

UNIVERSIDADE FEDERAL FLUMINENSE

EDHELMIRA LIMA MEDINA

METAMODELO SISTÊMICO COM ÊNFASE NA PRESERVAÇÃO DE
DADOS CLÍNICOS

Niterói

2019

UNIVERSIDADE FEDERAL FLUMINENSE

EDHELMIRA LIMA MEDINA

METAMODELO SISTÊMICO COM ÊNFASE NA PRESERVAÇÃO DE
DADOS CLÍNICOS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Doutor em Computação. Área de Concentração: Engenharia de Sistemas e Informação.

Orientador:

José Viterbo Filho

Co-orientador:

Orlando Loques

Niterói

2019

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

M491m Medina, Edhelmira Lima
Metamodelo Sistêmico com Ênfase na Preservação de Dados Clínicos / Edhelmira Lima Medina ; José Viterbo Filho, orientador ; Orlando Loques , coorientador. Niterói, 2019. 142 f. : il.

Tese (doutorado)-Universidade Federal Fluminense, Niterói, 2019.

DOI: <http://dx.doi.org/10.22409/PGC.2019.d.06142584725>

1. Metamodelo. 2. Preservação de dados clínicos. 3. Preservação do prontuário eletrônico do paciente. 4. Persistência de informação clínica. 5. Produção intelectual. I. Filho, José Viterbo, orientador. II. , Orlando Loques, coorientador. III. Universidade Federal Fluminense. Instituto de Computação. IV. Título.

CDD -

EDHELMIRA LIMA MEDINA

METAMODELO SISTÊMICO COM ÊNFASE NA PRESERVAÇÃO DE DADOS
CLÍNICOS

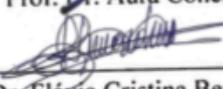
Tese de Doutorado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Doutor em Computação. Área de Concentração: Engenharia de Sistemas e Informação.

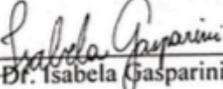
Aprovada em Dezembro de 2019.

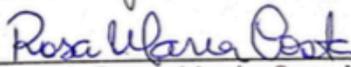
BANCA EXAMINADORA

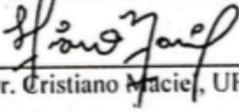

Prof. Dr. José Viterbo Filho – Orientador, UFF


Prof. Dr. Aura Conci, UFF


Prof. Dr. Flávia Cristina Bernardini, UFF


Prof. Dr. Isabela Gasparini, UDESC


Prof. Dr. Rosa Maria Esteves Moreira Costa, UERJ


Prof. Dr. Cristiano Maciel, UFMT

Niterói

2019

Dedicatória(s): À minha mãe Celia, meu esposo Giomar e minhas filhas Emily e Clarisse

AGRADECIMENTOS

A meus orientadores Prof. Dr. Viterbo e Prof. Dr. Orlando, pela ajuda e contribuições neste trabalho. À infinita paciência do Prof. Viterbo que teve comigo, pela confiança em meu trabalho e por nunca duvidar da minha capacidade. Sem ele, a conclusão deste trabalho não seria possível.

Agradeço a meu esposo Giomar pela força, apoio e companheirismo nesta fase da minha vida, sempre me tranquilizando e incentivando a enfrentar novos desafios. A minhas filhas que apesar da curta idade delas foram minha principal motivação.

Agradeço com muito carinho a minha família, minha mãe e irmãos por todo amor, carinho e apoio para a minha formação pessoal e profissional.

Agradeço a meus colegas da pós-graduação pela maravilhosa convivência e a todas as pessoas que de alguma forma contribuíram, direta ou indiretamente, positivamente, para a realização deste trabalho.

A CAPES pelo apoio financeiro concedido para a realização deste trabalho.

RESUMO

Existe um crescimento exponencial de informações clínicas que são geradas e armazenadas diariamente nos sistemas de informação de hospitais e centros de saúde. Estas informações muitas vezes estão distribuídas em banco de dados de diferentes provedores, fazendo com que os profissionais da saúde possuam conhecimento altamente incompleto do paciente. Nesse sentido, a presente tese aborda o tema da preservação de dados clínicos de um indivíduo durante seu ciclo de vida. De acordo à literatura identificamos que na área de saúde os documentos clínicos devem ser preservados historicamente, porém, essa preocupação é pouco explorada. Para isso, esta tese propõe um metamodelo, que em alto nível de abstração, permita representar a integração de processos clínicos (eventos temporais relacionados à saúde que geram documentos clínicos), a representação do prontuário eletrônico e sua preservação a longo prazo. A metodologia adotada nesta tese é baseada no *Design Science Research* (DSR) e consiste em cinco passos: identificação do problema, sugestão, desenvolvimento, validação e conclusão. A etapa de identificação do problema consiste em procurar fontes de conhecimento sobre a preservação digital no domínio da saúde e identificar suas carências. Na etapa de sugestão se propõe um metamodelo para facilitar a visualização de processos clínicos, prontuário eletrônico do paciente e preservação digital. Isto como resultado de uma revisão bibliográfica sistemática da literatura, visto que não se encontrou uma representação similar. Na etapa de desenvolvimento se formaliza o metamodelo que está baseado no meta-metamodelo *Ecore* e é construído no *Eclipse Modeling Framework*. A validação é realizada por meio da instanciação de modelos que descrevem cenários criados com base em questões de competência. Finalmente na conclusão se analisa os resultados obtidos, o qual se considera promissor para os cenários desenvolvidos. Portanto, o metamodelo ajudará a entender como é o processo de geração de dados clínicos e sua posterior preservação.

Palavras-Chave: Preservação de dados clínicos, preservação do prontuário eletrônico do paciente, persistência de informação clínica, metamodelo, processos clínicos.

ABSTRACT

There is an exponential growth of clinical information that is generated and stored daily in the information systems at hospitals and health centers. This information is often distributed in databases of different providers, making health professionals have highly incomplete knowledge of the patient. In this sense, the present thesis addresses the issue of clinical data preservation of an individual during his life cycle. According to the literature, we identified that in the health area, clinical documents should be preserved historically, however, this concern is little explored. For this reason, this thesis proposes a meta-model, which, at a high level of abstraction, allows representing the integration of clinical processes (temporal events related to health that generate clinical documents), the representation of the electronic health record and its long-term preservation. The methodology adopted in this thesis is based on Design Science Research (DSR) and consists of five steps: problem identification, suggestion, development, validation and conclusion. The problem identification stage consists of looking for sources of knowledge about digital preservation in the health field and identifying its needs. In the suggestion stage, a meta-model is proposed to facilitate visualization of clinical processes, electronic health records and digital preservation. This as a result of a systematic literature review, since no similar representation was found. In the development stage, the meta-model is formalized, based on the Ecore meta meta-model and built on the Eclipse Modeling Framework. Validation is performed by instantiating models that describe scenarios created based on competence questions. Finally, the conclusion analyzes the results obtained, which is considered promising for the scenarios developed. Therefore, the meta-model will help to understand the process of generating clinical data and its subsequent preservation.

Keywords: Clinical data preservation, electronic health records preservation, persistence of clinical information, meta-model, clinical process.

LISTA DE FIGURAS

Figura 1. Etapas da metodologia adotada nesta tese.....	20
Figura 2. Organização da tese.....	22
Figura 3. Modelo de Classificação das diferentes estratégias de preservação digital.	25
Figura 4. Características e estratégias identificados nos artigos para a preservação de dados clínicos	27
Figure 5. Especificação do Modelo de Referência OAIS	28
Figura 6. Modelo funcional OAIS	29
Figura 7. Relações dos componentes do Modelo de Referência ISO 13606.....	37
Figure 8. DCC Modelo de Ciclo de Vida	39
Figura 9. Camadas do metamodelo MOF	43
Figura 10. Exemplo de metamodelo de nível 2 e seus modelos derivados.....	44
Figura 11. Meta-metamodelo <i>Ecore</i>	45
Figura 12. Metodologia da revisão sistemática da literatura. Adaptado de	47
Figura 13. Distribuição dos artigos por ano, sobre a preservação de dados clínicos.....	53
Figura 14: Fluxograma para a colheita de dados no PEP	63
Figura 15: Modelo de processos clínicos ou de assistência clínica.	67
Figura 16: Modelo de referência <i>Electronic health record</i>	68
Figura 17: Modelo de Referência OAIS	71
Figura 18: Metamodelo de preservação proposto	77
Figura 19: Fragmento do metamodelo relacionado a processos clínicos	78
Figura 20: Fragmento do metamodelo relacionado ao PEP.....	79
Figura 21: Fragmento do metamodelo relacionado à preservação de dados	81
Figura 22: Visão do metamodelo relacionado às políticas de acesso	82
Figura 23: Exemplo de instanciação de modelos.....	84
Figura 24: Visão geral das formas de validação de metamodelos.	85
Figura 25: Linha do tempo de saúde de Sofia.....	88
Figura 26: Modelo de eventos clínicos	89
Figura 27: Sequência de eventos de saúde.....	90
Figura 28: Modelo de prontuário instanciado	91
Figura 29: Estrutura de um documento clínico.....	91

Figura 30: Modelo de preservação de um documento clínico	93
Figura 31: Exemplo de um fragmento de arquivo CDA.....	94
Figura 32: Exemplo de modelo de políticas de acesso a um documento clínico do PEP.....	95
Figura 33: Exemplo de laudo do exame de ultrassonografia obstétrica em forma de texto. ...	98
Figura 34: Modelo de preservação de DICOM.....	99

LISTA DE TABELAS

Tabela 1. Normas para armazenamento e comunicação de registros de saúde.	32
Tabela 2. Lista de questões de pesquisa.....	48
Tabela 3. Resultados da pesquisa considerando os critérios de inclusão e exclusão.....	50
Tabela 4. Lista dos estudos selecionados.....	51
Tabela 5. Estratégias e características da preservação de dados.....	55
Tabela 6. Modelos utilizados nos estudos selecionados	58
Tabela 7. Lista de atores envolvidos nos estudos abordados.....	59
Tabela 8. Descrição dos elementos do PEP.....	64
Tabela 9. Identificação de características para a preservação de dados clínicos	65
Tabela 10. Resumo de modelos utilizados para criação do metamodelo.	72
Tabela 11. Elementos conceituais para construção do metamodelo.....	74
Tabela 12. Exemplo de regras de políticas de acesso a um PEP.	96

LISTA DE SIGLAS E ABREVIATURAS

CCSD: *Consultative Comitee for Space Data Systems*

CDA: *Clinical Document Architecture*

CONARQ: *Conselho Nacional de Arquivos*

CFM: *Conselho Federal de Medicina*

DICOM: *Digital Imaging and Communication on Medicine*

EHR: *Electronic health record*

EMR: *Electronical medical record*

E-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos

EC: *Exclusion criteria*

HE: *Health Event*

HL7: *Health Level Seven*

ISO: *International Organization for Standardization*

IC: *Inclusion criteria*

MDE: *Model-driven engineering*

MDD: *Model- driven development*

MR: Modelo de Referência

OAIS: *Open Archival Information System*

OMG: *Object Management Group*

OMS: Organização Mundial de Saúde

OpenEHR: *Open Eletronic Health Record Foundation*

PEP: Prontuário eletrônico de paciente

PREMIS: *Preservation Metadata Implementation Strategies*

RES: Registro eletrônico de saúde

SAME: Serviço de Arquivo Médico e Estatística de Saúde

SIM: Sistema de Informações sobre Mortalidade

SIS: Sistema de Informações de Saúde

UML: *Unified Modeling Language*

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO	16
1.1 Motivação	17
1.2 Objetivo	19
1.3 Metodologia.....	20
1.4 Contribuições.....	22
1.5 Organização	22
CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA.....	24
2.1. Preservação Digital.....	24
2.1.1. Estratégias de Preservação Digital	24
2.1.2. Modelo de Referência – OAIS	27
2.2. Prontuario Eletrônico do Paciente	30
2.2.1 Padrões de Informação de Saúde.....	31
2.2.2 Padrão EN ISO 13940:2016	35
2.2.3 Norma ISO 13606.....	36
2.3. Curadoria Digital	37
2.3.1. Modelo do Ciclo de Vida da Curadoria Digital.....	38
2.4. Engenharia Dirigida por Modelos	42
2.4.1 Metamodelos	42
2.4.2 Eclipse Modeling Framework	44
2.5. Considerações Finais	46
CAPÍTULO 3 – REVISÃO DA LITERATURA.....	47
3.1 Planejamento da Revisão de Literatura.....	48
3.1.1 Questões de Pesquisa.....	48
3.1.2 Banco de Dados e Estratégias de Pesquisa	48
3.1.3 Critérios de Inclusão e Exclusão	49
3.2 Condução da Pesquisa.....	50

3.2.1	Seleção e Organização dos Estudos.....	50
3.2.2	Síntese da Informação	51
3.3	Relatório da Revisão	52
3.4	Considerações Finais.....	60
CAPÍTULO 4 – PRESERVAÇÃO DE DADOS CLÍNICOS		62
4.1	Identificação de Características.....	62
4.2	Integração de Modelos	66
4.2.1	Modelo de Processos Clínicos.....	66
4.2.2	Modelo de PEP	67
4.2.3	Modelo de Preservação Digital.....	69
4.2.4	Proposta (Integração de Modelos).....	71
4.3	Considerações Finais.....	73
CAPÍTULO 5 – PROPOSTA DO METAMODELO		74
5.1	Elementos Conceituais para a Preservação de Dados Clínicos.....	74
5.2	Elaboração do Metamodelo.....	75
5.3	Considerações Finais.....	83
CAPÍTULO 6 – VALIDAÇÃO DO METAMODELO.....		84
6.1	Instanciação de Modelos	84
6.2	Técnicas de Validação de Metamodelos	85
6.3	Definição de Questões de Competência.....	85
6.4	Cenário 1	86
6.4.1	Modelos do Cenário 1.....	88
6.5	Cenário 2	94
6.5.1	Modelo do Cenário 2	95
6.6	Cenário 3	96
6.6.1	Modelo do Cenário 3	97
6.7	Considerações Finais.....	99

CAPÍTULO 7 – CONCLUSÕES E TRABALHOS FUTUROS.....	101
7.1 Conclusões	101
7.2 Trabalhos Futuros.....	102
Referências.....	104
Apêndice A – Arquivo XML Ecore do Metamodelo.....	117
Apêndice B – Especificação do Metamodelo	125

CAPÍTULO 1 – INTRODUÇÃO

A implantação de Sistemas de Informação em Saúde (SIS) se disseminou nas últimas décadas em diferentes entidades de saúde, segundo dados da Organização Mundial da Saúde (OMS). Consequentemente há um crescimento exponencial de informações clínicas que são geradas e armazenadas diariamente nestes sistemas. Estas informações ou documentos clínicos (e.g. consultas, exames médicos, exames por imagem e laudos) são organizadas em prontuários médicos ou Prontuários Eletrônicos do Paciente (PEP). Tal registro contém resultados de exames e considerações pessoais do médico sobre o paciente e informação acerca do tratamento utilizado (Rudersdorf, 2012; Graña and Jackwoski, 2015; Lozano, 2016).

Os PEPs servem como repositórios de informações relativas à assistência em saúde. São coleções de fontes de informação relativa ao bem-estar físico, mental e social de um indivíduo (Galvão and Ricarte, 2012). A grande produção de dados e registros médicos no PEP o torna uma fonte de informação valiosa para as pesquisas e estudos clínicos epidemiológicos. Os resultados poderão ser utilizados para melhorar a assistência ao paciente, como fonte de extração de indicadores de saúde de uma comunidade e como informação estratégica para a tomada de decisões (Scott, 2007; Corn, 2009; Galvão and Ricarte, 2012), levando em consideração aspectos de privacidade dos dados do paciente, aspectos éticos e legais (Van-Gorp et al., 2014; Bheemanathini, 2019). Da mesma forma, esses prontuários muitas vezes são gerenciados por diferentes tipos de SIS que estão distribuídas em bancos de dados de diferentes provedores. Isso provoca uma alta fragmentação, duplicidade e inclusive perda de dados, fazendo que os profissionais da área da saúde possuam um conhecimento incompleto do paciente, o qual pode afetar à tomada de decisões nos diagnósticos (Evans, 2016; Mihailescu et al., 2017). Assim a preservação digital visa cumprir o objetivo de: integridade, autenticidade e confidencialidade (Gartner and Lavoie, 2013), porém a preservação dos PEPs por longo tempo é ainda pouco explorada na literatura.

Nesse contexto, surge a necessidade de implantação de sistemas de gestão arquivística visando a preservação, garantia de autenticidade, confiabilidade, privacidade, controle de acesso e a integridade desses documentos clínicos associados ao indivíduo (Scott, 2007; Corn, 2009; Goossen, 2014; Pinto and Sales, 2017). As soluções de preservação das informações clínicas do indivíduo continuam sendo objeto de pesquisa, conforme abordado por vários autores (Malone, et al., 2014; Chaczko et al., 2018; Jones, 2019). Os dados médicos podem ser

mantidos enquanto os respectivos pacientes estiverem vivos ou mesmo além desse período. Esta última afirmação refere-se na literatura como patrimônio de dados digitais (Fatoric and Seekamp, 2017; Chaczko et al., 2018, Gordon, 2018). Assim, o período de proteção necessário pode ser superior a 100 anos. Outros exemplos de dados confidenciais de longa duração são os dados gnômicos, segredos governamentais e dados fiscais (Becker et al., 2011; Winter, 2018).

1.1 MOTIVAÇÃO

Atualmente, grandes quantidades de dados são armazenadas em servidores baseados na nuvem (Celesti et al., 2016; Abouelmehdi et al., 2018; Chentharra et al., 2019; Zhu et al., 2019), e essa quantidade continuará crescendo massivamente no futuro. No domínio da área da saúde, por exemplo, no Japão, existem hospitais que usam armazenamento em nuvem para proteger seus dados médicos sensíveis contra perdas, isto devido a catástrofes naturais que pudessem acontecer (Askhoj et al., 2011; Kuroda et al., 2012). Assim os dados armazenados digitalmente precisam de proteção durante toda sua vida útil (Walters and Skinner, 2011; Gartner and Lavoie, 2013; Fatoric and Seekamp, 2017; Zhu et al., 2019) o que pode ser muito tempo.

De acordo com alguns autores, o valor clínico da maior parte das informações armazenadas nos registros médicos é perdido à medida que passa o tempo (Evans, 2016). Logo, a preservação e arquivamento a longo prazo irá representar um recurso valioso, pois o acesso a essa informação pode ajudar a melhorar a assistência ao paciente, reduzir custos dos cuidados de saúde e contar com um melhor acompanhamento dos resultados. Isso traz benefícios à consistência do tratamento e aumenta significativamente a informação disponível para a tomada de decisões (Burns, 2011; Misra et al., 2012; Bunawan and Nordin, 2015). Em consequência, será importante nas atividades de pesquisa, por exemplo em *big data*, que por meio do uso de técnicas de *data mining* e análise de dados torna possível extrair informações úteis, padrões e tendências de doenças, frequentemente desconhecidos e que podem ser prognosticados a partir de grandes quantidades de dados (Gonçalves et al., 2015; Celesti et al., 2016).

É assim que a preservação das informações clínicas dos pacientes procedentes dos PEP, tornam-se importantes para garantir a qualidade do atendimento e a eficiência nos diagnósticos e tratamentos (Kiefer et al., 2012; Evans, 2016). A preservação de dados compreende a combinação de políticas, estratégias e ações para garantir o acesso e reprodução que precisam de conteúdo autenticado, reformatado e originalmente digitais ao longo do tempo, independentemente dos desafios de falha das mídias e de mudança tecnológica (Ferreira, 2006;

Márdero, 2008; Lavoie, 2014). Portanto, esse conjunto de atividades ou processos responsáveis por garantir o acesso continuado, por longo prazo, à informação e ao patrimônio cultural existente em formatos digitais, exige proteção durante toda sua vida útil (Cunha and Lima, 2012; Cordeiro et al., 2018).

Um problema relacionado à preservação de dados clínicos é a interoperabilidade entre diferentes SIS, o qual dificulta o acesso ao PEP (Park et al., 2012). Os SIS existentes, além de lidar com a natureza altamente desconectada e complexa dos processos de saúde, devem contar com informações integradas e padronizadas sobre o PEP (Graña and Jackwoski, 2015). Por isso, seja em soluções tecnológicas ou na adoção de padrões de nomenclatura, os investimentos e pesquisas em interoperabilidade emergem como uma demanda cada vez mais urgente nas instituições de saúde (Park et al., 2012). Embora existam padrões bem estabelecidos, como é o caso do DICOM¹ (*Digital Imaging and Communications in Medicine*) e o HL7² (*Health Level Seven*), capazes de resolver aspectos específicos dos mecanismos para integração, estes padrões não resolvem os problemas de nomenclatura e semântica médica (Park et al., 2012; Bunawan and Nordin, 2015). Por exemplo, se um indivíduo é atendido pela primeira vez em um lugar diferente do habitual, ou mesmo se ele volte ao local do primeiro atendimento, nem sempre há como acessar seus dados históricos (exames passados, antecedentes, alergias, etc.), pois essas informações podem estar armazenadas em diferentes SISs sem nenhuma comunicação (Evans, 2016; Mihailescu et al., 2017). Deste modo, a atividade da preservação de dados digitais a partir do armazenamento organizado, visa proteger os conteúdos para serem acessados em um futuro distante, para assim poder perpetuar o ciclo de disseminação do conhecimento (Ferreira, 2006). O volume crescente, a heterogeneidade e as características da informação digital, não apenas podem ser independentes dos suportes, senão também dos formatos, particularmente a facilidade de manipulação, interligação e reutilização (Bunawan and Nordin, 2015; Evans, 2016)

Diante desse contexto, os padrões são essenciais para o desenvolvimento e a implantação de sistemas que visem implementar estratégias de preservação. Assim, modelos de informação genérica, modelos clínicos, ontologias e terminologias foram identificados como artefatos necessários para alcançar este objetivo. Por outro lado, um metamodelo permitiria a integração de diferentes propostas e tecnologias mantendo a integridade e consistência da arquitetura e

¹ <https://www.dicomstandard.org/>

² <https://www.hl7.com.br/>

facilitando evoluções progressivas (Rabbi et al., 2014; Cohen et al., 2015; Lozano, 2016; Chazcko et al., 2018).

