pages/2949149/An+archetype+or+a+template>. Acesso em 09 de jul. 2022.

LESLIE, H. Uso e reutilização do arquétipo OpenEHR em conjuntos de dados clínicos multilíngues: estudo de caso. **J Med Internet Res**. 2020 novembro. doi: 10.2196/23361: 10.2196/23361

LOPES, R. L. P; FERREIRA, P. D; PISA, T. I; TENÓRIO, M. J. Padrões de Normalização em Informática em Saúde - Especialização em Informática em Saúde – UNIFESP – São Paulo, 2018.

MARCONI, M. A; LAKATOS, E. M. **Metodologia Científica**. 8. ed. São Paulo: Atlas, 2022.

MEDEIROS, J. B. **Redação Científica: a prática de fichamentos, resumos, resenhas.** 13. ed. São Paulo: Atlas, 2019.

MICHAELIS, 2021. **Dicionário brasileiro da língua portuguesa.** Disponível em: <https://michaelis.uol.com.br/busca?id=WowM7>. Acesso em: 08 de set. 2021.

MIRANDA N. J. D. O; PINTO V. B. Prontuário eletrônico do paciente: padronização e interoperabilidade. **XVI Encontro Nacional de Pesquisa em Ciência da Informação (XVI ENANCIB).** 2015. Disponível em: <http://www.repositorio.ufc.br/handle/riufc/19187>. Acesso em: 25 jan. 2022.

MIYOSHI, N. S. B; **Arquitetura e métodos de integração de dados e interoperabilidade aplicados na saúde mental.** Tese (doutorado – Programa de Pós-Graduação em Clínica Médica) – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 2018.

MUKHIYA, K. S; LAMO, Y; An HL7 FHIR and GraphQL approach for interoperability between heterogeneous Electronic Health Record systems. **Health Informatics Journal** 2021.

NEVES, A. D; SANTOS, F. R; GUIMARÃES, J. B. I. **Práticas e reflexões sobre a representação da informação em cenários informacionais.** Ebook, São Leopoldo: Karywa, 2019.

OLIVEIRA, J. P; ALMEIDA, M. B; QUINTELA, E. L. Uma visão geral sobre fontes de informação em saúde. **I Congresso ISKO Espanha e Portugal/XI Congresso ISKO Espanha.** 2013. p. 993-1008. Disponível em: http://mba.eci.ufmg.br/downloads/isko_fontes.pdf. Acesso em: 25 jul 2021.

OPENEHR. **What is openEHR?** 2022. Disponível em: https://www.openehr.org/about/what_is_openehr. Acesso em: 14 out de 2022.

OPENEHR BRASIL. 2021. Disponível em: <http://openehr.org.br/index.php/sobre-nos/>. Acesso em: 14 out de 2022.

PAN, S; TRENTESAUX, D; MCFARLANE, D; MONTREUIL, B; BALLOT, E; HUANG, G. Q. Digital interoperability in logistics and supply chain management: state-of-the-art and research avenues towards Physical Internet. **Computers in Industry.** Volume 128. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0166361521000427>. Acesso em: 15 ago. 2021.

PRODANOV, C. C; FREITAS, E. C. **Metodologia do Trabalho Científico: métodos e técnicas da pesquisa e do trabalho acadêmico.** 2. ed. Novo Hamburgo: Feevale, 2013.

REEVE, A. **Managing Data in Motion: Data Integration best practice techniques and technologies.** Waltham, EUA: Elsevier, 2013.

CFM 2002a. **RESOLUÇÃO CFM nº 1.638/2002.** Define prontuário médico e torna obrigatória a criação da Comissão de Revisão de Prontuários nas instituições de saúde. Disponível em:

https://sistemas.cfm.org.br/normas/arquivos/resolucoes/BR/2002/1638_2002.pdf

Acesso em: 03 mar. 2022.

CFM 2002b. **RESOLUÇÃO CFM nº 1.639/2002**. Aprova as "Normas Técnicas para o Uso de Sistemas Informatizados para a Guarda e Manuseio do Prontuário Médico", dispõe sobre tempo de guarda dos prontuários, estabelece critérios para certificação dos sistemas de informação e dá outras providências. Disponível em:

<https://sistemas.cfm.org.br/normas/visualizar/resolucoes/BR/2002/1639>

Acesso em: 03 mar 2022.

CFM 2018. **RESOLUÇÃO CFM nº 2.218/2018**. Aprova as normas técnicas concernentes à digitalização e uso dos sistemas informatizados para a guarda e manuseio dos documentos dos prontuários dos pacientes, autorizando a eliminação do papel e a troca de informação identificada em saúde. Disponível em:

<https://sistemas.cfm.org.br/normas/visualizar/resolucoes/BR/2018/2218>. Acesso em: 10 mar. 2022.