Na literatura observamos que existem poucas iniciativas que abordam em profundidade o tema da preservação focado no domínio da saúde. Embora existam modelos de referência de preservação de dados genéricos (Ross, 2012), padrões interoperabilidade (Park et al., 2012) e representação do PEPs (Fung et al., 2010), não identificamos um modelo unificado que ajude entender o processo de preservação de dados clínicos. Portanto, essa ausência aparece como uma barreira importante para as iniciativas da preservação no domínio da área de saúde.

Na área de metamodelagem existem iniciativas que buscam aumentar o nível de abstração no processo de desenvolvimento de software. Um metamodelo permite a integração de diferentes propostas e tecnologias mantendo a integridade e consistência da arquitetura, facilitando evoluções progressivas (Rabbi et al., 2014; Cohen et al., 2015). Isso reforça a ideia de que propostas de metamodelos que ajudem a abstrair e esclarecer os processos de preservação de dados clínicos podem ser uma contribuição relevante para essa área.

1.2 OBJETIVO

O objetivo deste trabalho é desenvolver um metamodelo que permita representar o conhecimento associado ao processo de armazenamento e preservação da informação clínica associada a um indivíduo.

Para alcançar o objetivo geral acima, foram definidos os seguintes objetivos específicos:

- Estudar o domínio relacionado ao PEP e seus padrões;
- Estudar os conceitos básicos de metamodelagem e metodologias para a sua construção e validação;
- Analisar e identificar características e elementos conceituais que constituem a preservação de dados;
- Consolidar a fundamentação teórica da preservação de dados clínicos por meio de uma revisão sistemática da literatura;
- Integrar em um alto nível de abstração os processos clínicos, PEP e preservação de dados por meio da construção de um metamodelo;
- Validar o metamodelo proposto.

1.3 METODOLOGIA

A metodologia de pesquisa utilizada nesta tese se alinha com o *Design Science Research* (DSR) que é uma metodologia muito utilizada na área de modelagem e ontologias (Peffer et al., 2006; Silva, 2018; Oliveira, 2019). Ela envolve construir, investigar, validar e avaliar artefatos, a modo de resolver problemas práticos (Hevner et al., 2004; Peffer et al., 2006) criando artefatos inovadores. Os artefatos podem ser: construtos (entidades e relações), modelos (abstrações e representações), métodos (algoritmos e práticas) e instanciações (implementação de sistemas e protótipos) (Hevner et al., 2010).

Assim, em conformidade com o objetivo desta tese o resultado pode ser classificado como um artefato que representará o conhecimento associado aos processos de armazenamento e preservação da informação clínica relacionado a um indivíduo. Para tal, nessa tese a dividimos a metodologia em cinco etapas: entendimento ou conscientização do problema, sugestão, desenvolvimento, avaliação e conclusão conforme mostrado na Figura 1. Estas etapas representam uma esquematização das etapas do método DSR e são inspirados na proposta de Takeda et al. (1990), Vaishnavi and Kuechler (2004) e Peffer et al. (2006).

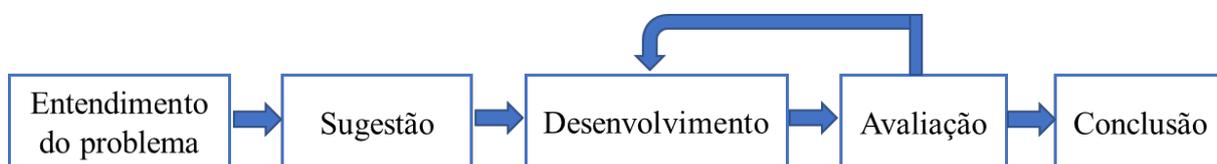


Figura 1. Etapas da metodologia adotada nesta tese. Adaptado de Takeda et al. (1990), Vaishnavi and Kuechler (2004)

A seguir descrevemos brevemente as etapas desta metodologia:

- **Entendimento do problema:** Esta etapa procura definir o problema de pesquisa específico previamente identificado e justificar o valor de uma solução. Assim, o entendimento é utilizado para desenvolver uma solução eficaz, justificando seu valor e importância. No contexto dessa tese, a pesquisa se inicia com uma motivação baseada em uma revisão informal da literatura e a identificação de fontes de conhecimento. Em seguida se realiza um estudo do contexto técnico da preservação digital no domínio da saúde e suas carências, para isto se realiza uma revisão bibliográfica sistemática e uma fundamentação teórica sobre o domínio de preservação de dados clínicos. Desta maneira, estabelecemos os recursos necessários para esta etapa que incluem o conhecimento do estado do problema e a importância da sua solução.

- **Sugestão:** Esta etapa busca inferir os objetivos de uma solução a partir da etapa anterior. As sugestões podem ser quantitativas (solução melhor do que as atuais) ou qualitativas (soluções para problemas não abordados até agora). Os recursos necessários nesta etapa são o conhecimento do estado do problema, soluções atuais e sua eficácia, se houver. Nesse sentido, como sugestão de solução ao problema identificado propomos criar um metamodelo que reutilize modelos padronizados existentes, seguindo o que sugere a engenharia de software orientada a modelos (MDE - *Model-Driven Engineering*), para abstrair os conceitos relacionados a processos clínicos, prontuário eletrônico do paciente e preservação digital.
- **Desenvolvimento:** Esta etapa visa criar a solução de artefatos, os quais podem ser: construções, modelos, métodos ou instâncias (Hevner et al., 2004; Peffers et al., 2006). As atividades realizadas incluem determinar a funcionalidade desejada do artefato e sua arquitetura e, em seguida, criar o artefato real. Uma vez que são identificados os elementos conceituais do domínio do problema, nesta etapa, o metamodelo proposto é formalizado. Este metamodelo é uma instância do meta-metamodelo *Ecore* e é construído no *Eclipse Modeling Framework* (EMF)³.
- **Avaliação.** Esta etapa tem por objetivo observar e medir o quão bem o artefato suporta uma solução para o problema. Para isso é necessário ter conhecimentos de métricas e técnicas de análise relevantes. Caso a solução não seja satisfatória é possível iterar e voltar à etapa anterior para tentar melhorar a eficácia do artefato. Dessa forma, para avaliar o metamodelo proposto nesta tese, utilizamos a estratégia de criar instâncias que modelam cenários, que por sua vez são criadas com base em questões de competência que o metamodelo deve responder.
- **Conclusão:** Nesta etapa se faz uma análise dos resultados obtidos para verificar se os resultados desejados foram atingidos. Na sequência são identificadas as suas limitações e finalmente se define trabalhos futuros. Por meio da análise dos resultados obtidos nesta tese, foi possível concluir que o metamodelo proposto responde aos parâmetros de avaliação sugeridos e se mostra adequado para resolver ao problema identificado.

³ <http://www.eclipse.org/emf>

1.4 CONTRIBUIÇÕES

A principal contribuição deste trabalho é a proposta de um metamodelo que permita instanciar modelos para construção de PEP que visem a preservação e acesso da informação clínica de indivíduos ao longo da sua vida.

As contribuições secundárias são:

- Verificação do estado da arte relacionado a abordagens existentes para preservação de PEPs.
- Integração de modelos de referência e padrões existentes para a construção de um metamodelo que permita abstrair os processos clínicos, PEPs e sua preservação considerando aspectos de privacidade;
- Elaboração de modelos de diferentes cenários que envolvem a preservação no domínio da saúde, criados em consequência do uso da técnica de formulação de questões de competência na validação do metamodelo;

1.5 ORGANIZAÇÃO

Esta tese é composta por sete capítulos e são estruturados de modo a se alinhar à metodologia proposta conforme apresentado na Figura 2.

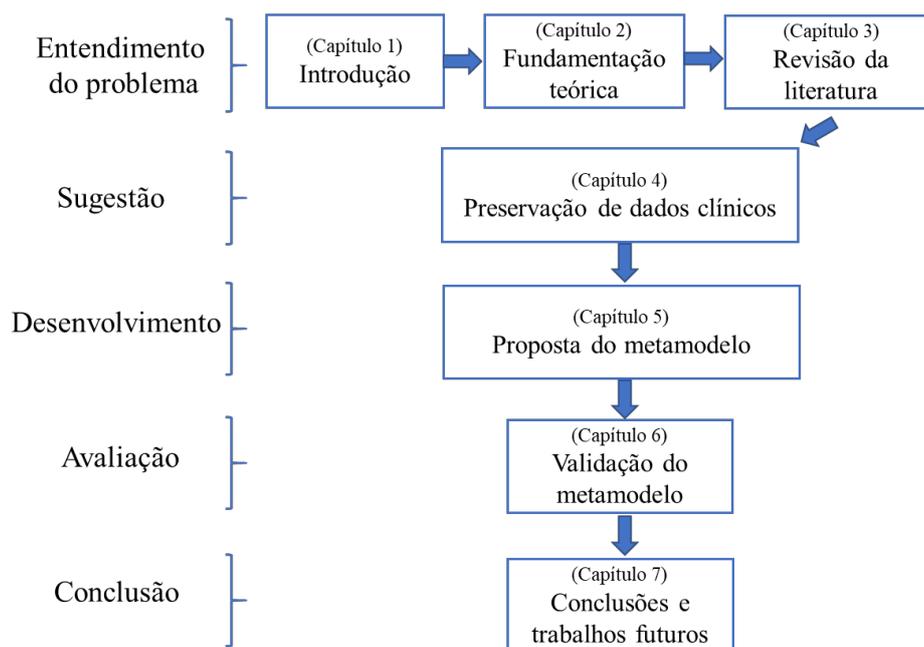


Figura 2. Organização da tese

Os três primeiros capítulos correspondem à etapa de entendimento do problema da metodologia utilizada. No Capítulo 1 se apresenta o assunto sobre o qual a pesquisa se realiza, os objetivos, metodologia adotada e as contribuições da tese. O Capítulo 2 corresponde à fundamentação teórica e contém conceitos básicos no contexto da preservação digital, PEP e meta modelagem. No Capítulo 3 se realiza uma revisão sistemática da literatura que visa identificar trabalhos relacionados e apresentar o estado da arte relacionado à preservação de PEPs. A etapa de sugestão é desenvolvida no Capítulo 4 onde são identificados um conjunto de características do contexto da preservação digital de PEPs e com base nos modelos existentes é proposto a elaboração de um metamodelo que vise abstrair as características identificadas e cenários propostos. Já no Capítulo 5 o metamodelo é formalizado como sendo uma instância do meta-metamodelo *Ecore*. Em seguida, no Capítulo 6, busca-se validar o metamodelo proposto por meio de questões de competência e a instanciação de modelos. Finalmente, no Capítulo 7 são apresentados as conclusões e trabalhos futuros.

CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA

Este capítulo aborda os fundamentos teóricos envolvidos nesta tese. Primeiramente são abordados conceitos relacionados à preservação de dados e suas estratégias, na sequência são descritos conceitos referentes aos sistemas de prontuário eletrônico. Por fim, são descritos os principais modelos de referência para preservação e curadoria digital.

2.1. PRESERVAÇÃO DIGITAL

A preservação digital pode ser definida como o planejamento da alocação de recursos e aplicação de métodos e tecnologias para assegurar que a informação digital de valor contínuo permaneça acessível e utilizável (Gartner and Lavoie, 2013). Entretanto, a *Association for Library Collections & Technical Services – ALCTS (2007)* define a preservação digital como a combinação de políticas, estratégias e ações para garantir o acesso e a reprodução de uma determinada informação. Sendo que esta informação deve ser autenticada, reformatada e digitalmente carregada ao longo do tempo, independentemente dos desafios de falha de mídia e mudança tecnológica.

Por outro lado, Ferreira (2006) afirma que a preservação digital se trata de um conjunto de atividades ou processos aplicados para garantir o acesso contínuo a informações a longo prazo e o patrimônio cultural remanescente em formatos digitais. Dessa forma, deve-se garantir o acesso contínuo e confiável a objetos digitais autênticos enquanto eles forem considerados valiosos. Isso deve incluir não apenas atividades técnicas, mas também um conjunto de considerações estratégicas e organizacionais relacionadas à sobrevivência e gerenciamento de material digital (Pennock, 2006).

2.1.1. ESTRATÉGIAS DE PRESERVAÇÃO DIGITAL

De acordo com Piconi (2014) que resgata o trabalho de Lee et al. (2002), as estratégias mais conhecidas para resolver o problema da preservação digital no sentido estrutural são a emulação, migração e encapsulamento. Entretanto, Thibodeau (2002), além de considerar mais duas classes (preservação da tecnologia e atualização/refrescamento) ele organiza as diferentes estratégias em um mapa bidimensional, posicionando no seu extremo esquerdo as estratégias centradas na preservação do objeto físico/lógico e no extremo oposto as estratégias centradas na preservação do objeto conceitual (ver Figura 3). No eixo vertical as diversas estratégias são dispostas por meio do seu grau de especificidade, isto é, se são estratégias apenas aplicáveis a

uma dada classe de objetos digitais ou se trata de estratégias genéricas, passíveis de ser administradas a qualquer classe de objetos digitais.

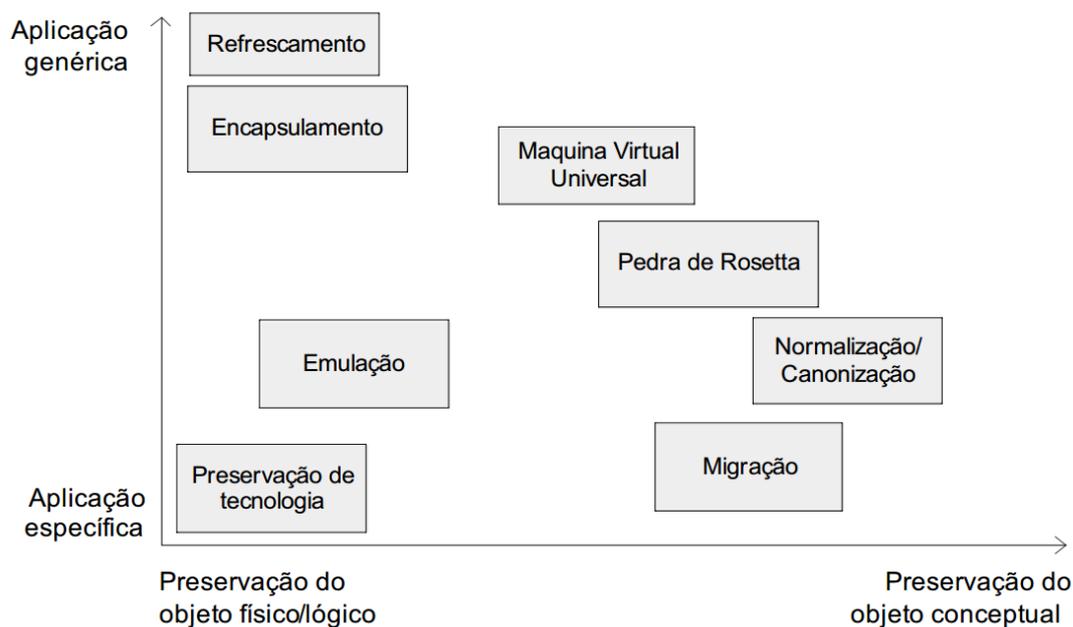


Figura 3. Modelo de Classificação das diferentes estratégias de preservação digital. Adaptado de Thibodeau (2002)

Observamos que o objeto conceptual conhecido também como objeto semântico é o entendimento desde o ponto de vista humano e constitui aquilo que deve ser preservado (Ferreira, 2006). A seguir são descritas brevemente as classes mencionadas na Figura 3:

- **Preservação de tecnologia:** Esta estratégia consiste na conservação e manutenção de todo o hardware e software necessários à correta apresentação dos objetos digitais (Ferreira, 2006). Observe-se que o foco não se concentra no objeto conceptual, mas sim na preservação do objeto digital na sua forma original.
- **Refrescamento / atualização:** Consiste na transferência de informação de um suporte físico de armazenamento para outro mais atual, antes que o primeiro se deteriore ou se torne irremediavelmente obsoleto (Thibodeau, 2002).
- **Emulação:** As estratégias de emulação baseiam-se na utilização de um software, designado emulador, capaz de reproduzir o comportamento de uma plataforma de hardware e/ou software, em uma outra que seria incompatível (Rothenberg, 1999). A capacidade de preservar, com um elevado grau de fidelidade e com as características e as funcionalidades do objeto digital original é uma grande vantagem (Lee et al., 2002).

- **Migração/conversão:** A migração consiste na transferência periódica de material digital de uma dada configuração de hardware/software para uma outra, ou de uma geração de tecnologia para outra subsequente (Ferreira, 2006).
- **Encapsulamento:** Consiste em preservar, juntamente com o objeto digital, toda a informação necessária e suficiente para permitir o futuro desenvolvimento de conversores, visualizadores ou emuladores. Esta informação poderá consistir, por exemplo, em uma descrição formal e detalhada do formato do objeto preservado (Time, 2002).
- **Máquina virtual Universal:** Refere-se ao desenvolvimento de máquinas virtuais que podem executar funções essenciais em diversas plataformas. Segundo Thibodeau (2002), a máquina virtual geralmente fornece apenas um conjunto limitado de funções básicas o que termina afetando o desempenho do software construído para esse fim.
- **Pedra de Roseta:** Esta abordagem começa com tipos de dados, mas, no lugar de articular as propriedades essenciais de cada tipo de dado, constrói uma amostra representativa de objetos desse tipo. Acrescenta amostras paralelas dos mesmos objetos em linguagens diferentes, a segunda amostra criada é chamada, como no grego, de Pedra de Roseta original (Thibodeau, 2002).
- **Normalização/Canonização:** A normalização de formatos refere-se que na preservação existem diversos tipos de formatos. Thibodeau (2002) afirma que deveria ser criada uma superclasse de tipos de dados originais de forma que permita incorporar atributos e métodos essenciais do formato original.

Sob perspectiva semelhante os autores Scott (2007), Corn et al. (2009) e Goossen (2014) coincidem e descrevem um conjunto de características (armazenamento, meios de armazenamento, obsolescência tecnológica, escalabilidade e custos/financiamento) e estratégias para a preservação de dados digitais, conforme se mostra na Figura 4. A premissa deles é que todas essas estratégias devem garantir que o documento mantenha sua autenticidade, integridade, funcionalidade e acesso contínuo pelo tempo que for necessário.

No que se refere a dados clínicos, as estratégias maiormente consideradas são a interoperabilidade semântica, acessibilidade, compatibilidade, filtragem e manutenção dos dados, migração, emulação e extensibilidade. Por exemplo, a interoperabilidade permite aumentar a qualidade do atendimento e diminuir a duplicação de exames e prescrições que o paciente realiza. A acessibilidade se refere à prevenção de acesso desautorizado às informações. A compatibilidade se refere a que um documento deve ser compatível em suporte e formato. A

filtragem e manutenção visa o gerenciamento das informações. A migração considera manter o documento acessível a longo prazo por meio de políticas de atualização de hardware e software. A emulação reproduz um ambiente computacional para que seja possível a execução de outros sistemas sobre ele. E a extensibilidade é a adoção antecipada de abordagens baseados em padrões para desenvolvimento de tecnologia com a finalidade de minimizar o impacto da obsolescência tecnológica.