RICHARDSON. J. **Pesquisa Social: Métodos e Técnicas**. 4 ed. São Paulo: Atlas, 2017.

SALES, L. F; VEIGA, V. S; HENNING, P; SAYÃO, L. F. Um panorama histórico da iniciativa GO FAIR: da Europa para o Brasil. *In*: SALES, Luana Farias; VEIGA, Viviane dos Santos; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 9 -22.

SAMPIERI, R. H; COLLADO, C. F; LUCIO, P. B. **Metodologia de Pesquisa**, Tradução de Fátima Conceição Murad, Melissa Kassner, Sheila Clara Dystyler Ladeira. 5. ed. São Paulo: McGraw-Hill, 2013.

SANTOS, L. O. B. S. Prefácio. *In*: SALES, Luana Farias; VEIGA, Viviane dos Santos; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 5-6.

SANTOS, M. **Sistema de registro eletrônico de saúde baseado na norma ISO 13606: Aplicações na Secretaria de Saúde de Estado de Minas Gerais** (Tese de Doutorado em Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte) 2011.

Disponível em: https://repositorio.ufmg.br/bitstream/1843/ECIC-8L8HFJ/1/tese_eci_ufmg_marcelo_rodrigues_dos_santos_2011.pdf. Acesso em 10 fev. 2022.

SARIPALLE, R; RUNYAN, C; RUSSELL, M; Using HL7 FHIR to achieve interoperability in patient health record. **Journal of Biomedical Informatics**. Available online 04 May 2019 1532-0464/ © 2019 Elsevier Inc.

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.

SETZER, V; W. Dado, **Informação, Conhecimento e Competência**. 2015. Disponível em: <https://www.ime.usp.br/~vwsetzer/dado-info.html>. Acesso em: 15 dez. 2022.

SURF. FAIR DATA ADVANCED USE CASES: **From Principles to Practice in the Netherlands**. 2019. Disponível em: <https://zenodo.org/record/1250535#.WyI3RkiFNaQ>. Acesso em: 25 ago. 2021

TOLBOOM, I. **The impact of digital transformation. The Netherlands: Delf University of Technology**, Delft, 2016. Disponível em: <http://resolver.tudelft.nl/uuid:d1d6f874-abc1-4977-8d4e-4b98d3db8265>. Acesso em: 16 maio 2019.

TUTE E; SCHEFFNER I; MARSCHOLLEK M. A method for interoperable knowledgebase data quality assessment. **BMC Med Inform Decis Mak**. 2021.

VENTURA, M. M; O estudo de caso como Modalidade de Pesquisa. **Rev SOCERJ**. 2007.

WILKINSON, M. D. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. **SCIENTIFIC DATA**. DOI: 10.1038. 2016.

WILKINSON, M. D. *et al*. **FAIR Metric FM-R1.1**. 2018. Disponível em: https://github.com/FAIRMetrics/Metrics/blob/master/Distributions/FM_R1.1.pdf. Acesso em: 01 de jun. 2022.

YIN, R. **Estudo de caso: planejamento e métodos**. 5. ed. Porto Alegre: Bookman; 2015.

ZENG, M. L; QIN, J. **Metadata**. New York: Neal-Schuman Publishers, 2008

ANEXOS

ANEXO A - Alinhamento do FHIR com os princípios FAIR

Continue a avaliação de onde parou, basta ler o QRCode.



*Obrigatório

Avaliação Fair

PRINCÍPIO *FINDBLE*

F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente? 1 2 3 4
 5 6 *

- Identificador Persistente
- Endereço *Web*
- Identificador Local
- Sem identificador

F2. Os conjuntos de dados são descritos com metadados? 1 3 4 5 *

- Metadados Ricos
- Metadados Estruturados
- Metadados Simples
- Sem metadados

26/01/23, 20:26

FairDataBR

F3. O identificador está incluído em todos os registros / arquivos de metadados que descrevem os dados? 1 4 6 *

 Sim Não

F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web? 1 6 *

 Sim Não

F5. O conjunto de (meta)dados está publicado em um repositório? 1 2 *

 Público Geral Específico de Domínio Institucional de Local Em Nenhum Repositório

PRINCÍPIO ACCESSIBLE

A1. Quão acessíveis são os (meta)dados? 2 *

 Acessível ao Público Acesso com Embargo Acesso somente aos dados ou somente aos metadados Sem acesso a dados e metadados

A2. Os (meta)dados estão acessíveis *online* sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?² *

- API de Web Service padrão
- Web Service* não padrão
- Download* de arquivo *online*
- Através de acordo individual
- Sem acesso aos (meta)dados

A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?^{1 6} *

- Sim
- Não

A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?^{1 3} *

- Sim
- Não
- Não se aplica

A5. O protocolo (ex: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?^{1 6} *

- Sim
- Não

26/01/23, 20:26

FairDataBR

A7. É possível realizar o *download* dos conjuntos de (meta)dados?² *

Sim

Não

PRINCÍPIO *INTEROPERABLE*

I1. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?^{1 2 4 5} *

Sim

Não

I2. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?