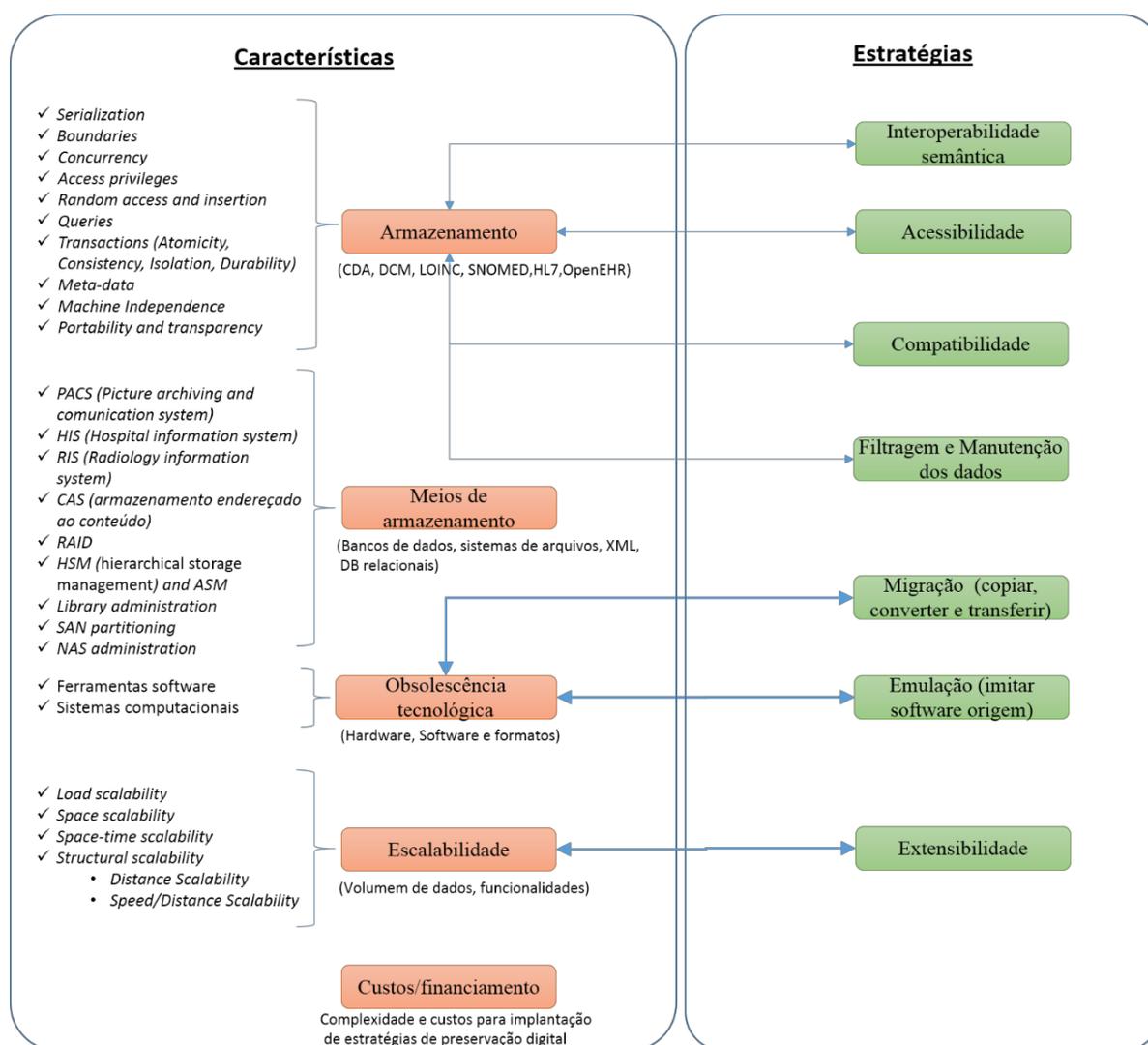


Figura 4. Características e estratégias identificados nos artigos para a preservação de dados clínicos (Fonte: autor)

2.1.2. MODELO DE REFERÊNCIA - OAIS

De acordo com Earl (1990) um modelo de referência é uma representação externa e explícita de parte da realidade do ponto de vista de um indivíduo. Serve para apoiar a execução de tarefas relacionadas com aquela parte da realidade, sejam operacionais ou gerencias. Um modelo de

referência proporciona uma sintaxe (estrutura de dados e formatos básicos) para poder representar qualquer tipo de informação em qualquer domínio. São formalizados por meio de uma linguagem definido por construtos de modelagem.

O modelo de referência *Open Archival Information System* – OAIS (ISO 14721, 2012) nasceu como uma iniciativa do Comitê Consultivo para Sistemas de Dados Espaciais (*Consultative Comitee for Space Data System - CCSDS*) com a *International Organization for Standardization* (ISO) cujo objetivo era de desenvolver um conjunto de padrões capazes de regular o armazenamento a longo prazo de informações digitais produzidas em missões espaciais (CCSDS, 2012; Lavoie, 2014).

O modelo de referência OAIS, é um modelo conceptual que visa identificar os componentes funcionais que devem fazer parte de um sistema de informação dedicado à preservação digital. A especificação do modelo de referência do ambiente para um arquivo do tipo OAIS é mostrada na Figura 5.

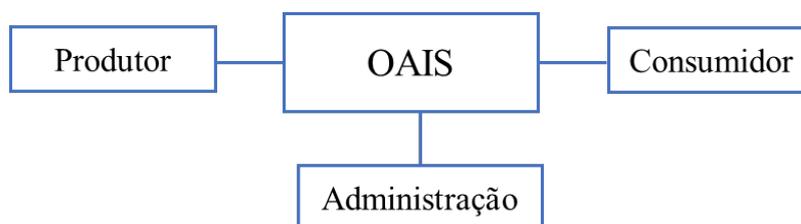


Figure 5. Especificação do Modelo de Referência OAIS. Fonte: (CCSDS, 2012)

Como se observa na Figura 5, o modelo está composto por três componentes principais: Produtor, Administrador e Consumidor. E por pacotes de informação trocados no interior de um repositório digital compatível com o modelo de referência OAIS (Ferreira, 2006). Estes componentes são descritos a seguir (CCSDS, 2012):

- **Produtor:** É uma entidade externa e representa o papel desempenhado por pessoas ou sistemas clientes que fornecem as informações a serem preservadas.
- **Administração:** Ou gerenciamento, é responsável pelas operações diárias de manutenção e sobretudo pela parametrização e monitorização dos processos desencadeados no interior do repositório.
- **Consumidor:** É o papel desempenhado por pessoas ou sistemas de clientes que interagem com os serviços do OAIS para encontrar e adquirir informações preservadas

de interesse. Uma classe especial de consumidores é a comunidade-alvo, que são um conjunto de consumidores capazes de entender as informações preservadas. Vale mencionar que um determinado indivíduo ou sistema pode atuar no papel tanto de consumidor como de produtor.

A Figura 6 apresenta a representação do OAIS mais difundida na literatura. Está composta por seis entidades funcionais, assim como os pacotes de informação (Pacote de Informação de Submissão, Pacote de Informação de Arquivamento, Pacote de Informação de Disseminação) trocados dentro de um repositório digital compatível com este modelo de referência. Estas entidades gerenciam o fluxo de informação entre as entidades que conformam o ambiente OAIS e identificam os componentes funcionais dos arquivos relacionados com a preservação dos objetos digitais (Márdero, 2008).

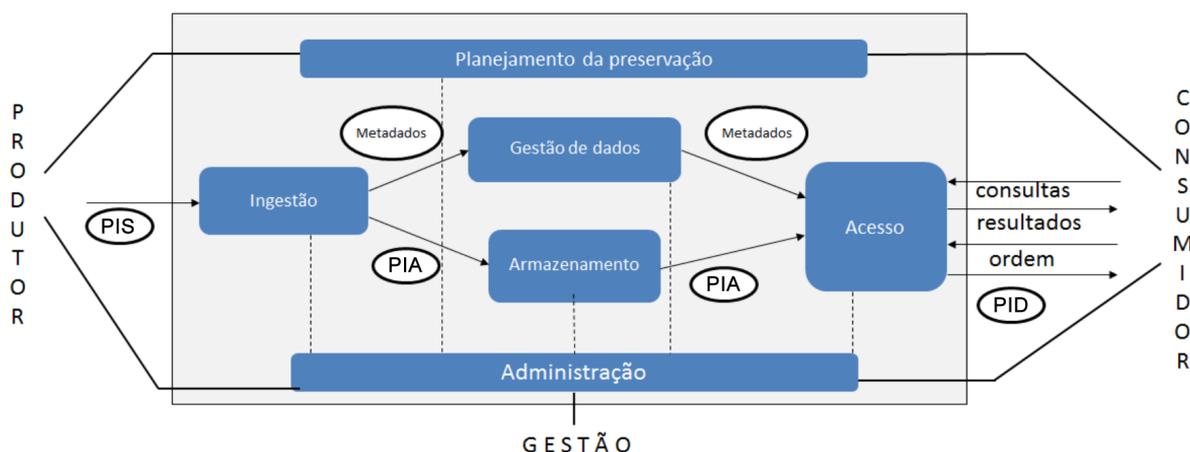


Figura 6. Modelo funcional OAIS. Fonte (CCSDS, 2012)

As descrições das funções de cada entidade presentes no modelo de referência OAIS são descritos a seguir:

- **Ingestão:** Fornece os serviços e funções para aceitar os pacotes de informações de submissão (PIS) dos produtores e preparar o conteúdo para armazenamento e gerenciamento dentro do arquivo.
- **Armazenamento:** Fornece os serviços e funções para armazenamento, manutenção e recuperação do pacote de informações de arquivamento (PIA).
- **Gestão de Dados:** Fornece os serviços e funções para preencher, manter e acessar as informações descritivas (ou metainformação) que identificam e documentam os acervos do arquivo e os dados usados para gerenciar dito arquivo.
- **Administração:** É responsável pelas operações diárias de manutenção e sobretudo pela parametrização e monitorização dos processos desencadeados no interior do repositório.

Este componente interage com todos os restantes de modo a assegurar o correto funcionamento do mesmo.

- **Plano de Preservação:** Fornece os serviços e funções para monitorar o ambiente do OAIS, fornecendo recomendações e planos de preservação para garantir que as informações armazenadas no OAIS permaneçam acessíveis e compreensíveis pela Comunidade-alvo a longo prazo mesmo que o ambiente de computação original se torne obsoleto.
- **Acesso :** Fornece os serviços e funções que dão suporte aos consumidores na determinação da existência, descrição, localização e disponibilidade das informações armazenadas no OAIS e permitem que os consumidores solicitem e recebam produtos de informação. Os pacotes que são entregues ao consumidor assumem a forma de Pacotes de Informação de Disseminação (PID).

2.2. PRONTUARIO ELETRÔNICO DO PACIENTE

Segundo a organização mundial da saúde (OMS, 2003) o Prontuário Eletrônico do Paciente (PEP), do inglês *Electronic Health Record* (EHR), define-se como um conjunto de informações de saúde e assistência de um paciente durante toda sua vida. Nela se registram todo tipo de procedimentos desde consultas, administração de medicamentos, resultados de exames e informações demográficas. Na literatura (Yina, 2010; Araujo et al., 2014; Evans, 2016) é possível encontrar uma diversidade de nomes para o PEP, como Registro Eletrônico de Paciente (EPR - *Electronic Patient Record*) ou Registro Médico Eletrônico (EMR – *Electronic Medical Record*). Nesta tese, estes conceitos serão considerados sinônimos.

Um PEP é uma coleção longitudinal de informações eletrônicas de saúde dos pacientes ou populações individuais. Esses registros descrevem sistematicamente a documentação do histórico médico e dos cuidados de um único paciente ao longo do tempo, dentro da jurisdição de um profissional de saúde em particular (OMS, 2003). Os documentos que compõem o PEP são chamados de registros médicos ou prontuários médicos, que segundo o Conselho Federal de Medicina (CFM)⁴, incluem uma variedade de tipos de "anotações" inseridas ao longo do tempo pelos profissionais de saúde, registrando observações e administração de medicamentos e terapias, pedidos de administração de medicamentos, resultados de testes, raios-x, relatórios, etc. (Yina, 2010).

⁴ <http://portal.cfm.org.br/>

Os autores Araujo et al. (2014) e Graña and Jackwoski (2015) afirmam que o PEP é um instrumento fundamental para apoiar a saúde pública, representa um grande repositório de informações que cria diretrizes de conduta em saúde, incorpora todos os cuidados de saúde e doença onde quer que tenha sido fornecido. Ele pode capturar informações clínicas diretamente no ponto em que são geradas e integrar informações de várias fontes como: o hospital, o laboratório, outros profissionais de saúde e o próprio paciente. Segundo Araujo et al. (2014) algumas características da adoção do PEP são:

- Acesso rápido à informação do estado de saúde de um paciente,
- Melhoria do processo de tomada de decisões,
- Proporciona ao médico a gestão de conhecimento dentro do seu consultório, aumentando a eficácia e efetividade de tratamento e cuidado do paciente,
- Contribui para a redução dos custos do setor, principalmente reduzindo o uso excessivo (duplicação de testes de laboratório) e o uso indevido de serviços (fraude e erros de diagnóstico).

Contudo, para que esta ferramenta atinja plenamente o seu potencial é fundamental que a sua adoção seja disseminada, para que diferentes sistemas possam trocar informações de saúde, em um meio seguro, onde e quando se precisem delas. Algumas aplicações-chaves são: checagem de medicamentos, agendamento baseado em necessidades, suporte à decisão, *Business Intelligence* e gerenciamento de doenças ou de casos que necessitam que a informação tenha uma representação semântica no banco de dados.

2.2.1 PADRÕES DE INFORMAÇÃO DE SAÚDE

No PEP existem problemas técnicos relacionados à variedade de padrões de saúde para diferentes propósitos, isto pode ser visto na Tabela 1. Os padrões visam sistematizar conjuntos de dados clínicos e definir protocolos para uniformizar as informações de saúde. Normalmente, eles são dedicados à padronização da interoperabilidade, estrutura, terminologia e à regulamentação dos dados clínicos e demográficos dos pacientes. Os tipos de padrões mais utilizados se dividem em padrões de mensagens, padrões de estrutura e conteúdo e padrão de terminologias médicas (ver Tabela 1).

Tabela 1. Normas para armazenamento e comunicação de registros de saúde. Adaptado de (Gomes, 2013; Roehrs et al., 2017)

Organização	Domínio	Padrões E-health
<i>European Committee for Standardization</i>	Modelo de informação	ISO13606/CEN/TC 251
<i>International Health Terminology Standards Development Organization</i>	Terminologia	SNOMED CT
<i>Health level 7 Integrating the Healthcare Enterprise</i>	Comunicação e arquitetura	HL7 v2.x, v3.0, CDA, FHI, IHE
<i>Digital Imaging and Communication in Medicine</i>	Imagem	DICOM
<i>Open Electronic Health Record</i>	Arquitetura PEP	OpenEHR
<i>Logical Observation Identifiers Names and Codes</i>	Banco de dados	LOINC
<i>The Continuity of Care Record</i>	Padrão de registro	CCR
<i>International Classification of Primary Care The International Classification of Diseases</i>	Modelo de classificação	ICPC ICD/ICF/ICHI
<i>The Health Insurance Portability and Accountability Act</i>	Modelo de portabilidade	HIPAA
<i>Open Archival Information System International Research on Permanent Authentic Records in Electronic Systems Preservation Metadata Implementation Strategies</i>	Padrões de preservação	OAIS, INTERPARES, PREMIS

O HL7 (*Health Level Seven*) é um padrão que reúne normas, diretrizes e metodologias internacionais para a representação e a transferência de dados clínicos e administrativos entre SIS, de modo a permitir interoperabilidade entre esses sistemas. Estas normas estão concentradas na camada de aplicação (camada 7) do modelo OSI.

A *Health Level Seven International* dá suporte a este padrão. Trata-se de uma organização internacional de normalização cujas normas também são adotadas por outros órgãos de emissão de padrões como o *American National Standards Institute* (ANSI) e a *International Organization for Standardization* (ISO).

O HL7 desenvolve padrões conceituais, padrões de documentos, normas de aplicativos e padrões de mensagens. Padrões de mensagens (HL7 versão 3) são particularmente importantes porque definem como as informações são empacotadas e comunicadas de uma parte à outra. Tais normas definem o idioma, estrutura e tipos de dados necessários para a perfeita integração de um sistema a outro. Por outro lado, o Modelo de Informação de Referência (*RIM - Reference Information Model*) e o Framework de Referência de Desenvolvimento HL7 (*HDF - Health*

Development Framework) são a base do processo de desenvolvimento dos padrões de mensagens HL7 que está na Versão 3 e que utilizam a linguagem de marcação XML. O RIM especifica como deve ser feita a representação dos dados clínicos (domínios) e o ciclo de vida de mensagens ou grupos de mensagens. Já o HDF é um projeto para especificar os processos e metodologia usados por todas as comissões de definição de padrões da organização internacional HL7 para: o início do projeto, análise de requisitos, concepção dos padrões, sua implementação, o processo de aprovação do padrão, etc.

Por outro lado, o HL7 possui a versão HL7 CDA (*Clinical Document Architecture*) que está baseado no XML para a marcação de documentos. O HL7 CDA define a estrutura e semântica de documentos clínicos, para isso utiliza o RIM e tipos de dados, de forma a facilitar seu compartilhamento e integração em um ambiente de interoperabilidade.

Por outro lado, o openEHR⁵ é um padrão de estrutura e conteúdo. Trata-se de um padrão que permite criar PEPs interoperáveis e independentes da tecnologia. Está conformado por um modelo dual, formado por um modelo de referência e um modelo de arquétipos. O modelo de referência está composto por um conjunto de classes que representam estruturas genéricas do registro eletrônico de saúde ou de dados demográficos, além de fornecerem mecanismos para controle de versão, auditoria e controle de acesso aos dados. O modelo de arquétipos está representado por metadados, usados para definir padrões que caracterizam dados clínicos cuja especificidade constitui um padrão ISO 13606-2. O suporte a este padrão está dado pela fundação openEHR que fornece um conjunto de especificações, ferramentas e recursos para lidar com a interoperabilidade.

Outro padrão é o SNOMED CT⁶ cujo foco são as terminologias médicas. Seu objetivo está na codificação de termos clínicos e mapeamento das terminologias nacionais e internacionais em uso, visando suportar a interoperabilidade semântica entre outros sistemas de codificação por exemplo o CID-10. Por outra parte, o CID⁷ (Classificação Internacional de Doenças) fornece códigos alfanuméricos relativos à classificação de doenças e de uma diversidade de sinais, sintomas, aspectos anormais, queixas, ferimentos, entre outras características de saúde. É por meio desta classificação que a OMS faz o seguimento das estatísticas de morbidade e mortalidade a nível mundial.

⁵ https://www2.openehr.org/pt/what_is_openehr

⁶ <https://www.snomed.org/snomed-ct/>

⁷ <https://icd.who.int/browse10>

Para o caso de codificação de exames laboratoriais se tem o padrão LOINC⁸ (*Logical Observation Identifiers Names and Codes*) que a diferença do CID, que registra diagnósticos, o LOINC é utilizado como um banco de dados para identificar as observações de testes laboratoriais.

O DICOM⁹ (*Digital Imaging and Communications in Medicine*) é um padrão internacional para troca de imagens médicas. O padrão permite definir, por meio de um conjunto de normas, o tratamento, armazenamento e transmissão de informação médica relativa aos exames de imagens médicas. Assim, o padrão permite que instituições que o utilizam tenham registros de saúde de seus pacientes em um formato que favorece a busca de informações e troca de documentos entre instituições (Horiil et al., 1994; Mildenerger, 2002).

O padrão DICOM SR (*Structured Reporting*) estabelece como devem ser formados, armazenados e transferidos documentos estruturados (e.g. laudos, ou qualquer tipo de observação clínica). Estes documentos contêm informações de contexto, tais como procedimentos que devem ser executados para o sucesso de um tratamento, e dados sobre profissionais de saúde envolvidos (Clunie, 2000). Um objeto no padrão DICOM SR pode conter referências embutidas a imagens, arquivos de formato de onda, como eletrocardiogramas e hemogramas, e arquivos de áudio bem como a outros documentos no mesmo padrão. O padrão estabelece métodos de referência a coordenadas destas imagens e formatos de onda, de forma integrada ao laudo (Clunie, 2000). O DICOM SR também permite incluir codificações internacionais e proprietárias para criar relatórios estruturados e padroniza a maneira de criar um documento estruturado radiológico, bem como os elementos que agrupam as informações relacionadas ao relatório.

No entanto, o ICPC¹⁰ (*The International Classification of Primary Care*) é um padrão que classifica elementos importantes no PEP. Elementos como razões de encontro, visão do paciente, processo de atendimento (decisão, ação, intervenção ou plano) e a avaliação ou diagnóstico que reflete a visão do médico.

Compartilhar dados de saúde dos pacientes organizações e pesquisadores são obrigados a cumprir com as diretrizes de privacidade especificadas na Lei de Portabilidade e Responsabilidade do Seguro de Saúde (*Health Insurance Portability and Accountability -*

⁸ <https://loinc.org/international/> (Acessado em 20/11/2019)

⁹ <https://www.dicomstandard.org/> (Acessado em 20/11/2019)

¹⁰ <https://www.globalfamilydoctor.com/groups/WorkingParties/wicc.aspx> (Acessado em 20/11/2019)

HIPAA). O HIPAA delinea duas abordagens para proteger informações de saúde individualmente identificáveis. As regras estabelecem normas para proteger os registros pessoais de saúde, salvaguardando a privacidade das informações pessoais nas transações por via eletrônica, como também estabelecem limites e condições sobre o uso e divulgação que podem ser feitos de tais informações sem a autorização do paciente. HIPAA também dá aos pacientes os direitos sobre suas informações de saúde, incluindo os direitos de examinar e obter uma cópia dos seus registros de saúde e solicitar correções.

Finalmente, existem estratégias e iniciativas para preservação de dados, não exclusivamente focados no domínio da saúde. Estes são o OAIS, INTERPARES e PREMIS. O OAIS, como já mencionado na seção 2.1.2, possui um esquema conceitual que orienta a sistemas de arquivos para a preservação de informação a longo prazo. Já o INTERPARES, cuja base conceitual tem origem no modelo OAIS, caracteriza-se por definir requisitos funcionais para os sistemas informatizados de gestão de documentos para preservação (Duranti, 2008). Por outro lado, o PREMIS é um padrão de metadados e baseia-se no XML. Este padrão é considerado como uma tradução semântica em unidades de implementação do modelo OAIS (Guenther, 2010).

2.2.2 PADRÃO EN ISO 13940

A ISO 13940¹¹ é um padrão europeu que está focado em conceitos que representam todos os processos de assistência médica (i.e., atividade profissional de saúde, atendimento social como autocuidado, cuidado de terceiros). Para poder representar o conteúdo clínico e o contexto clínico a ISO 13940 está relacionada a um modelo genérico de processos de assistência médica, bem como a definições abrangentes de conceitos e modelos conceituais para os aspectos clínicos, gerenciais e de recursos dos serviços de saúde. Esta norma divide-se em oito modelos que representam o contexto dos serviços de saúde e são descritos brevemente a seguir:

- **Healthcare actors:** Representa a uma organização ou pessoa que participa dos cuidados de saúde. Portanto, todo profissional de saúde de uma instituição é um agente de saúde com vários graus de responsabilidade no gerenciamento de cuidados dos pacientes.
- **Healthcare matters:** Representa ao assunto de assistência médica que pode ser uma doença ou diversos tipos/condições de doenças.
- **Healthcare activities:** Representa as atividades destinadas direta ou indiretamente ao tratamento e melhora do estado de saúde do paciente.