^{1 5} *

Esquema de Comunidade Padrão

Esquema Explícito

Esquema não formalizado

26/01/23, 20:26

FairDataBR

13. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores? *

Links Internos/Externos

Sem *Links*

14. Para representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico? *

Uso de instrumentos de controle terminológico COM identificador global

Uso de instrumentos de controle terminológico SEM identificador global

Nenhum recurso/instrumento

Elementos de dados não descritos

PRINCÍPIO *REUSABLE*

R1. Os conjuntos de (meta)dados estão licenciados? *

Sim

Não

R2. Qual a licença de uso dos conjuntos de (meta)dados? *

Aberta

Restrita

26/01/23, 20:26

FairDataBR

R4. Os conjuntos de (meta)dados possuem proveniência detalhada? 1 6 *

Sim

Não

R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados? 1 5 6 *

Rastreamento de Proveniência

Declaração de Proveniência

Sem Informações de Proveniência

R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio? 1 4 6 *

Padrões de Domínio

Padrões Mínimos de Domínio

Padrões Genéricos

Nenhum Padrão Utilizado

Pontuação Calculada

Aproveite e verifique como ficará sua pontuação em FairDataBR⁺

Seção 1 FINDABLE: Seção 2 ACCESSIBLE: Seção 3 INTEROPERABLE:

4.60

8.50

4.00

Seção 4 REUSABLE:

MÉDIA FAIR:

8.20

6.33

ANEXO B - Alinhamento do OpenEHR com os princípios FAIR

Continue a avaliação de onde parou, basta ler o QRCode.



*Obrigatório

Avaliação Fair

PRINCÍPIO *FINDBLE*

F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente? 1 2 3 4
 5 6 *

- Identificador Persistente
- Endereço *Web*
- Identificador Local
- Sem identificador

F2. Os conjuntos de dados são descritos com metadados? 1 3 4 5 *

- Metadados Ricos
- Metadados Estruturados
- Metadados Simples
- Sem metadados

26/01/23, 20:33

FairDataBR

F3. O identificador está incluído em todos os registros / arquivos de metadados que descrevem os dados? 1 4 6 *

 Sim Não

F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web? 1 6 *

 Sim Não

F5. O conjunto de (meta)dados está publicado em um repositório? 1 2 *

 Público Geral Específico de Domínio Institucional de Local Em Nenhum Repositório

PRINCÍPIO ACCESSIBLE

A1. Quão acessíveis são os (meta)dados? 2 *

 Acessível ao Público Acesso com Embargo Acesso somente aos dados ou somente aos metadados Sem acesso a dados e metadados

A2. Os (meta)dados estão acessíveis *online* sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?² *

- API de Web Service padrão
- Web Service não padrão
- Download de arquivo *online*
- Através de acordo individual
- Sem acesso aos (meta)dados

A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?^{1 6} *

- Sim
- Não

A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?^{1 3} *

- Sim
- Não
- Não se aplica

A5. O protocolo (ex: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?^{1 6} *

- Sim
- Não

26/01/23, 20:33

FairDataBR

A7. É possível realizar o *download* dos conjuntos de (meta)dados?² *

Sim

Não

PRINCÍPIO *INTEROPERABLE*

I1. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?^{1 2 4 5} *

Sim

Não

I2. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?

^{1 5} *

Esquema de Comunidade Padrão

Esquema Explícito

Esquema não formalizado

26/01/23, 20:33

FairDataBR

13. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores? 1 2 3 4 5 *

Links Internos/Externos

Sem *Links*

14. Para representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico? 1 2 6*

Uso de instrumentos de controle terminológico COM identificador global

Uso de instrumentos de controle terminológico SEM identificador global

Nenhum recurso/instrumento

Elementos de dados não descritos

PRINCÍPIO *REUSABLE*

R1. Os conjuntos de (meta)dados estão licenciados? 1 2 4 5 6*

Sim

Não

R2. Qual a licença de uso dos conjuntos de (meta)dados? 1 2 4 *

Aberta

Restrita

26/01/23, 20:33

FairDataBR

R4. Os conjuntos de (meta)dados possuem proveniência detalhada? 1 6 *

- Sim
- Não

R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados? 1 5 6 *

- Rastreamento de Proveniência
- Declaração de Proveniência
- Sem Informações de Proveniência

R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio? 1 4 6 *

- Padrões de Domínio
- Padrões Mínimos de Domínio
- Padrões Genéricos
- Nenhum Padrão Utilizado

Pontuação Calculada

Aproveite e verifique como ficará sua pontuação em FairDataBR⁺

Seção 1 FINDABLE: Seção 2 ACCESSIBLE: Seção 3 INTEROPERABLE:

4.60

8.17

8.75

Seção 4 REUSABLE:

MÉDIA FAIR:

10.00

7.88

ANEXO C - Alinhamento entre o modelo proposto OpenFairEHR e os princípios FAIR

Continue a avaliação de onde parou, basta ler o QRCode.