¹¹ <https://www.iso.org/obp/ui/#iso:std:iso:13940:ed-1:v1:en> (Acessado em 20/11/2019)

- **Healthcare process:** Refere-se ao conjunto de processos de assistência médica inter-relacionadas com a especificação de um ou mais problemas de saúde.
- **Healthcare planning:** Representa o gerenciamento de atividades de saúde em que um plano de assistência é criado ou modificado.
- **Time:** Representa o intervalo de tempo durante o qual as atividades de saúde são realizadas para os cuidados do paciente.
- **Responsibilities:** São mandados com base em um compromisso e consentimento informado ou autorização por lei, que define os direitos e obrigações de um ator da área de saúde em relação ao seu envolvimento nos processos de saúde realizados para um assunto específico de atendimento.
- **Information management:** Representa o gerenciamento do PEP. Todas as informações (ex. registros médicos, registros dentários, de assistência clínicas) são armazenadas numa mídia eletrônica.

2.2.3 NORMA ISO 13606

A norma ISO 13606¹² é um modelo dual de arquitetura de PEP. Está composto por estruturas de dados, conhecidas como arquétipos (semelhante à proposta do openEHR), os quais representam os conceitos clínicos aceitos dentro da comunidade com o objetivo de garantia de interoperabilidade, consistência e qualidade da informação (Kalra, 2006; Muñoz et al., 2011). Seu modelo de referência representa as características gerais de componentes de um PEP, como eles são agregados, e o requerido contexto de informação para reunir os requisitos éticos, legais e de originalidade (Muñoz et al., 2011). Podemos identificar as seguintes características:

- Está conformado por um conjunto de tipos primitivos;
- Suas entidades principais são: *EHR_extract*, *folders*, *compositions*, *sections*, *entries*, *cluster* e *elements*;
- Conta com classes auxiliares e, opcionalmente;
- Classes de descrição de dados demográficos e de comunicação.

A Figura 7 mostra uma visão geral do modelo de referência da ISO 13606, logo são descritas suas principais entidades.

¹² <http://www.en13606.org/> (Acessado em 20/11/2019)

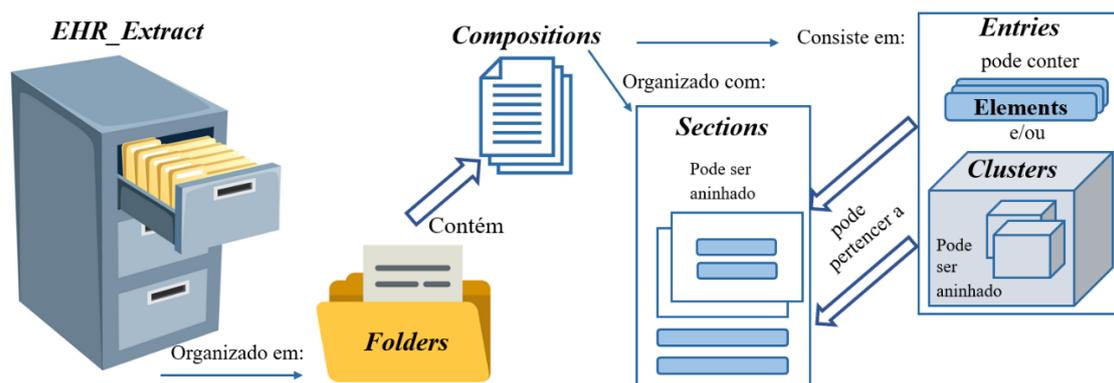


Figura 7. Relações dos componentes do Modelo de Referência ISO 13606. Adaptado de (Muñoz et al., 2011)

- **Ehr_Extract:** é a representação de uma parte ou toda a informação do prontuário eletrônico do paciente.
- **Folder:** refere-se à entidade que organiza o PEP ao primeiro nível e é geralmente associada a eventos clínicos ou especialidades médicas.
- **Composition:** representa um documento do PEP, como um relatório ou um exame, que geralmente está relacionado a uma sessão clínica. As composições são organizadas em pastas (*Folder*).
- **Section:** dentro de uma composição, são as entradas clínicas durante um encontro clínico e que formam um documento clínico (*Composition*) são organizados em seções ou títulos que facilitam o registro e a apresentação da informação.
- **Entry:** é uma entrada ou declaração clínica. É a unidade de registro de informações em um PEP. Ele pode corresponder, por exemplo, a uma observação, um pedido de procedimento ou um diagnóstico.
- **Cluster:** é a estrutura de dados que organiza as informações em séries, tabelas ou árvores.
- **Element:** representa os nodos, que são os valores que contém um único valor de um tipo de dados.

2.3. CURADORIA DIGITAL

Historicamente, as práticas de curadoria apareceram no início dos anos 90 focadas quase completamente nos dados, chamados curadoria de dados. Assim o campo de curadoria cresceu e a curadoria digital tornou-se reconhecida como uma prática generalizada, selecionando recursos que variam de objetos de dados a conteúdo multimídia (Rusbridge et al., 2005; Abbott, 2014).

Segundo Márdero (2008), a curadoria digital é apresentada como uma resposta ao crescimento de dados digitais de diferentes tipos de fontes. Ele evolui das noções de preservação digital e da necessidade de distribuir e gerenciar dados digitais desde o início, garantindo sua preservação, descoberta, interpretação e reutilização a longo prazo. Em suma, a curadoria digital é o conjunto de atividades e processos necessários para garantir a preservação das informações digitais ao longo do tempo. A curadoria de dados também considera e aborda funções de nível superior, como gerenciamento de repositório, preservação digital, governança e financiamento.

Na área da saúde, os dados médicos são cada vez mais dispersos e há uma redundância nas solicitações, por exemplo, testes para diagnóstico médico. Todos os testes realizados anteriormente pelo paciente, além de seus dados médicos, devem ser integrados e sempre disponíveis. A implantação das tarefas de curadoria poderia ajudar no gerenciamento dessas informações, porém deve-se tomar as medidas de proteção necessárias já que são considerados como dados sensíveis (Li et al., 2018).

2.3.1. MODELO DO CICLO DE VIDA DA CURADORIA DIGITAL

Digital Curation Center (DCC)¹³ é um centro de especialização reconhecido internacionalmente em curadoria digital, com foco na criação de capacidades e habilidades para gerenciamento de dados de pesquisa. O DCC propôs um Modelo de Ciclo de Vida de Curadoria (Higgins, 2008) que fornece uma visão geral gráfica de alto nível das etapas necessárias para uma curadoria e preservação bem-sucedidas de dados desde sua conceituação inicial. A Figura 8 apresenta as etapas e ações de curadoria definidas por este modelo.

Diferentemente de outros modelos de curadoria, o modelo de ciclo de vida do DCC fornece uma visão não linear da curadoria digital. O elemento chave deste modelo são os dados, que representam qualquer informação em formato digital binário, incluindo:

- **Objetos digitais** (*Digital Objects*): objetos digitais simples (por exemplo, arquivos de texto, arquivos de imagem ou arquivos de som, juntamente com seus identificadores e metadados relacionados) ou objetos digitais complexos (por exemplo, objetos digitais discretos criados pela combinação de vários outros objetos digitais, como sites)
- **Bancos de dados** (*Databases*): coleções estruturadas de registros ou dados armazenados em um sistema de computador

¹³ <http://www.dcc.ac.uk/>

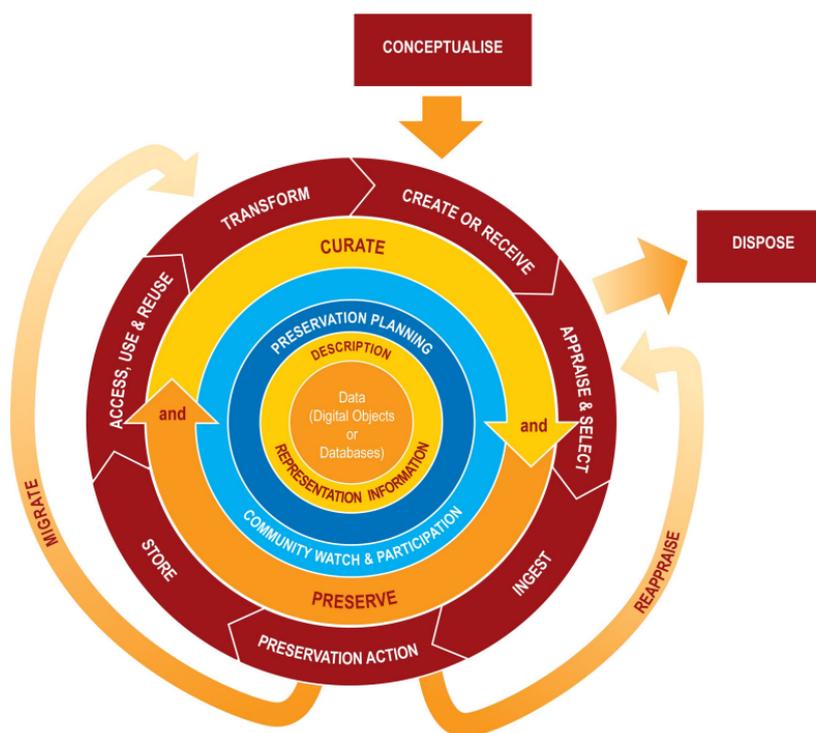


Figure 8. DCC Modelo de Ciclo de Vida (Fonte: Higgins, 2008)

Observando a Figura 8, podemos ver vários anéis concêntricos em sua estrutura. As ações do ciclo de vida completo são representadas por informações de descrição e representação, planejamento de preservação, vigilância e participação da comunidade e organização e preservação.

- **Descrição e representação da informação** (*Description and representation information*): Trata-se do gerenciamento, descrição, técnicas, estrutura e preservação dos metadados por meio da aplicação de padrões para garantir um adequado controle e descrição ao longo prazo. Também, coleta e atribui uma representação da informação necessária para compreender e processar tanto o material digital quanto aos metadados associados.
- **Plano de preservação** (*Preservation planning*): É o plano de preservação ao longo do ciclo de vida de um material digital. Inclui os planos para gerenciamento e administração de todas as ações do ciclo de vida de curadoria.
- **Acompanhamento e participação da comunidade** (*Community watch & participation*): Trata da necessidade de atenção para as atividades que se desenrolam no âmbito das comunidades envolvidas com o problema de curadoria, bem como a necessidade de participação no desenvolvimento de padrões, de ferramentas e de software adequados ao problema e que possam também serem compartilhados.

- **Preservação e curadoria** (*Preserve and curate*): Estar continuamente alerta e empreender as ações administrativas e gerenciais planejadas para a curadoria e preservação por todo o ciclo de vida da curadoria.

Em relação as ações sequenciais no anel externo, estes representam as principais ações necessárias para selecionar dados, desde a criação até o uso e reutilização. Além disso, essas ações não são realizadas apenas uma vez. Eles podem ser repetidos enquanto os objetos digitais estiverem sendo selecionados. As ações sequenciais não se referem exclusivamente à curadoria, mas representam estágios do ciclo de vida dos dados que devem ter um componente de curadoria.

- **Conceitualização** (*Conceptualise*): Conceber e planejar a criação do dado, incluindo os métodos de captura e as opções de armazenamento;
- **Criação e/ou recebimento** (*Create or receive*): Refere-se a criar dados, incluindo metadados administrativos, descritivos, estruturais e técnicos. Metadados de preservação podem, também, serem adicionados no momento da criação. Receber dados, de acordo com políticas bem documentadas, de criadores de dados, de outros arquivos, de repositórios ou de centros de dados e, se necessário atribuir metadados apropriados.
- **Avaliação e seleção** (*Appraise & Select*): Avaliar o dado e selecionar o que será objeto dos processos de curadoria e de preservação por longo prazo. Manter-se aderente tanto às boas práticas quanto às políticas pertinentes e também às exigências legais.
- **Arquivamento** (*Ingest*): Avaliar e transferir os dados para um arquivo, ou repositório, ou centro de dados ou outro custodiante apropriado.
- **Ações de preservação** (*Preservation action*): Sugere empreender ações para garantir a preservação a longo prazo e a retenção do dado de natureza oficial. Ações de preservação devem garantir que o dado permaneça autêntico, confiável e utilizável, mantendo a sua integridade.
- **Armazenamento** (*Store*): Armazenar os dados de forma segura mantendo em conformidade aos padrões relevantes.
- **Ações de uso e reuso** (*Access, use & refuse*): Está relacionado à ação de garantir que o dado possa ser acessado tanto pela sua comunidade alvo, quanto pelos demais usuários interessados na reutilização do dado.
- **Transformação** (*Transform*): Criação de novos dados a partir do original.

Por último temos o conjunto final de ações que é representado por “ações ocasionais”, que podem ocorrer diante de condições específicas. Além disso, essas ações não são necessárias para todos os dados.

- **Eliminação/descarte** (*Dispose*): Descarte os dados que não foram selecionados para curadoria e preservação a longo prazo, de acordo com políticas, orientações ou requisitos legais documentados. Normalmente, os dados podem ser transferidos para outro arquivo, repositório, centro de dados ou outro custodiante. Em alguns casos, os dados são destruídos. A natureza dos dados pode, por razões legais, exigir destruição segura.
- **Reavaliação** (*Reappraise*): Retorna dados que falhem nos procedimentos de validação para posterior avaliação e re-seleção.
- **Migração** (*Migrate*): Migrar os dados para um formato diferente.

Este modelo pode ser usado para planejar atividades dentro de uma organização ou consórcio de forma a garantir que todos os estágios necessários sejam realizados, cada um na sequência correta. Segundo Higgins (2008), o modelo permite que a funcionalidade granular seja mapeada; definir papéis e responsabilidades; e construir uma estrutura de padrões e tecnologias para implementar. Pode ajudar no processo de identificação de etapas adicionais que podem ser necessárias ou ações que não são necessárias em determinadas situações ou disciplinas e garantir que os processos e políticas sejam adequadamente documentados.

Vale ressaltar que existem outros modelos como: o DataONE que é uma base que visa apoiar à ciência, envolvendo as comunidades relevantes de ciência, dados e políticas, além de fornecer armazenamento fácil, seguro e persistente de dados (Allard, 2012). A preservação dos dados é realizada pela disseminação de ferramentas integradas e fáceis de usar para descoberta, análise, visualização e tomada de decisão de dados (Allard, 2012). Além dessas ferramentas, o DataONE também fornece um modelo de ciclo de vida de dados como um auxílio para gerenciar e preservar dados.

Outra iniciativa focada no arquivamento de dados é a DDI (*Data Documentation Initiative*) que é um padrão para dados de ciências sociais que combina requisitos de metadados e requisitos de empacotamento (DDI, 2012).

O DDI é um modelo de curadoria mais linear para dados de pesquisa que representa o ciclo de vida completo dos dados. Este modelo considera que os metadados devem ser mantidos

e versionados ao longo do tempo para suportar a comparação entre estudos em uma coleta de dados maior (por exemplo, estruturas geográficas, locais geográficos, conceitos, hierarquias do universo e organizações e indivíduos).

2.4. ENGENHARIA DIRIGIDA POR MODELOS

A Engenharia Dirigida por Modelos (*Model-Driven Engineering* - MDE) considera a utilização de modelos como os principais artefatos no desenvolvimento de aplicações e sistemas. De acordo a esta abordagem, os modelos não somente descrevem ou documentam um software, mas também servem para atuar no seu desenvolvimento, manutenção e operação (Seidewitz, 2003; Schmidt, 2006). As técnicas de MDE, tais como Desenvolvimento Dirigido por Modelos (*Model-Driven Development* - MDD) e Arquitetura Dirigida por Modelos (*Model-Driven Architecture* - MDA) propõem o uso de abstrações mais próximas do domínio do problema como uma forma de mitigar a distância semântica existente entre o problema a ser solucionado e o ferramental (software) utilizado para tal. Um aspecto importante da MDE é a ênfase na integração entre a parte tecnológica e o conhecimento específico de um determinado domínio (Favre, 2004; Favre and NGuyen, 2005). A utilização de seus princípios aumenta a qualidade dos sistemas de software, o grau de reuso e, como resultado implícito, a eficiência em seu desenvolvimento (Bézivin, 2005).

2.4.1 METAMODELOS

Segundo Guedes (2018) um modelo é capaz de capturar a visão de um sistema, representa uma abstração com um certo propósito, como por exemplo descrever aspectos estruturais ou comportamentais de um sistema. Este propósito determina o que deve ser incluído no modelo e o que é irrelevante no nível apropriado de detalhe.

Entretanto, um metamodelo define uma linguagem para expressar modelos, ou seja, é um modelo de modelos e deve definir as construções de modelagem e seus relacionamentos (Völter, 2013). O papel de um metamodelo é definir a semântica para modelar elementos dentro de um modelo e seus relacionamentos, os quais devem ser definidos na forma da sintaxe abstrata. A semântica pode ser estática, que define regras e restrições que garantem o uso correto dos elementos do metamodelo, ou dinâmica que define o comportamento associado aos elementos do metamodelo (Völter, 2013).

Assim, os metamodelos apresentam um potencial relevante para lidar de forma adequada com esses problemas pelo fato de oferecerem um nível mais alto de abstração no

desenvolvimento de software (OMG, 2004). Tal característica proporciona reusabilidade, portabilidade e interoperabilidade, aumentando a produtividade do processo e a qualidade das aplicações. Porém, dada a popularidade do uso de modelos, surge a necessidade de uma padronização para a construção de metamodelos e modelos. Dessa forma, a *Object Management Group* (OMG) apresentou uma arquitetura de metamodelagem de quatro camadas, denominada *Meta-Object Facility* (MOF)¹⁴. Na MOF, cada elemento de uma camada inferior é uma instância de um elemento de uma camada superior (Seidewitz, 2003), conforme mostrado na Figura 9.

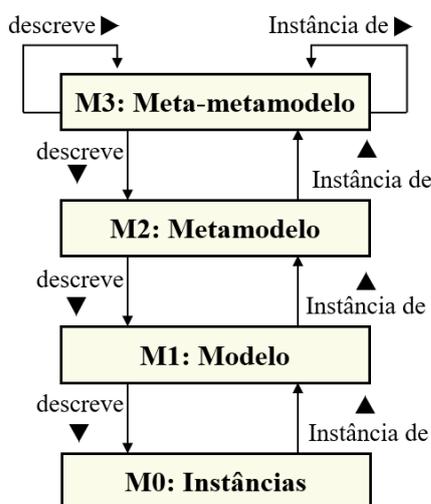


Figura 9. Camadas do metamodelo MOF

- **Camada M3:** representa o meta-metamodelo da MOF utilizado para construção de metamodelos. O modelo MOF formaliza suas próprias abstrações eliminando a necessidade de um nível superior (Steinberg et al., 2008).
- **Camada M2:** nesta camada estão os metamodelos que podem ser utilizados para modelar sistemas de domínio específico, ou seja, comporta uma instância do meta-metamodelo. Um exemplo de linguagem definida conforme o modelo MOF é a Linguagem de Modelagem Unificada (*Unified Modeling Language* - UML).
- **Camada M1:** composta por modelos que descrevem sistemas utilizando as definições presentes em seus respectivos metamodelos presentes em M2.
- **Camada M0:** contém as entidades ou objetos que formam o sistema em tempo de execução que são criadas a partir das definições presentes em M1.

¹⁴ <http://www.omg.org/mof>

A Figura 10 mostra um exemplo de metamodelo na camada 2 para abstrair as entidades classe e atributo do UML, para isso são criadas duas metaclasses chamadas *Classe* e *Atributo* relacionadas de 1 a muitos (uma classe pode ter muitos atributos). Já no nível 1 é criada uma instância deste metamodelo por meio da classe *Pessoa* com atributo *idade* do tipo *int*. Por fim no nível 0, é criada uma instância da classe *Pessoa*, o objeto *João* com atributo *idade=25*.

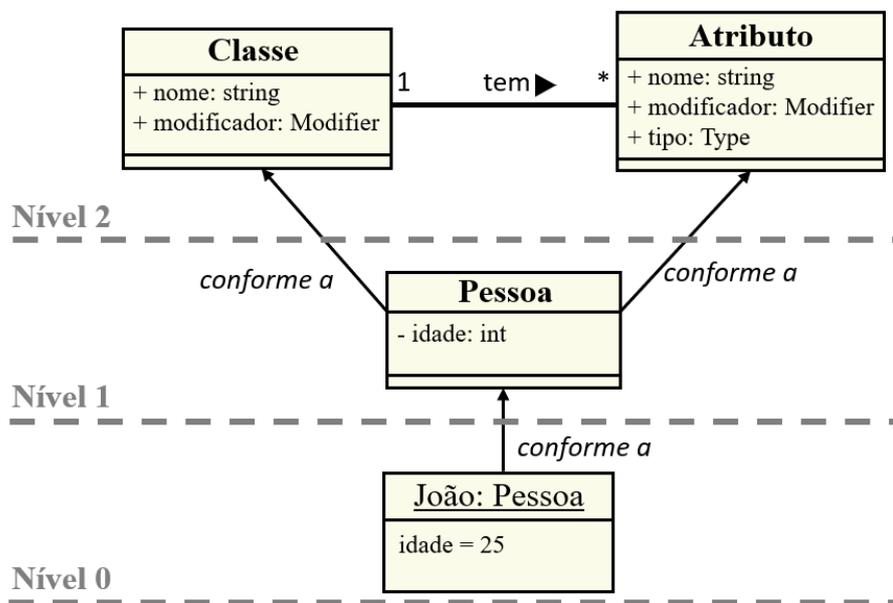


Figura 10. Exemplo de metamodelo de nível 2 e seus modelos derivados

Em resumo, os metamodelos no Nível 2 são utilizados para definir os modelos no Nível 1. Assim, a partir da linguagem de definição (por exemplo a UML) pode-se definir diagramas de classes e diagramas de objetos. Nesse sentido, o metamodelo para preservação de dados clínicos proposto nesta tese e apresentado no Capítulo 4, foi construído usando o EMF (*Eclipse Modeling Framework*) que é uma instância do meta-metamodelo *Ecore*. Assim, com o metamodelo definido usando o *Ecore*, os modelos instanciados podem ser gerados a partir de uma API e portados para XMI (*XML Metadata Interchange*)¹⁵ ou terem código-fonte gerados em Java. Desta forma, os modelos do Nível 1 podem ser instanciados e validados, ou de forma programática ou por meio de um editor, tendo como referência o metamodelo da aplicação construído.

2.4.2 ECLIPSE MODELING FRAMEWORK

O *Eclipse Modeling Framework* (EMF) é um arcabouço de modelagem construído sobre o ambiente de desenvolvimento integrado (*Integrated Development Environment - IDE*) Eclipse.

¹⁵ <https://www.omg.org/spec/XMI>

O EMF fornece mecanismos para a criação, edição e validação de modelos e metamodelos, além de permitir a geração de código a partir dos modelos. Para tal, o EMF possibilita a geração de uma implementação em linguagem Java, de maneira que cada uma das classes do metamodelo (chamadas de metaclasses) corresponde a uma classe em Java. Dessa forma, estas classes podem ser instanciadas para criar modelos em conformidade com o metamodelo. O EMF também possibilita criar ferramentas de modelagem (editores) que permitem criar modelos a partir de seus metamodelos.

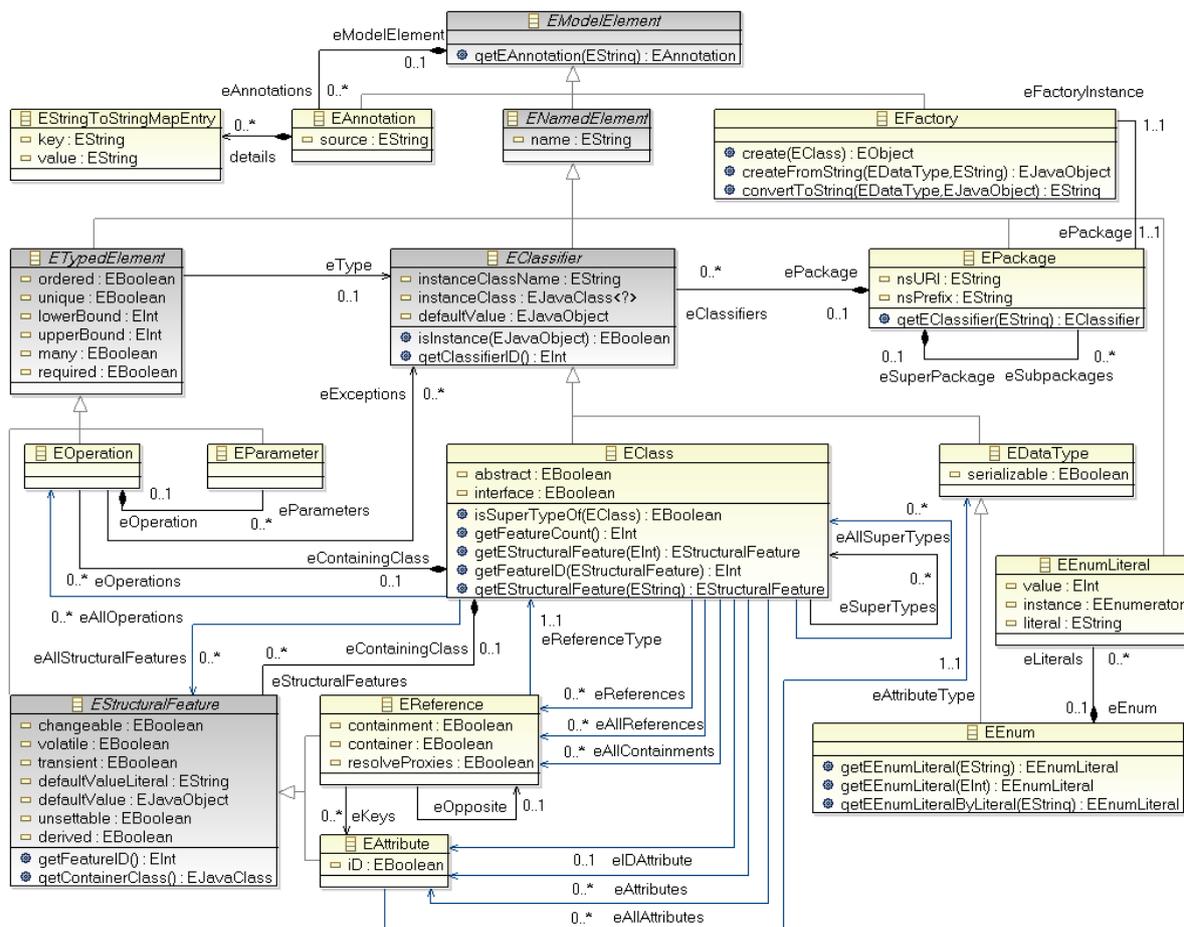


Figura 11. Meta-metamodelo *Ecore*

Por outro lado, os metamodelos construídos no EMF são instâncias do meta-metamodelo *Ecore* que, por sua vez, é baseado no meta-metamodelo MOF. O *Ecore* pode ser utilizado para construir modelos que representam o metamodelo de uma linguagem. Sendo assim, um metamodelo com base no *Ecore* é definido por meio de instâncias de classes do tipo: *EClass*, *EAttribute*, *EReference*, *ESuperType* e *EDataType*. Uma *EClass* representa uma classe composta por atributos e referências. Um *EAttribute* é um atributo que possui um nome e um tipo. Uma *EReference* define uma associação entre classes e, no caso da definição de uma

superclasse, para se valer dos conceitos de herança, essa associação é do tipo *ESuperType*. Por fim, um *EDataType* é o tipo de um atributo, cujo valor pode ser um tipo primitivo (número inteiro ou real, string, booleano, etc.), uma enumeração *EEnum* ou uma referência a uma *EClass*. A Figura 11 mostra os componentes do meta-metamodelo *Ecore* (Merks and James, 2009), com seus atributos, operações e relacionamentos.

2.5. CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados conceitos básicos relacionados à preservação digital, modelos e estratégias. O modelo de referência OAIS é um dos mais representativos na literatura e é amplamente utilizado para implantação de estratégias de preservação de documentos digitais (Korb and Strodl, 2010; Askhoj et al., 2011; Voinov et al., 2017). Seguidamente, se descreveram as características dos PEPs e seus padrões de interoperabilidade relacionados como por exemplo HL7 (que é um conjunto de padrões para intercâmbio de informação clínica) e o padrão DICOM (que permite a transferência e arquivamento de imagens médicas). Logo a seguir foram descritos as normas ISO 13940 e ISSO 13606. O primeiro aborda os conceitos que representam os processos clínicos e o segundo é um modelo de referência que inclui um conjunto de classes que definem os blocos de construção genéricos dos PEPs. A norma ISO 13606 também detalha como os dados clínicos devem ser agregados para criar estruturas de dados mais complexas e as informações de contexto que devem acompanhar todos os dados para atender aos requisitos éticos e legais (Muñoz et al., 2011).

Outro tema abordado é a curadoria digital que de acordo à literatura é a encarregada da gestão dos documentos digitais e sua conceptualização evolui das noções de preservação digital e da necessidade de estabelecer um conjunto de processos de forma a garantir a preservação, descoberta, interpretação e reutilização da informação a longo prazo. Nesse sentido, a curadoria digital é importante porque ajuda a estabelecer planos de preservação. Em nosso domínio, o papel do curador é importante para definir quais serão os dados e metadados do PEP a serem preservados e ajudará a definir as políticas de acesso.

Por fim, neste Capítulo abordamos os conceitos sobre a Engenharia Dirigida por Modelos, cujo foco é a utilização de modelos como principais artefatos no desenvolvimento de sistemas. Vimos que os modelos não são apenas artefatos de documentação utilizados para auxiliar na atividade de desenvolvimento, são também parte integrante do software, pois permite que o arquiteto se concentre em modelos de alto nível ou metamodelos de forma a se proteger da complexidade inerente à implementação nas diferentes plataformas.

CAPÍTULO 3 – REVISÃO DA LITERATURA

Neste capítulo apresentamos uma revisão sistemática da literatura (janeiro 2010 até outubro 2019) seguindo como guia as recomendações dos autores Kitchenham and Charters (2007) e Jalali and Wohlin (2012) para proporcionar uma visão abrangente e válida das evidências existentes sobre metamodelos de preservação de PEPs.

A metodologia de revisão adotada neste estudo consiste em um protocolo dividido em três etapas: planejamento, condução e relatório (ver Figura 12). Durante a fase de planejamento, questões de pesquisa foram formuladas em conformidade aos objetivos deste trabalho. Em seguida, um conjunto de palavras-chave foi selecionado para realizar a pesquisa em diferentes fontes de dados.

Durante a fase de condução da pesquisa, foram utilizados critérios de inclusão e exclusão para remover estudos irrelevantes. Os trabalhos selecionados foram organizados e sintetizados de acordo com um conjunto de características. Na fase de relatório, as respostas para todas as questões de pesquisa foram descritas. Após isso, todas as informações coletadas foram classificadas e mapeadas. Finalmente, foram identificadas as tendências e limitações atuais no estado da arte, o que nos levou a propor um metamodelo que será descrito nos próximos capítulos.

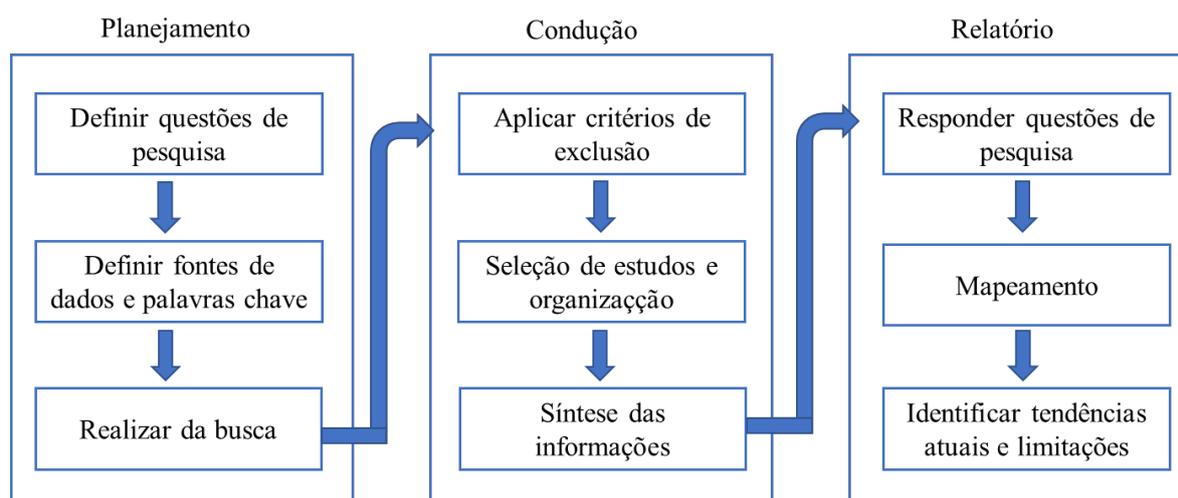


Figura 12. Metodologia da revisão sistemática da literatura. Adaptado de Malhotra and Chug (2016)

3.1 PLANEJAMENTO DA REVISÃO DE LITERATURA

3.1.1 QUESTÕES DE PESQUISA

O principal objetivo desta revisão sistemática da literatura é realizar um mapeamento no campo da Preservação de Dados Clínicos. Para isso, foi definida a seguinte questão de pesquisa geral: *Qual é o estado da arte atual em relação à preservação de dados clínicos?*

Para responder a essa questão principal de pesquisa, definimos seis questões específicas (listadas na Tabela 2), para orientar e estruturar a extração e análise de dados de pesquisa e a síntese de evidências:

Tabela 2. Lista de questões de pesquisa

No.	Questões de pesquisa
QP1	Qual é a evolução das pesquisas relacionadas à preservação de PEP durante os últimos anos?
QP2	Quais estratégias são utilizadas na preservação digital dos PEPs?
QP3	Qual é o papel dos metadados na preservação digital de dados de saúde?
QP4	Que tipos de modelos de referência são usados para representar a preservação digital de PEPs?
QP5	Quais são os principais papéis dos atores nos estudos de preservação em PEP?
QP6	Qual é o papel da curadoria de dados nos estudos de preservação do PEP?

3.1.2 BANCO DE DADOS E ESTRATÉGIAS DE PESQUISA

Para realizar a pesquisa foram utilizados como fonte de dados os maiores repositórios bibliográficos como: ACM, IEEEExplore, PubMed, Springer e também realizamos a busca na Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) e Google Scholar para coletar artigos que não foram recuperados nas outras fontes.

A estratégia de busca foi definida com base no uso palavras-chave como “preservação digital” e “prontuário eletrônico do paciente” e seus sinônimos e abreviações. Também foram utilizadas as palavras-chave de “curadoria de dados” e “metamodelo” com suas respectivas variações, para recuperar trabalhos que tratem da curadoria digital e modelo de referência para preservação digital.

Uma vez que as palavras-chave foram selecionadas, foram criadas cadeias de pesquisas (*string query*). Estas cadeias estão formadas pela composição de palavras-chave mediante

operadores lógicos AND (e lógico) e OR (ou lógico) e o uso de parênteses para agrupação. A seguir se mostra estas:

E1: *((long-term OR metadata OR data OR lifelong) AND "digital preservation")*

E2: *(("electronic health record" OR EHR OR "electronic healthcare record" OR "personal healthcare record" OR "electronic medical record" OR EMR OR "electronic patient record" OR EPR OR "clinical data" OR "medical data" OR "health data") AND (preservation OR archival))*

E3: *(("meta-model" OR "metamodel" OR "ontology" OR "model") AND "digital preservation")*

E4: *(("curation" OR "digital curation" OR "metadata curation" OR "data curation") AND "digital preservation")*

A seguir se mostra a cadeia de pesquisa final:

T = E1 **OR** E2 **OR** E3 **OR** E4

É importante mencionar que as pesquisas foram realizadas na língua inglesa, a exceção das fontes de dados locais, onde também foram utilizadas palavras chave na língua portuguesa.

3.1.3 CRITÉRIOS DE INCLUSÃO E EXCLUSÃO

Os critérios de inclusão foram estabelecidos com base na busca de conteúdos razoáveis e relevantes no contexto da preservação dos dados no âmbito da saúde. Os trabalhos que foram selecionados são dos últimos dez anos e devem apresentar conteúdos relacionados à preservação de prontuários eletrônicos, modelos ou estratégias para preservação de dados clínicos. A seguir são listados os critérios de inclusão (CI) que foram considerados durante a pesquisa.

- CI1: Estudos empíricos sobre preservação de dados ou metadados relacionados à saúde.
- CI2: Estudos que descrevem um caso de estudo de preservação de dados clínicos.
- CI3: Estudos que propõem um metamodelo ou modelo de referência para preservação de dados clínicos.

Além disso, um conjunto de critérios de exclusão (CE) foi usado para remover artigos irrelevantes da pesquisa. Os critérios de exclusão elaborados foram os seguintes:

- CE1: O título e o resumo do estudo não apresentam nenhum tipo de constatação sobre o tema da pesquisa;
- CE2: Estudos contendo menos de duas páginas;
- CE3: Estudos publicados antes de 2010.
- CE4: Estudos de revisão/*survey*
- CE5: Estudos duplicados ou artigos estendidos do mesmo autor.

3.2 CONDUÇÃO DA PESQUISA

3.2.1 SELEÇÃO E ORGANIZAÇÃO DOS ESTUDOS

Após realizar a pesquisa nos bancos de dados bibliográficos listados na seção 3.1.2, de um total de 2741 estudos, inicialmente 88 estudos potencialmente relevantes foram selecionados. Isto foi realizado com base na leitura do título, palavras chave e caso necessário a leitura do resumo do estudo.

Os 88 estudos pré-selecionados foram organizados cronologicamente e por biblioteca. Também foram utilizados os critérios de inclusão e exclusão nos textos completos destes estudos. Com isso, na etapa de filtragem muitos destes estudos recuperados foram eliminados, por exemplo, estudos duplicados (publicados em mais de uma biblioteca acadêmica), estudos que não apresentaram uma abordagem de preservação para dados relacionados à saúde.

Tabela 3. Resultados da pesquisa considerando os critérios de inclusão e exclusão

Fonte de dados	Estudos recuperados	Estudos Excluídos	Estudos Incluídos
ACM	233	229	4
BDTD	36	30	6
IEEE	192	189	3
PubMed	679	675	4
Springer	1592	1587	5
Google scholar	9	0	9
Total	2741	2710	31

A Tabela 3 mostra um resumo de todos os estudos recuperados utilizando os diferentes mecanismos de pesquisa. Considerando o total de 2741 estudos, finalmente foram selecionados

31 estudos como sendo os mais relevantes e que estão em conformidade aos objetivos deste capítulo.

3.2.2 SÍNTESE DA INFORMAÇÃO

Cada estudo foi analisado cuidadosamente pelo autor desta tese, extraindo suas características mais relevantes de acordo com as questões de pesquisa definidas na Tabela 2. Todas as informações coletadas foram registradas em planilhas do Microsoft Excel. Os dados extraídos de cada estudo foram organizados e sintetizados de forma estruturada com a ajuda de tabelas e mapas mentais. A Tabela 4 apresenta as referências dos estudos que foram selecionados nesta etapa e são dispostos em ordem cronológica.

Tabela 4. Lista dos estudos selecionados

ID estudo	Autor	ID estudo	Autor
S1	Santos et al., (2010)	S17	Bunawan and Nordin (2015)
S2	Joemon et al. (2010)	S18	Celesti et al. (2016)
S3	Yina (2010)	S19	Celesti et al. (2017)
S4	Bote et al. (2011)	S20	Lo et al. (2017)
S5	Cruz et al. (2011)	S21	Pinto and Sales (2017)
S6	Kiefer et al. (2012)	S22	Roehrs et al. (2017)
S7	Misra et al. (2012)	S23	Sassi et al. (2017)
S8	Rudersdorf (2012)	S24	Abouelmehdi et al. (2018)
S9	Bote et al. (2012)	S25	Chaczko et al. (2018)
S10	Gomes (2013)	S26	Gordon (2018)
S11	Piconi et al. (2013)	S27	Li et al. (2018)
S12	Goossen (2014)	S28	Limo et al. (2019)
S13	Malone et al. (2014)	S29	Oliveira (2019)
S14	Piconi (2014)	S30	Ou et al. (2019)
S15	Graña and Jackwoski (2015)	S31	Zhu et al. (2019)
S16	Gonçalves et al. (2015)		

3.3 RELATÓRIO DA REVISÃO

QP1: Qual é a evolução das pesquisas relacionadas à preservação de PEP durante os últimos anos?

Nesta revisão, analisamos a evolução do número de estudos relacionados à preservação de dados relacionados à saúde publicados no período de 2010-2019. O paradigma da preservação digital surgiu oficialmente no final de 1996 (Waters and Garret, 1996) com iniciativas governamentais realizadas por pesquisadores da área da Ciência da Informação. Porém, é a partir da década de 2000 que recém é possível encontrar estudos que discutem a relevância da preservação digital (Araujo et al., 2013). É assim que o Comitê Consultivo para Sistemas de Dados Espaciais (CCSDS) iniciou um esforço conjunto para desenvolver um conjunto de padrões capazes de regular o armazenamento a longo prazo de informações digitais produzidas em missões espaciais. A partir desse esforço é que nasceu o modelo de referência OAIS (descrito no Capítulo 2). O OAIS é um modelo de referência que visa identificar os componentes funcionais que devem fazer parte de um sistema de informação dedicado à preservação digital (CCSDS, 2012; Lavoie, 2014). O OAIS tornou-se a ISO 14721:2003 e posteriormente atualizada para a ISO 14721: 2012 (ISO, 2012), que é um modelo padrão para sistemas de preservação digital em muitas instituições e organizações.

A Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) liderou a questão sobre a preservação digital e lançou no 2003 a “*Charter on the Preservation of the Digital Heritage*” e, no 2007 o “*Towards an Open Source Repository and Preservation System*” (Bradley et al., 2007) que teve como foco ações de preservação do patrimônio digital, as quais devem ser mantidas para as gerações futuras e possuem as seguintes características (Misra et al., 2012): conteúdo original totalmente gerado e apresentado digitalmente e sem equivalente analógico; dependente de hardware, software, mídia e formatos de armazenamento. Por outro lado, na mesma década, surge a proposta do Modelo de Ciclo de Vida de Curadoria do DCC (Higgins, 2008; Abrams. et. al., 2008; Sayão and Sales, 2013), que propõe um ciclo de vida completo com atividades definidas por: descrição e representação da informação, planejamento da preservação, participação da comunidade alvo e a curadoria e preservação.