*Obrigatório

Avaliação Fair

PRINCÍPIO *FINDBABLE*

F1. Os conjuntos de (meta)dados possuem um identificador único, global e persistente? 1 2 3 4
 5 6 *

- Identificador Persistente
- Endereço *Web*
- Identificador Local
- Sem identificador

F2. Os conjuntos de dados são descritos com metadados? 1 3 4 5 *

- Metadados Ricos
- Metadados Estruturados
- Metadados Simples
- Sem metadados

26/01/23, 20:39

FairDataBR

F3. O identificador está incluído em todos os registros / arquivos de metadados que descrevem os dados? 1 4 6 *

 Sim Não

F4. O recurso digital pode ser encontrado em mecanismos de pesquisa na Web? 1 6 *

 Sim Não

F5. O conjunto de (meta)dados está publicado em um repositório? 1 2 *

 Público Geral Específico de Domínio Institucional de Local Em Nenhum Repositório

PRINCÍPIO ACCESSIBLE

A1. Quão acessíveis são os (meta)dados? 2 *

 Acessível ao Público Acesso com Embargo Acesso somente aos dados ou somente aos metadados Sem acesso a dados e metadados

A2. Os (meta)dados estão acessíveis *online* sem a necessidade de intermediação de protocolos ou ferramentas especializadas a partir do momento em que o acesso é permitido?² *

- API de Web Service padrão
- Web Service* não padrão
- Download* de arquivo *online*
- Através de acordo individual
- Sem acesso aos (meta)dados

A3. É possível acessar o conjunto de (meta)dados pelo identificador fornecido?^{1 6} *

- Sim
- Não

A4. Os metadados estão disponíveis mesmo quando o conjunto de dados não estiverem mais acessíveis?^{1 3} *

- Sim
- Não
- Não se aplica

A5. O protocolo (ex: HTTP, SAML, OAI-PMH) é aberto, gratuito e universalmente implementável?^{1 6} *

- Sim
- Não

26/01/23, 20:39

FairDataBR

A7. É possível realizar o *download* dos conjuntos de (meta)dados?² *

 Sim Não

PRINCÍPIO *INTEROPERABLE*

I1. Os conjuntos de (meta)dados estão disponíveis em formatos preferenciais?^{1 2 4 5} *

 Sim Não

I2. Os conjuntos de dados estão estruturados a partir de um esquema de metadados ou modelos de dados aprovados pela comunidade?

^{1 5} *

 Esquema de Comunidade Padrão Esquema Explícito Esquema não formalizado

13. Os conjuntos de (meta)dados estão vinculados a outros (meta)dados usando identificadores? 1 2 3 4 5 *

Links Internos/Externos

Sem *Links*

14. Para representação dos dados ou conjuntos de dados, utilizou-se recursos/instrumentos de controle terminológico? 1 2 6*

Uso de instrumentos de controle terminológico COM identificador global

Uso de instrumentos de controle terminológico SEM identificador global

Nenhum recurso/instrumento

Elementos de dados não descritos

PRINCÍPIO *REUSABLE*

R1. Os conjuntos de (meta)dados estão licenciados? 1 2 4 5 6*

Sim

Não

R2. Qual a licença de uso dos conjuntos de (meta)dados? 1 2 4 *

Aberta

Restrita

26/01/23, 20:39

FairDataBR

R4. Os conjuntos de (meta)dados possuem proveniência detalhada? 1 6 *

- Sim
- Não

R5. Os conjuntos de (meta)dados são acompanhados ou vinculados a uma descrição da origem do fluxo de trabalho que produziu os (meta)dados? 1 5 6 *

- Rastreamento de Proveniência
- Declaração de Proveniência
- Sem Informações de Proveniência

R6. Os (meta)dados estão de acordo com padrões relevantes para o Domínio? 1 4 6 *

- Padrões de Domínio
- Padrões Mínimos de Domínio
- Padrões Genéricos
- Nenhum Padrão Utilizado

Pontuação Calculada

Aproveite e verifique como ficará sua pontuação em FairDataBR⁺

Seção 1 FINDABLE: Seção 2 ACCESSIBLE: Seção 3 INTEROPERABLE:

7.60	10.00	10.00
------	-------	-------

Seção 4 REUSABLE: MÉDIA FAIR:

10.00	9.40
-------	------