A Figura 13 mostra um gráfico de distribuição em relação ao número de publicações nos últimos anos. De acordo com os parâmetros de nossa pesquisa, consideramos um total de 31 estudos. Esse número em um tópico tão importante como este pode ser um número muito

pequeno, mas é limitado a estudos que apresentam a abordagem relacionada à preservação de dados clínicos ou à preservação do PEP.

Observamos que não existe um padrão claro de evolução no estado da arte da preservação de dados clínicos, pois está sujeito à adoção de um ou mais modelos de referência [S09] [S15] ou à proposta de um [S11] [S13] [S21]. Vale ressaltar que estudos recentes [S27] [S31], além de abordar a preservação de dados, abordam a privacidade e a segurança.

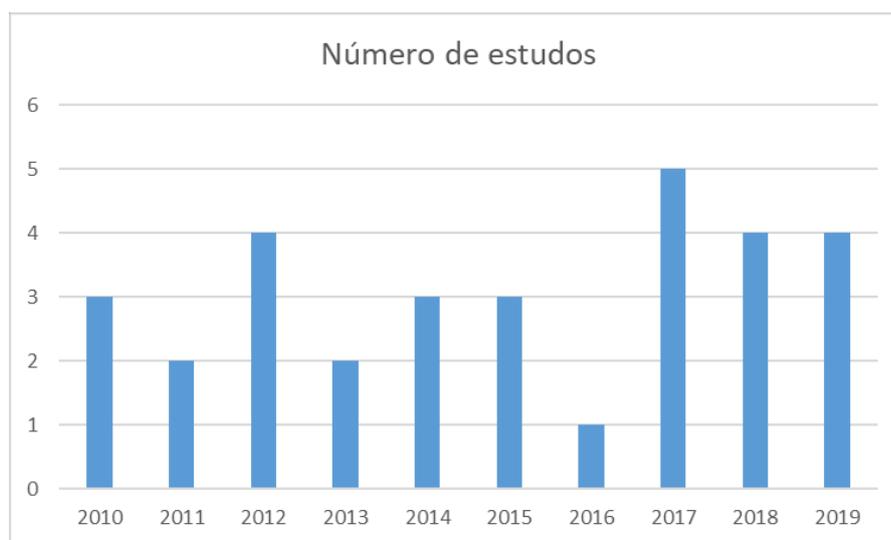


Figura 13. Distribuição dos artigos por ano, sobre a preservação de dados clínicos.
(Fonte: O autor)

QP2: Quais estratégias são utilizadas na preservação digital dos PEPs?

De acordo com a revisão de literatura, observamos que também não existe um consenso na adoção de modelos ou estratégias relacionados à preservação de dados digitais clínicos por longo prazo. No entanto, alguns esforços foram feitos nas últimas duas décadas (Scott, 2007; Corn, 2009; Giaretta, 2011) para definir estratégias genéricas (conforme visto no Capítulo 2) que devem ser consideradas para abordar a preservação digital. A maioria das propostas, por exemplo, menciona a influência do modelo de referência OAIS (CCSDS, 2012).

O seguimento dessas estratégias tem a finalidade de manter as características dos documentos digitais (Suderman, 2010; Dos Santos, 2012), garantindo sua autenticidade, integridade, funcionalidade e acesso contínuo pelo tempo que seja necessário, documentando seus procedimentos em estruturas de metadados conforme foi estudado em Gordon (2018).

De outra perspectiva, Bote et al. (2011) aborda os custos da preservação digital de PEP e apresenta uma metodologia para análise de custos em preservação digital. Para isso, o autor descreve estratégias de preservação detalhadas em quatro fases: entrada, preservação, acesso e saída. A fase de entrada consiste na entrada e revisão das informações médicas. Após um processo de conversão (digitalização), as novas informações digitais são incorporadas ao armazenamento de arquivos (ingestão) (CCSDS, 2012). A fase de preservação consiste em manter as informações em boas condições, seguindo estratégias de armazenamento e custódia, auditoria e migração conforme instruções do OAIS. Considerando somente as informações médicas que se encontram distribuídas em quatro formatos PDF, WAV, MPEG, DICOM. A fase de acesso ou solicitação refere-se ao acesso à uma informação específica, recuperando uma cópia dela. Finalmente, a fase de saída visa a destruição de informações após a data de vencimento. Embora este artigo fale sobre custos, que foge de nosso escopo, o interessante da pesquisa é que nos ajuda a identificar quais formatos devem ser preservados e o valor que gera sua guarda por longo tempo.

Entretanto, Piconi (2014) descreve que as estratégias de implantação e substituição/migração aplicadas em sistemas de PEPs devem ser planejadas seguindo políticas de arquivamento, como padrões e ISOs, para evitar a perda de informações. Essas estratégias devem ser usadas para garantir, com o tempo, a recuperação do contexto original em que as informações clínicas foram produzidas usando esses sistemas.

Diversos autores (Sayão and Sales, 2013; Piconi, 2014; Oliveira, 2019) destacam que não basta preservar apenas o documento digital, também é necessário preservar os metadados que descrevem o ciclo de vida deste documento, bem como seus contextos tecnológicos (hardware e software), histórico, criação e atualização, desde o momento de sua produção até seu destino final.

Outros autores (Barateiro et al., 2010; Bote et al., 2012; Rudersdorf, 2012; Bunawan and Nordin, 2015; Fatima and Colomo-Palacios, 2018) que abordam as estratégias de preservação, se baseiam no trabalho de Scott (2007), quem fez uma análise das características de preservação de dados e apresentou recomendações e estratégias de políticas que devem ser consideradas para preservar essas informações em ambientes digitais. (ver Tabela 5).

Tabela 5. Estratégias e características da preservação de dados (Scott, 2007)

Estratégias	Características
Características dos dados de preservação	<ul style="list-style-type: none"> ✓ Opções de armazenamento ✓ Meios de armazenamento ✓ Obsolescência tecnológica ✓ Escalabilidade ✓ Custo do financiamento
Soluções para preservação de dados	<ul style="list-style-type: none"> ✓ Atualização de dados ✓ Migração de dados ✓ Emulação ✓ Extensibilidade ✓ Filtragem de dados ✓ Compatibilidade com versões anteriores
Estratégias, opções de política e recomendações	<ul style="list-style-type: none"> ✓ Nível prático: deve ser utilizado guidelines para facilitar a identificação de registros relevantes e sua jurisdição (ex. hospital, posto de saúde regional, estadual, etc.) ✓ Nível do programa: representado por uma organização nacional supervisora que deve se encarregar de liderar o caminho ao uso de tecnologias e soluções para manter uma base de conhecimento de livre acesso ✓ Nível político: representantes e responsáveis para definir a rota nacional ou jurisdicional de e-saúde para garantir que a legislação não restrinja inadvertidamente opções futuras para preservar os dados. ✓ Setores público e privado: parcerias e definição de estratégias para implementar soluções em pro da preservação de dados a baixo custo e de longo prazo.

QP3: Qual é o papel dos metadados na preservação digital de dados de saúde?

Entre os mecanismos tecnológicos existentes para apoiar a preservação digital está a associação de metadados a documentos eletrônicos. Segundo Kiefer et al. (2012), eles são considerados uma parte essencial de qualquer arquivo de preservação ao longo prazo. Os metadados (Riley, 2017) são informações estruturadas que descrevem e facilitam a recuperação

e gerenciamento de um recurso de informação, neste caso os PEPs (Doyle et al., 2008). No contexto internacional, existem padrões de metadados focados na preservação de documentos (Bunawan and Nordin, 2015), por exemplo: PREMIS (*PREservation Metadata: Implementation Strategies*)¹⁶ que é um dicionário de dados para preservação de metadados ao igual que o e-ARQ¹⁷ Brasil e o ISO31606 abordado no trabalho de Gomes (2013). Eles descrevem um recurso prático e abrangente para implementar metadados de preservação em sistemas de arquivamento digital. No entanto, esses padrões são genéricos e não atendem necessariamente às demandas impostas pelo PEPs (Barateiro et al., 2010).

O principal objetivo dos metadados é manter os recursos importantes das informações. Assim, Bunawan and Nordin (2015) apresenta dez desafios de preservação de metadados para registros digitais que são: padrão e ferramentas de metadados, precisão de categoria, resistência do usuário à mudança, conflito de metadados, extração de metadados principais, data do último acesso, metadados automáticos, repositórios de planos de arquivos, predefinidos metadados e autenticidade dos metadados. A abordagem do autor coincide com as de Gordon (2018) e Limo et al. (2019) que afirmam que, ao gerenciar e usar registros digitais, é importante incorporar metadados para ajudar na recuperação futura desses registros. Porém, Bunawan and Nordin (2015) enfatizam que a preservação de metadados deve ser herdada em qualquer processo de gerenciamento de registros eletrônicos. Esse recurso deve se tornar um componente padrão dentro da organização para o futuro e deve passar por um processo de validação em implementações de preservação de registros digitais. Além disso, para Kiefer et al. (2012), é necessário desenvolver padrões para codificar anotações explicativas que atuem como metadados, que permita a interpretação desses documentos e a padronização de formatos para sua produção.

Similarmente, Gartner and Lavoie (2013) afirma que a preservação e disseminação de metadados consiste em processos necessários para integrar, ingerir, preservar metadados e garantir que estes sejam descobertos e acessíveis aos colaboradores e não colaboradores. No entanto, essas posições teóricas são genéricas e não atendem necessariamente às demandas específicas impostas no domínio dos PEPs. Além disso, Oliveira (2019) aborda o uso de metadados na curadoria digital. Ele afirma que a coleta, criação, avaliação e seleção de metadados é importante para garantir a qualidade da preservação dos metadados no ciclo de

¹⁶ <http://www.loc.gov/standards/premis/>

¹⁷ <http://www.siga.arquivonacional.gov.br/images/publicacoes/e-arq.pdf>

vida da curadoria de documentos digitais. E o estudo de Gordon (2018) enfoca os conceitos de patrimônio digital. Nesse contexto, propõe-se um modelo para explicar que a relação entre preservação e padrões digitais do patrimônio está bem relacionada devido aos aspectos de acessibilidade e contexto.

QP4: Que tipos de modelos de referência são usados para representar a preservação digital de PEPs?

Na revisão percebemos que um dos modelos de referência para preservação de dados digitais mais utilizado é o OAIS (Joemon et al., 2010; Bote et al., 2011; Misra et al., 2012; Piconi, 2014; Pinto and Sales, 2017; Oliveira, 2019). OAIS é um modelo conceitual que visa identificar os componentes funcionais que devem fazer parte de um sistema de informação dedicado à preservação digital (CCSDS, 2012; Lavoie, 2014). Similar enfoque adotou Celesti et al. (2016), que destaca a preservação de metadados usando o padrão OAIS, que serve como uma descrição associada ao documento a ser preservado. Celesti et al. (2016) considera um caso de estudo para demonstrar que, em um cenário escalável, um armazenamento de arquivo para processamento de big data é mais conveniente para hospitais do que os sistemas tradicionais de armazenamento de arquivo. Porém os domínios em que é aplicado este modelo não está relacionado à preservação de dados clínicos, somente os trabalhos de Gomes (2013) e Pinto and Sales (2017) consideram os PEPs. O primeiro propõe que o modelo de referência OAIS deve ser aplicado em PEPs, assim como curadoria, com o objetivo de preservar o conteúdo digital, a fim de garantir confidencialidade, confiabilidade, autenticidade e acesso às informações registradas nesses documentos. Por sua vez, Gomes (2013) centra-se em questões de acesso e preservação da informação clínica e apresenta recomendações referentes à definição de uma política de gestão de dados e de metadados em registros clínicos digitais. Na sua análise aborda diferentes padrões da área da saúde entre eles HL7, CDA, PREMIS, ISO 31606, DICOM, OpenEHR, SNOMED.

Outros estudos usam ontologias para estabelecer e avaliar a presença de metadados e estratégias de preservação para PEPs produzidos e gerenciados por sistemas de informação de saúde. Kiefer et al. (2012) apresenta uma ontologia para preservação digital e gerenciamento de metadados. Os autores usaram o padrão de preservação PREMIS e fizeram uma comparação com o padrão OAIS. Por outra parte Malone et al. (2014) usou um SWO (*Software Ontology*) para armazenar, gerenciar e analisar dados de ciências da vida. Os autores usaram técnicas ágeis para reunir informações para o SWO e manter o envolvimento com os usuários. O principal uso

do SWO está na descrição dos recursos utilizados no armazenamento, gerenciamento e análise de dados.

Por outro lado, Piconi (2014) propôs uma ontologia para formalizar conceitos capazes de capturar os principais relacionados à definição de metadados para preservação de documentos clínicos. O autor definiu um conjunto de metadados, visando à preservação dos documentos (nascidos digitais ou digitalizados) que compõem os PEP, utilizando como modelo a *Arquitetura de Documentos Clínicos (CDA)*¹⁸ e os metadados do e-ARQ-Brasil: Requisitos de Modelo para Gerenciamento de documentos de arquivamento de sistemas computadorizados.

Outras abordagens consideram outros tipos de modelos, como é o caso de Gordon (2018) que propõe um modelo de herança de dados (DHP) para explicar como construir uma estrutura precisa para preservação de dados. A ideia desta pesquisa está relacionada ao gerenciamento do patrimônio, diretamente em termos de preservação. Outro trabalho que considera os modelos DHP é a proposta de Chaczko et al. (2018). Oliveira (2019) já propõe um metamodelo para ecossistemas de dados como parte de um objetivo principal, que consiste em uma estrutura que oferece uma ampla gama de processos para auxiliar a curadoria de metadados, entre eles metadados de saúde. A Tabela 6 lista os modelos encontrados na revisão de literatura.

Tabela 6. Modelos utilizados nos estudos selecionados

Modelo	Referências
OAIS	S07, S15, S17, S18, S21
Ontologias	S09, S10, S13,
ISO31606	S10
CDA	S10
Outros	S25, S26, S29

QP5: Quais são os principais papéis dos atores nos estudos de preservação de PEPs?

Atores são entidades autônomas, como empresas, instituições ou indivíduos, que atuam na preservação de dados de saúde em uma ou mais funções específicas. Os atores são motivados

¹⁸ CDA: É uma especificação do padrão HL7. Trata-se de um padrão de marcação em documentos que especifica a estrutura e semântica de ‘documentos clínicos’ para fins de troca entre prestadores de serviços de saúde e pacientes com base em XML. Fonte: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=7. Acessado 20/11/2019.

por um conjunto de interesses e envolvidos por um objetivo comum, para nosso escopo, a preservação de dados clínicos (Bote et al. 2011; Gartner and Lavoie, 2013; Zhu et al. 2019).

A preservação a longo prazo do PEP é complicada pela possibilidade de um dia os registros serem usados longitudinalmente e integrados nos locais de atendimento. Os registros têm o potencial de serem criados, usados, editados e visualizados por várias entidades ou atores independentes. Essas entidades incluem, entre outros, médicos de atenção primária, hospitais, companhias de seguros e pacientes. Portanto, as escolhas sobre a estrutura e a propriedade desses registros terão um impacto profundo na acessibilidade e privacidade das informações do paciente (Yina, 2010; Abouelmehdi et al., 2018; Zhu et al., 2019).

De acordo com o modelo de referência da OAIS (CDSC, 2012), temos atores e componentes (Giaretta, 2011). Produtor e Consumidores são os Gerentes. No contexto da preservação dos dados de saúde, o Produtor é a equipe profissional de saúde, pois produz e registra dados, informações e conhecimentos no prontuário do paciente (Gomes, 2013; Celesti et al., 2017). A Administração é o papel desempenhado por aqueles que definem a política geral do OAIS como um componente em um domínio político mais amplo, por exemplo, um hospital. O Consumidor será o papel desempenhado por pessoas ou sistemas clientes que interagem com os serviços da OAIS para encontrar e adquirir informações preservadas de interesse, por exemplo, equipe profissional de saúde, Ministério da Saúde, paciente, pesquisadores, etc. A Tabela 7 identifica os atores envolvidos nos estudos.

Tabela 7. Lista de atores envolvidos nos estudos abordados

Atores	Referência
Hospitais, centros de saúde	S23, S20, S19, S15
Serviços públicos	S21
Computação na nuvem	S17, S18
Governo	S16
Bibliotecas públicas	S07, S08, S11

QP6: Qual é o papel da curadoria de dados nos estudos de preservação do PEP?

A curadoria é a atividade que visa gerenciar e promover o uso de dados desde sua criação, de modo a garantir que este seja adequado para fins contemporâneos e disponível para a descoberta e reuso (Walters and Skinner, 2011; Oliveira, 2019).

Bote et al. (2011) afirma que uma vez que a adoção dos processos de preservação digital seja prioridade nas organizações de saúde, elas serão afetadas por obrigações legais que os forçará a desenvolver políticas de curadoria de dados. Entretanto, Rudersdorf (2012), relata o desenvolvimento de uma aplicação que incorpora tecnologias de curadoria, de forma a coletar e autenticar registros na fase de ingestão. Por outro lado, Malone et al. (2014) aborda a construção de uma ontologia biomédica cujo objetivo foi descrever como os dados são analisados para descobrir uma trilha de auditoria, proveniência e reprodutibilidade. E para alcançar sua proposta eles utilizaram um modelo de curadoria. Porém não menciona como foi realizado esse processo, mas sim que utilizaram como base a ontologia EDAM¹⁹ que define conceitos da bioinformática.

3.4 CONSIDERAÇÕES FINAIS

A preservação digital de dados clínicos, presente no PEP, é uma demanda real para ajudar o acesso atual e futuro ao crescente da produção de registros digitais de saúde, pois essas informações representam um recurso valioso de pesquisa para garantir assistência médica e eficiência diagnóstica no tratamento do paciente (Piconi, 2013; Bunawan and Nordin, 2015).

Devido a que os registros médicos contêm informações sensíveis (Bheemanathini, 2019), sua preservação digital requer cuidados especiais sobre aspectos éticos e legais. Dessa forma, é necessário um entendimento claro dos aspectos relacionados à preservação digital de PEPs (Scott, 2007), bem como a definição de estratégias, políticas e diretrizes que garantam o acesso e o uso dessas informações a longo prazo (Kiefer et al., 2012; Lo et al., 2017; Abouelmehdi et al., 2018). A preservação digital é sobre manter a capacidade de armazenar o essencial de objetos digitais e proteger esse material contra potenciais ameaças (Gartner and Lavoie, 2013).

Nesse sentido, organizações nacionais e internacionais (Bote et al., 2011; Gartner and Lavoie, 2013; Malone et al., 2014; Celesti et al., 2016) desenvolveram várias estratégias para preservar os registros digitais a nível nacional ou institucional. Entretanto, com base na revisão

¹⁹ <https://github.com/edamontology/edamontology/>

da literatura, conforme mostrado na seção anterior (3.3), foi possível identificar que os PEP não devem apenas compartilhar informações, mas também ser armazenados por um longo período para servir, entre outras coisas, como fontes de assistência e monitoramento da qualidade da saúde do paciente e também como fonte de pesquisas científicas (Celesti et al., 2016; Pinto and Sales, 2017; Sassi et al., 2017). A preservação dos dados clínicos de pacientes a longo prazo é essencial para a sobrevivência e o desenvolvimento humano. Nessa perspectiva, o PEP gerado nas instituições de saúde são documentos sensíveis (Bheemanathini, 2019) extremamente importantes para pacientes, familiares, médicos e pesquisadores. E estes documentos digitais devem ser armazenados permanentemente, garantindo sua preservação, garantindo sua autenticidade, confiabilidade e integridade (Santos et al., 2010; Van Gorp et al., 2014; Pinto and Sales, 2017; Gordon, 2018)

Portanto, é necessário desenvolver e aplicar regras para codificar anotações explicativas, ou metadados (Kiefer et al., 2012; Gomes, 2013; Bunawan and Nordin, 2015; Riley, 2017; Gordon, 2018; Oliveira, 2019), que permitam a interpretação desses documentos e a padronização de formatos, desde a produção até o armazenamento permanente (Waters and Garrett, 1996). A importância do registro de metadados de preservação garante a compatibilidade de mídia e formato, interpretação de documentos (processamento e estrutura de dados), descrição do documento, presunção de autenticidade e informações de migração que possam ter ocorrido nos documentos (CONARQ, 2015), além de facilitar a localização e a busca de informações, pois esses dados clínicos são, por sua vez, diversificados e inscritos em diversos suportes como: imagens geradas em radiologia, ultrassonografia, tomografia computadorizada, gráficos, planilhas, de observações e evolução clínica, os quais constituem os registros médicos que compõem o PEP.

CAPÍTULO 4 – PRESERVAÇÃO DE DADOS CLÍNICOS

Os capítulos anteriores estabeleceram o referencial teórico necessário para o entendimento do PEP e a preservação digital com seus respectivos padrões e modelos de referência. Assim, este capítulo corresponde à etapa de sugestão da metodologia utilizada onde se justifica a criação de um metamodelo. Para isso, primeiramente se identifica um conjunto de características pertinentes à preservação de dados clínicos e na sequência se reúnem modelos e padrões já existentes que permitam abstrair essas características. Estes modelos são padrões e normas já estabelecidas na literatura médica, gestão arquivística e outros permitindo reunir as ferramentas necessárias para formalizar o metamodelo.

4.1 IDENTIFICAÇÃO DE CARACTERÍSTICAS

No Capítulo anterior foi possível observar que existe uma carência de um entendimento claro dos aspectos relacionados à preservação digital de PEPs que garantam o acesso e o uso dessas informações a longo prazo. Mas existem algumas iniciativas que destacam a sua importância (Kiefer et al., 2012; Lo et al., 2017; Abouelmehdi et al., 2018). Portanto, para que haja um entendimento global do problema é necessário identificar as características que descrevem os processos relacionados à assistência médica de um indivíduo, os documentos clínicos gerados e sua preservação.

Segundo o estudo de Lærum and Faxvaag (2004) existem 11 tarefas básicas que são realizadas pelos profissionais de saúde com maior frequência em ambientes hospitalares, elas são apresentadas a seguir:

- Revisar os problemas do paciente
- Procurar informações específicas no registro do paciente
- Obter os resultados de novos testes e investigações
- Adicionar notas diárias sobre condições do paciente
- Requisitar análises laboratoriais
- Requisitar exames de vídeo/imagem
- Obter resultados laboratoriais
- Obter resultados de exames de vídeo/imagem
- Requisitar tratamentos
- Escrever prescrições

- Registro de códigos para diagnóstico e procedimentos executados

As tarefas listadas geram documentos clínicos que fazem parte do PEP, os tipos de informações que compõem estes documentos são referentes a dados de identificação do profissional e do paciente; dados da anamnese, hipóteses diagnósticas e diagnóstico definitivo; exames clínicos, laboratoriais ou de imagem; receitas/prescrição de medicamento; fichas de atendimento ambulatorial ou de urgência; folhas de observação médica, boletins médicos, laudos biopsias e lâminas; planos de tratamento, raciocínio médico, descrições cirúrgicas e fichas anestésicas; evolução do tratamento, conduta terapêutica e resumos de alta (OMS, 2006). Considerando estes itens foi criado um fluxograma para mostrar o funcionamento em ambientes hospitalares ver Figura 14.

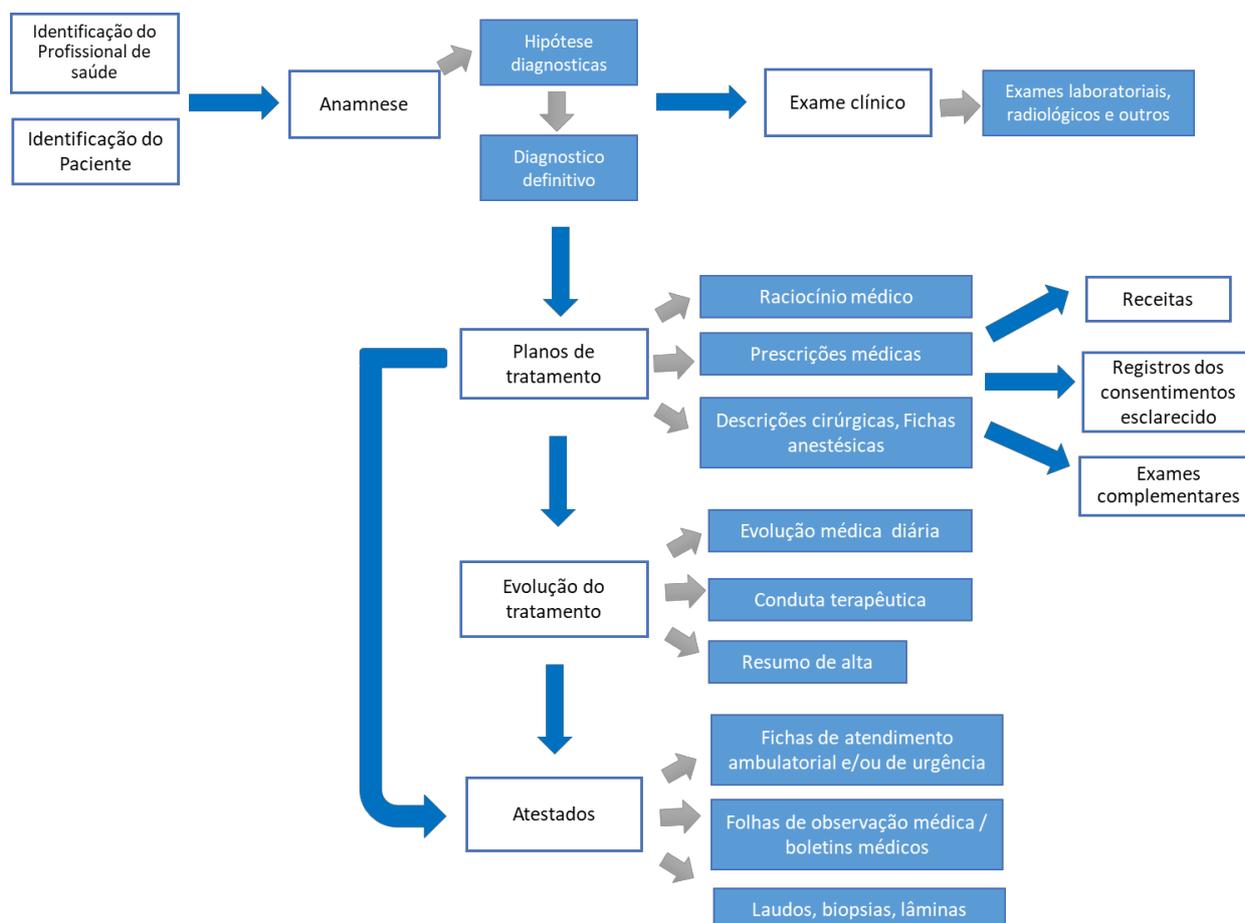


Figura 14: Fluxograma para a colheita de dados no PEP (Fonte: O autor)

Na Tabela 8 se descreve os elementos do fluxograma da Figura 14, assim como os principais elementos do PEP.

Tabela 8. Descrição dos elementos do PEP

Nome	Descrição
Identificação do profissional de saúde	Através do código de identificação do profissional pode-se recuperar a informação dos envolvidos no registro do PEP
Identificação do paciente	Dados pessoais do paciente (ex. nome completo, data de nascimento, naturalidade, endereço, entre outros.)
Anamnese → Hipóteses diagnósticas e diagnóstico definitivo	Histórico de doenças tanto do paciente quanto dos incidentes na família. Hábitos relacionados à saúde, tratamentos e descrição do estado físico do paciente.
Exame Clínico → Exames laboratoriais, → Imagens radiológicas e outros	A descrição é realizada em uma ficha de identificação com os principais apontamentos feitos pelos agentes envolvidos. Os relatórios tanto de exames de imagens quanto de laboratórios servem para nortear um tratamento ou um diagnóstico. A informação é apresentada em ordem cronológica.
Planos de tratamento → Raciocínio médico → Prescrições médicas → Descrições cirúrgicas, fichas anestésicas	Evolução diária do paciente com registro de data e hora. Relatório de todos os procedimentos nos quais o paciente foi submetido e identificação dos profissionais participantes (médicos, enfermeiras, assistentes sociais e outros profissionais da saúde) com assinatura eletrônica se o documento é elaborado e/ou armazenado em meio eletrônico.
Receitas/prescrição de medicamento	Consiste em uma lista com medicamentos que já foram prescritos ao paciente. Também indica quais os medicamentos que o paciente tem alergia ou intolerância.
Registros dos consentimentos esclarecido	Apresentam informações referentes ao consentimento, por parte do paciente ou responsável sobre formas de tratamento e procedimentos cirúrgicos.
Exames complementares	Exames de apoio para o diagnóstico. Complementam os dados da anamnese e de exame físico para confirmação de uma hipótese diagnóstica e tratamento.
Evolução do tratamento	Observação de sinais de melhora ou piora do estado de saúde do paciente. Assim como a identificação de efeitos adversos.
Atestados → Fichas de atendimento ambulatorial e/ou atendimento de urgência → Folhas de observação médica, boletins médicos → Laudos biopsias, lâminas	O registro de atendimentos ou assistência clínica gerará um determinado atestado.

Os planos ou protocolos de atendimento usados nas entidades de saúde dependem do tipo da doença ou sintoma que o paciente apresenta (OMS, 2006). Assim o registro dessas informações é gerenciado por diferentes SIS que devem lidar com requisitos complexos de

interoperabilidade usando protocolos padronizados. Preservar documentos clínicos e torná-la acessível é uma tarefa complexa (Bunawan and Nordin, 2015) e para assegurar condições mínimas de preservação destes documentos é necessário o uso de metadados que permitam a interpretação destes documentos e a padronização de formatos para sua produção (Piconi, 2014). Os metadados são importantes em qualquer arquivo de preservação a longo prazo, pois servem para descrever os objetos arquivados e suas relações com as aplicações e apoiam a presunção de autenticidade do documento (CONARQ, 2011).

Diante disso, de acordo à revisão bibliográfica realizado nos Capítulos 2 e 3, identificamos um conjunto de características para preservação de dados clínicos e são mostrados na Tabela 9.

Tabela 9. Identificação de características para a preservação de dados clínicos

Características de preservação	Referência
PEPs servem como repositórios de informações relativas à assistência em saúde.	(OMS, 2003; Galvão and Ricarte, 2012; Gomes, 2013; Araujo et al., 2014; Graña and Jackwoski, 2015; Lozano, 2016; Pinto and Sales, Pinto and Sales, 2017)
PEPs são fonte de informação valiosa para as pesquisas e estudos clínicos epidemiológicos.	(Galvão and Ricarte, 2012; Yina, 2010; Araujo et al., 2014; Evans 2016)
PEPs são gerenciados por diferentes tipos de SIS tanto de entidades públicas ou privadas	(Bote et al., 2011)
PEPs servem como fontes que norteiam diagnósticos e tratamentos	(Yina, 2010; Araujo et al., 2014; Graña and Jackwoski, 2015)
PEPs contém o histórico de informação clínica de um indivíduo a longo de sua vida	(OMS, 2003; Bote et al, 2011; Malone et al., 2014)
PEPs são elementos de comunicação entre setores de uma instituição de saúde e entre diferentes atores envolvidos.	(Bote et al., 2011; Lozano, 2016)
Preservação representa informação estratégica para a tomada de decisões.	(Scott, 2007; Corn, 2009; Galvão and Ricarte, 2012)
Preservação deve garantir a autenticidade, confiabilidade, privacidade, controle de acesso e a integridade de documentos clínicos associados ao indivíduo.	(Scott, 2007; Muñoz et al., 2011; Corn, 2009; Gartner and Lavoie, 2013; Goossen, 2014; Pinto, 2017; Chaczko et al., 2017; Gordon, 2018; Abouelmehdi et al., 2018; Zhu et al., 2019; Chenthara et al., 2019)
Preservação compreende a combinação de políticas, estratégias e ações para garantir o acesso e reprodução.	(Ferreira, 2006; Márdero, 2008; Bote et al, 2011; Lavoie, 2014; Cohen et al., 2015)

Permite a tomada de decisões nos diagnósticos.	(Burns, 2011; Misra et al., 2012; Bunawam and Nordin, 2015; Evans, 2016; Mihailescu et al., 2017)
Dados armazenados digitalmente precisam de proteção durante toda sua vida útil.	(Walters and Skinner, 2011; Gartner and Lavoie, 2013; Kiefer et al, 2012; Fatoric and Seekamp, 2017; Zhu et al., 2019)
Dados clínicos de um paciente devem ser preservados.	(Corn, 2009; Bote et al, 2011; Kiefer et al., 2012; Cohen et al., 2015; Chaczko et al., 2017)
Deve existir um mecanismo de recuperação de dados preservados	(Barateiro et al., 2010; Bote et al, 2011; Rudersdorf, 2012; Gartner and Lavoie, 2013; Malone et al., 2014)
A preservação digital é independente da tecnologia	(Rabbi et al., 2014; Celesti et al., 2017)
Existência de um plano de preservação	(Rudersdorf, 2012; Kiefer et al., 2012; Malone et al., 2014)
Existência de estratégias de preservação	(Time, 2002; Barateiro, 2010; Rudersdorf, 2012; Bunawan and Nordin, 2015)
Objetos preservados devem ter políticas de acesso bem definidas	(Chaczko et al., 2017; Gordon, 2018; Abouelmehdi et al., 2018; Zhu et al., 2019; Chenthara et al., 2019)
Armazenamento e preservação de metadados	(Barateiro et al., 2010; Kiefer et al, 2012; Gomes, 2013; Gartner and Lavoie, 2013; Bunawan and Nordin, 2015; Celesti et al., 2016; Riley, 2017; Gordon, 2018; Limo et al., 2019)
Curadoria de dados	(Rusbridge et al., 2005; Higgins, 2008; Abrams et al., 2009; Barateiro et al., 2010; Walters and Skinner, 2011; Rudersdorf, 2012; Sayão, 2013; Abbott, 2014; Malone et al., 2014; Oliveira, 2019)

4.2 INTEGRAÇÃO DE MODELOS

Conforme visto na literatura, não existe uma representação ou abstração que faça a junção de todas as características identificadas na seção anterior. Mas os processos clínicos, PEPs e a preservação de dados possuem um conjunto de regras, diretrizes ou características estabelecidas em normas separadamente, com seus respectivos modelos. Diante disso, esta seção tem por objetivo apresentar os modelos que abstraem essas características.

4.2.1 MODELO DE PROCESSOS CLÍNICOS

A norma ISO 13940 (2016), descrita na seção 2.2.2, abstrai o processo de assistência médica abordando temas relativos à organização de uma entidade de saúde, atores envolvidos (profissionais da saúde diretamente ou indiretamente relacionados) e outras atividades de assistência clínica. Apesar desta norma conceitualizar vários aspectos relacionados ao cuidado da saúde, nesta tese somente tomamos como referência o contexto de processos de saúde do

paciente. Na Figura 15 podemos observar parte de um modelo desta norma referente ao modelo de processos de saúde. A classe *healthcare process* representa um conjunto de atividades de saúde interrelacionadas ou em interação com ela mesma que transforma insumos em produtos. Os insumos podem ser dados referentes ao estado de saúde do paciente (*input health state*), atores como um paciente em tratamento (*subject of care*) ou alocação e gerenciamento de recursos (*healthcare administration*) no processo clínico. Os produtos se traduzem em: tratamentos e resultados do estado de saúde (*output health state*); plano de cuidados (*care plan*) é um conjunto de estratégias para melhorar ou resolver um problema de saúde; prontuário eletrônico (*health record*). Este último, é um documento clínico concreto que reúne todas as informações clínicas do paciente.

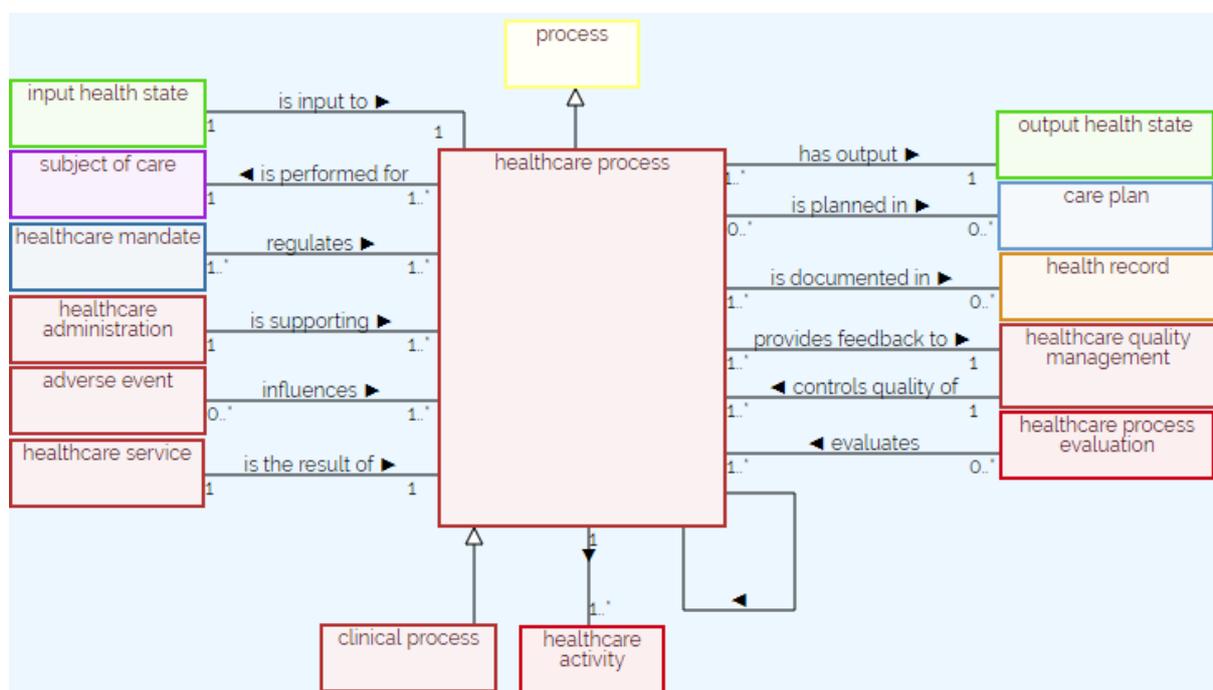


Figura 15: Modelo de processos clínicos ou de assistência clínica. (Fonte: ISO 13940:2016)

4.2.2 MODELO DE PEP

Seguidamente temos a norma ISO/EN13606-1 (descrita na seção 2.2.3), fornece um modelo de referência para representar as características globais dos componentes dos PEPs, como eles são agregados e as informações de contexto necessárias para atender aos requisitos éticos, legais e de proveniência (Munõz et al., 2011). Este modelo define o conjunto de classes que formam os blocos de construção genéricos do PEP (ver Figura 16), portanto, são capazes de abstrair a informação clínica contida dentro dos documentos clínicos gerados nos processos de saúde. A classe principal neste modelo é o *Composition* que representa um documento clínico que pode

ser por exemplo um laudo ou um exame. Documentos clínicos são organizados em pastas (*Folder*) que fazem parte do PEP representado por *EHR_Extract*. Documentos clínicos estão estruturados em seções (*Section*) que por sua vez agrupam um conjunto de elementos (*Element* ou *Cluster*), por exemplo sintomas do paciente.

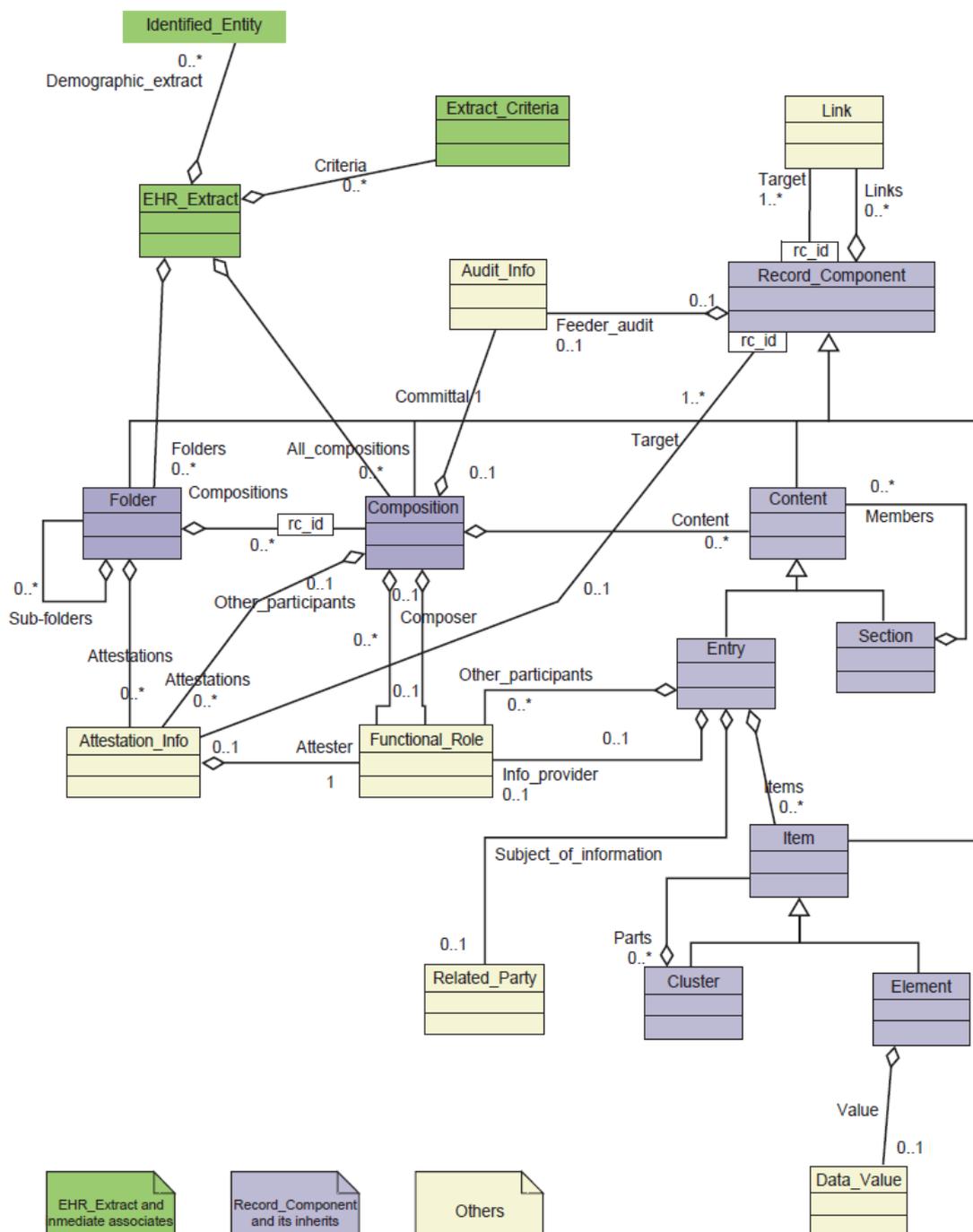


Figura 16: Modelo de referência *Electronic health record*. (Fonte: ISO 13606-1:2019)

4.2.3 MODELO DE PRESERVAÇÃO DIGITAL

Um projeto de preservação digital, independentemente da natureza dos dados, é complexo, posto que envolve uma quantidade enorme de variáveis em todo o seu desenvolvimento e acompanhamento a longo prazo. A partir dos estudos da literatura sobre preservação digital, verificamos que o modelo de referência OAIS (ISO 14721, 2012) integra diversos esforços em torno de um sistema conceitual (Lavoie, 2014), que serve de base comum para iniciativas de arquivamento da informação digital. OAIS é um modelo abstrato, que proporciona flexibilidade aos desenvolvedores de sistemas, que desejam utilizar repositórios permitindo interpretar seu uso de forma relevante em qualquer campo de conhecimento e conteúdo (Cruz and Díez-Carrera, 2016). Dessa forma, este modelo é útil para abstrair a preservação de dados clínicos provenientes do PEP conforme também foi mostrado nos trabalhos de Misra et al. (2012), Graña and Jackowski (2015), Bunawan and Nordin (2015), Celesti (2016) e Pinto and Sales (2017).

Desta maneira, em conformidade ao nosso escopo de estudo, os elementos do OAIS (descritos na Seção 2.1.2) podem ser adaptados para preservação de dados clínicos realizando os seguintes ajustes:

- **Produtor:** consideramos os atores envolvidos nos processos clínicos como por exemplo a equipe multiprofissional da saúde, posto que ela produz e registra dados, informações e conhecimentos no PEP.
- **Administração:** equipe de curadoria da entidade de saúde, são componentes internos da aplicabilidade do modelo OAIS. Por exemplo, na esfera hospitalar, o Serviço de Arquivo Médico e Estatística de Saúde (SAME) está conformado por uma equipe multidisciplinar que se ocupa por garantir que os prontuários sejam registrados seguindo as normas determinadas pela Comissão de Revisão de Prontuários (CFM, 2002). Portanto, o SAME é encarregado de lidar com os aspectos sigilosos, éticos e legais dos dados no prontuário e da sua preservação.
- **Consumidor:** equipe multiprofissional de saúde, entidades governamentais Ministério da Saúde, paciente, pesquisadores (devidamente autorizados) e estudantes da área de saúde de modo geral.
- **Ingestão:** entrada dos documentos clínicos que, após o aval da Comissão de Prontuários de cada entidade de saúde, formam parte do chamado Pacote de Submissão de Informação (SIP).

- **Gestão de dados:** Neste bloco é feito o gerenciamento dos dados e metadados descritivos e temáticos definidos na ingestão/entrada, visando garantir, de forma padronizada, todo o conteúdo registrado no documento, a fim de facilitar a preservação e a curadoria dessa documentação.
- **Armazenamento:** Fornecer serviços de modo que se garanta a autenticidade, segurança, integridade e confidencialidade/privacidade dos documentos clínicos preservados.
- **Plano da preservação:** É responsável pelo monitoramento das tecnologias (hardware e software) visando à atualização do sistema de preservação dos prontuários. Para tanto, elabora e decide as políticas a serem adotadas na organização de saúde que devem estar em consonância com as políticas de saúde do Estado-Nação.
- **Acesso:** Por se tratar de informações clínicas cujo acesso é de certa forma restrito, deve-se atentar para a necessidade de interface que indique o acesso por meio de senha que certifique e garanta legalmente a autenticidade, segurança, integridade, confidencialidade/privacidade dos documentos e dos metadados atribuídos.

Observe-se que produtores e consumidores podem ser instâncias do modelo *Healthcare actors* da norma ISO 13940. O Pacote de Submissão de Informação (SIP) é a entrada do sistema de preservação e é representado pela classe *Archival information package* do modelo de referência OAIS (veja Figura 17). Esta classe corresponde à unidade lógica que inclui tanto um objeto digital quanto outros tipos de informações que devem ser associados ao objeto digital, a fim de preservar e proporcionar acesso significativo a ele ao longo do tempo. Esta classe por sua vez está composta por informações de conteúdo (*Content Information*) que no contexto de saúde corresponderia a um documento clínico a ser preservado. As informações de conteúdo estão compostas por objetos de dados (*Data object*) que representam a informação digital e por uma representação (*Representation Information*) que permite sua interpretação e por sua vez é necessário para transformar completamente os bits nas informações de conteúdo.

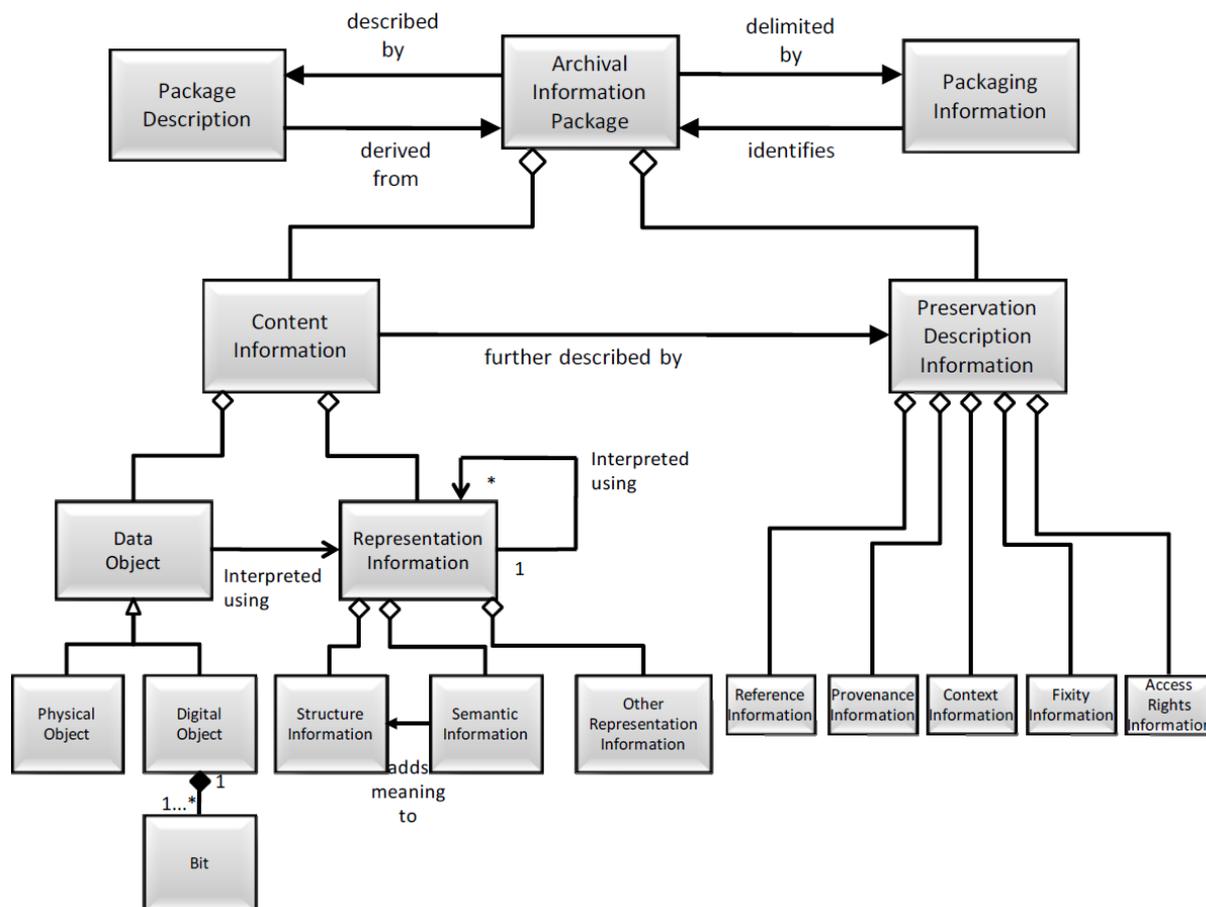


Figura 17: Modelo de Referência OAIS (Fonte: CCSDS, 2012)

4.2.4 PROPOSTA (INTEGRAÇÃO DE MODELOS)

Considerando o exposto acima, se propõe criar um metamodelo por meio de abstrações de modelos já existentes na literatura que permita ter uma visão global da preservação de dados clínicos em um alto nível de abstração independente da arquitetura e implementação. Para isto utilizamos os modelos de processos clínicos da norma ISO 13940 (2016) para representar eventos de saúde associados a uma pessoa ao longo da sua vida. Estes processos clínicos também são abstraídos reutilizando o metamodelo proposto por Behnam and Badreddin (2013) para representar caminhos clínicos (*clinical pathways*). Os documentos clínicos gerados nestes processos são abstraídos pela norma ISO 13606 (Muñoz, 2011) e o modelo de referência OAIS (CCSDS, 2012) é utilizado para preservação destes dados.

Um aspecto importante a considerar são as políticas de segurança, pois é fundamental salvaguardar a privacidade da informação do paciente. De forma geral, políticas de segurança refere-se a um conjunto de diretrizes, regulamentos, regras e práticas que prescrevem como uma organização gerencia, protege e distribui informações (Karadsheh, 2012; Abouelmehdi et

al. 2018; Li et al., 2018). Apesar dos modelos citados acima já considerarem aspectos de políticas de segurança, se utilizam abstrações de políticas de acesso como o metamodelo de Li et al., (2015). Este metamodelo é orientado à proteção de dados sensíveis e está fundamentado nas propostas de Ardagna et al. (2010) e Jin et al. (2011) que propõem arquiteturas de controle de acesso com foco no PEP.

Conforme visto na seção anterior, é importante considerar também os metadados no processo de preservação de dados clínicos, pois estes têm a função de apoiar sistemas que gerenciam objetos digitais (Piconi, 2014). Para isso é necessário estabelecer metadados essenciais para os documentos arquivísticos digitais. Diante disso, existem iniciativas como é o caso do projeto InterPARES (*International Research on Permanent Authentic Records in Electronic Systems* - Pesquisa Internacional sobre Documentos Arquivísticos Autênticos Permanentes em Sistemas Eletrônicos) que fornece um conjunto de especificações e modelos conceituais para auxiliar tanto produtores quanto preservadores de documentos arquivísticos na gestão de metadados associados com estes documentos.

O contexto da preservação de dados clínicos exposto inclui atores como produtores e consumidores trocando dados, isto se encaixa na definição de ecossistema de dados que de acordo com Oliveira (2019) é um conjunto de redes compostas por atores autônomos que consomem direta ou indiretamente, produzem ou fornecem dados e outros recursos relacionados, por exemplo, software, serviços e infraestrutura. Cada ator desempenha um ou mais papéis e é conectado a outros atores por meio de relacionamentos, de maneira que a colaboração e a concorrência dos atores promovam a auto regulação do ecossistema de dados. Portanto, a junção dos modelos abordados neste capítulo seriam também parte do metamodelo de ecossistema de dados proposto por Oliveira (2018). Os recursos a serem trocados pelos atores são os documentos clínicos com seus respectivos metadados.

A Tabela 10 mostra resumidamente os modelos e como estes serão utilizados na criação do metamodelo.

Tabela 10: Resumo de modelos utilizados para criação do metamodelo.

Modelos	Utilização	Referências
ISO 13940 e metamodelo processos clínicos.	Visa abstrair o processo de assistência médica de uma entidade de saúde e os atores diretamente ou indiretamente relacionados. Por exemplo uma consulta médica. A classe principal é denominada de <i>Health care process</i> .	REF ISO (Behnam and Badreddin, 2013)

ISO 13606	Visa abstrair os elementos do PEP, a classe principal é <i>Composition</i> que representa um documento clínico como um todo.	(Muñoz, 2011)
OAIS (ISO 14721)	Visa abstrair a preservação de dados. A classe principal é <i>ContentInformation</i> que representa o objeto digital a ser preservado e está diretamente relacionado a um documento clínico (<i>Composition</i>).	(CCSDS, 2012)
Ecosistema de dados	Permite abstrair entidades que compartilham dados em uma rede complexa de colaboração. A classe principal é <i>Resource</i> que generaliza os PEPs e objetos digitais como sendo entidades que podem ser produzidas, fornecidas ou consumidas pelos atores.	(Oliveira et al., 2019)
Políticas de acesso	Permite abstrair regras de acesso a recursos. A classe principal é <i>Rule</i> .	(Li et al., 2015)

4.3 CONSIDERAÇÕES FINAIS

Neste capítulo foi possível identificar um conjunto de características que descrevem a preservação de dados clínicos. Os processos clínicos ao qual um paciente é submetido desde a visita a uma entidade de saúde geram documentos clínicos que eventualmente farão parte do PEP. Este por sua vez é um documento arquivístico de preservação permanente e é mais que uma demanda legal, representa um valioso recurso patrimonial e de pesquisa.

Os modelos existentes na literatura fornecem um entendimento da estrutura semântica dos atores e entidades envolvidos nos processos clínicos, na produção e preservação de dados clínicos. Os modelos selecionados neste capítulo (maioritariamente formados por normas internacionais ISO) não descrevem nenhuma arquitetura, tecnologia, design de banco de dados ou implementação concreta de sistemas de informação com foco arquivístico. Eles são estruturas conceituais e representações abstratas focados em aspectos específicos do ecossistema da saúde. Em consequência, por meio da integração e generalização destes modelos e de seus componentes, no próximo capítulo se formalizará um metamodelo em alto nível de abstração do entendimento da preservação de dados clínicos.

Todos esses módulos deverão estar vinculados a fim de que o fluxo de informação e de comunicação do sistema de preservação de dados clínicos se efetive com o menor ruído possível e assegure os aspectos legais de autenticidade, segurança, integridade, confidencialidade/privacidade dos documentos e dos metadados atribuídos. Ademais, ratifique a interoperabilidade entre cada um desses módulos.

CAPÍTULO 5 – PROPOSTA DO METAMODELO

Continuando com a metodologia adotada nesta tese, este capítulo tem por objetivo descrever o desenvolvimento de um metamodelo capaz de abstrair a preservação de dados clínicos. Primeiramente se identifica um conjunto de elementos conceituais extraídos das etapas anteriores da metodologia para que o metamodelo proposto utilize uma terminologia comum aos modelos existentes na literatura. Na sequência o metamodelo é formalizado como sendo uma instância do meta-metamodelo *Ecore* descrito no Capítulo 2.

5.1 ELEMENTOS CONCEITUAIS PARA A PRESERVAÇÃO DE DADOS CLÍNICOS

A partir das características extraídas no capítulo anterior são identificados um conjunto de elementos conceituais listados na Tabela 11. Estes elementos servirão como base para a nomeação de criação das classes, atributos e relações do metamodelo aproveitando a terminologia presente nos modelos de referência e padrões já existentes na literatura.

Tabela 11. Elementos conceituais para construção do metamodelo

Elementos	Descrição do constructo	Referências
Ator	Representa uma entidade autônoma como paciente, curador e médico.	(Higgins, 2008; Thibodeau, 2010; Celesti et al. 2016; Oliveira, 2019)
Documento clínico	É um registro que faz parte do PEP.	(Corn, 2009; Yina, 2010; Gomes, 2013; Goossen, 2014; Van-Gord and Comuzzi, 2012; Saasi et al., 2017; Celesti et al., 2017; Lozano, 2016; Evans, 2016; Chaczko et al., 2018; Roehrs et al., 2017; Abouelmehdi et al., 2018; Gordon, 2018)
Processo clínico	Qualquer evento que gera dados clínicos que devem fazer parte do PEP.	(Goossen, 2014; Celesti et al., 2016; Saassi et al., 2017)
Prontuário	Histórico clínico de um paciente durante sua vida.	(Soares et al., 2011; Kiefer et al., 2012; Piconi, 2014; Lozano, 2016; Roehrs, 2017; Pinto and Sales, 2017; Limo et al., 2019)
Objeto digital	É a unidade de informação a ser preservada. Por exemplo, um laudo, imagem médica	(Korb and Strodl, 2010; Askhoj et al., 2011; Lavoie, 2014; Voinov et al., 2017).
Metadado	São atributos adicionados ou marcados em documentos clínicos (laudos, exames, imagens)	(Walters and Skinner, 2011; Bote, 2011; Kiefer et al., 2012; Gomes, 2013; Goossen, 2014; Bunawan and Nordin, 2015; Gordon and Chaczko, 2017;

		Chaczko et al., 2018; Gordon, 2018; Oliveira, 2019)
Recurso	Objeto ou generalização de um documento clínico. Objeto a ser preservado.	(Márdero 2008; Carvalho, 2013; Oliveira, 2019)
Padrão	Representa especificações, procedimentos e diretrizes. São projetados para garantir que produtos, serviços e sistemas sejam seguros, confiáveis e consistentes.	(Horiil et al., 1994; Mildenberger, 2002; Joemon et al., 2010; Bote et al., 2011; Misra et al., 2012; Gomes, 2013; Piconi, 2014; Pinto and Sales, 2017; Roehrs et al., 2017; Oliveira, 2019)
Semântica	Pode ser uma ontologia que define um recurso	(Ferreira, 2006; Márdero 2008; Suderman, 2010; Dos Santos, 2012; CCSDS, 2012)
Política	Conjunto de normas de política de acessos de um recurso	(Ardagna et al., 2010; Jin et al., 2011; Li et al., 2015)
Regra	Conjunto de preceitos que regulam o acesso a recursos	(Kiefer et al., 2012; Carvalho, 2013; Gomes, 2013; Bunawan and Nordin, 2015; Li et al., 2015; Riley, 2017; Gordon, 2018; Oliveira, 2019)
Atividade	Tarefas que devem ser realizadas para a preservação de dados clínicos.	(Carvalho, 2013; Völter, 2013)

5.2 ELABORAÇÃO DO METAMODELO

A abordagem de modelagem adotada nesta neste é baseada no MOF (*Meta Object Facility*) descrito no Capítulo 2 que serve como padrão para escrever metamodelos. Os elementos fundamentais do MOF incluem, por exemplo, definições para conceitos como Classe ou Propriedade aplicáveis à descrição de classes e seus atributos. O metamodelo proposto é formalizado no EMF, um framework de modelagem baseado no meta-metamodelo *Ecore*. A notação de classes do UML é adequada para expressar as entidades básicas do metamodelo como: classes, atributos e relações de generalização, associação regular (seta) e composição (losango cheio) (Galvão, 2018).

Em conformidade aos elementos conceituais identificados na Tabela 11 propõe-se um metamodelo sistêmico inicial em um alto nível de abstração que se encaixa na camada M2 do MOF. Este metamodelo visa representar a preservação de dados ao longo do ciclo de vida de um ser humano segundo aspectos de saúde ressaltando requisitos de preservação. O ciclo vital de um ser humano começa com a gestação, parto e puerpério, seguido da etapa da infância, posteriormente puberdade, adolescência, adultos jovens, maturidade, velhice e óbito (Eizirik and Bassols, 2013).

O metamodelo também é construído por meio de abstrações de modelos já presentes na literatura na forma de normativas internacionais (normas ISO) e padrões de referência na área médica e arquivologia, tais como: modelos de processos clínicos (ISO 13940, 2016; Behnam and Badreddin, 2013), modelo de referência OAIS (CCSDS, 2012), especificação ISO 13606 (Muñoz, 2011), políticas de acesso (Li et al., 2015) e ecossistema de dados (Oliveira, 2018) conforme foi apresentado na Tabela 10 no Capítulo 4.

As Figuras 18-22 apresentam o metamodelo proposto. Uma visão completa é mostrado na Figura 18 cujo código fonte *Ecore* no padrão XML (*Extensible Markup Language*) é disponibilizado no Apêndice A. As classes, atributos e relações são descritos em detalhe no Apêndice B. A Figura 19 mostra um fragmento do metamodelo responsável por representar eventos de saúde (*HealthEvent*) de uma pessoa durante seu ciclo de vida. Denominamos *HealthEvent* a todo evento que acontece na vida de uma pessoa e leva à produção de dados clínicos. Uma entrada por emergência, triagem, um procedimento médico são exemplos de instâncias de *HealthEvent*. Estes eventos geralmente acontecem dentro de uma organização ou entidade de saúde (i.e., hospitais, clínicas, consultórios) representado pela classe *HealthCareEntity*. Em consequência os eventos de nascimentos (*BirthEvent*) e de óbito (*DeadEvent*) são classes derivadas de *HealthEvent*.

A classe *HealthEvent* também visa abstrair processos clínicos e de saúde (*clinical pathways*) que podem estar compostos por outros processos derivados, ou seja, um evento de saúde pode conter outros sub eventos de saúde (em conformidade ao modelo de processos clínicos visto na seção 4.2.1). A auto relação *hasSubHealthEvent* da classe *HealthEvent* é essencial para modelar a complexidade dos processos de assistência à saúde. Como resultado, o metamodelo é capaz de capturar processos clínicos complexos, como tratamentos de quimioterapia, compostos por etapas derivadas mais detalhadas e essas etapas serão também instâncias de processos de *HealthEvent*.

A classe *Person* é um ator e visa representar uma pessoa (paciente) e seus relacionamentos familiares por meio das auto relações *hasParent*, *hasChildren* e a classe *Family*. Uma instância de pessoa deve ser identificada inequivocamente por meio de um atributo *ID* obrigatório que pode ser o número de documento nacional. Os dados demográficos de uma pessoa são modelados por meio da classe *DemographicData* e pode representar valores de contato ou residência que junto aos outros relacionamentos do metamodelo pode ser

utilizado para fins estatísticos por órgãos competentes e responsáveis pelos dados cartográficos e demográficos de um país, como por exemplo o IBGE no caso do Brasil.

Um evento de saúde pode ter múltiplas características (*Feature*). Uma característica representa um atributo ou aspecto que caracteriza um evento de saúde. Estas características, por sua vez podem ser derivadas em outras sub características que podem ser instanciadas, por exemplo, por taxonomias ou protocolos que descrevem um procedimento médico (*HealthEvent*).

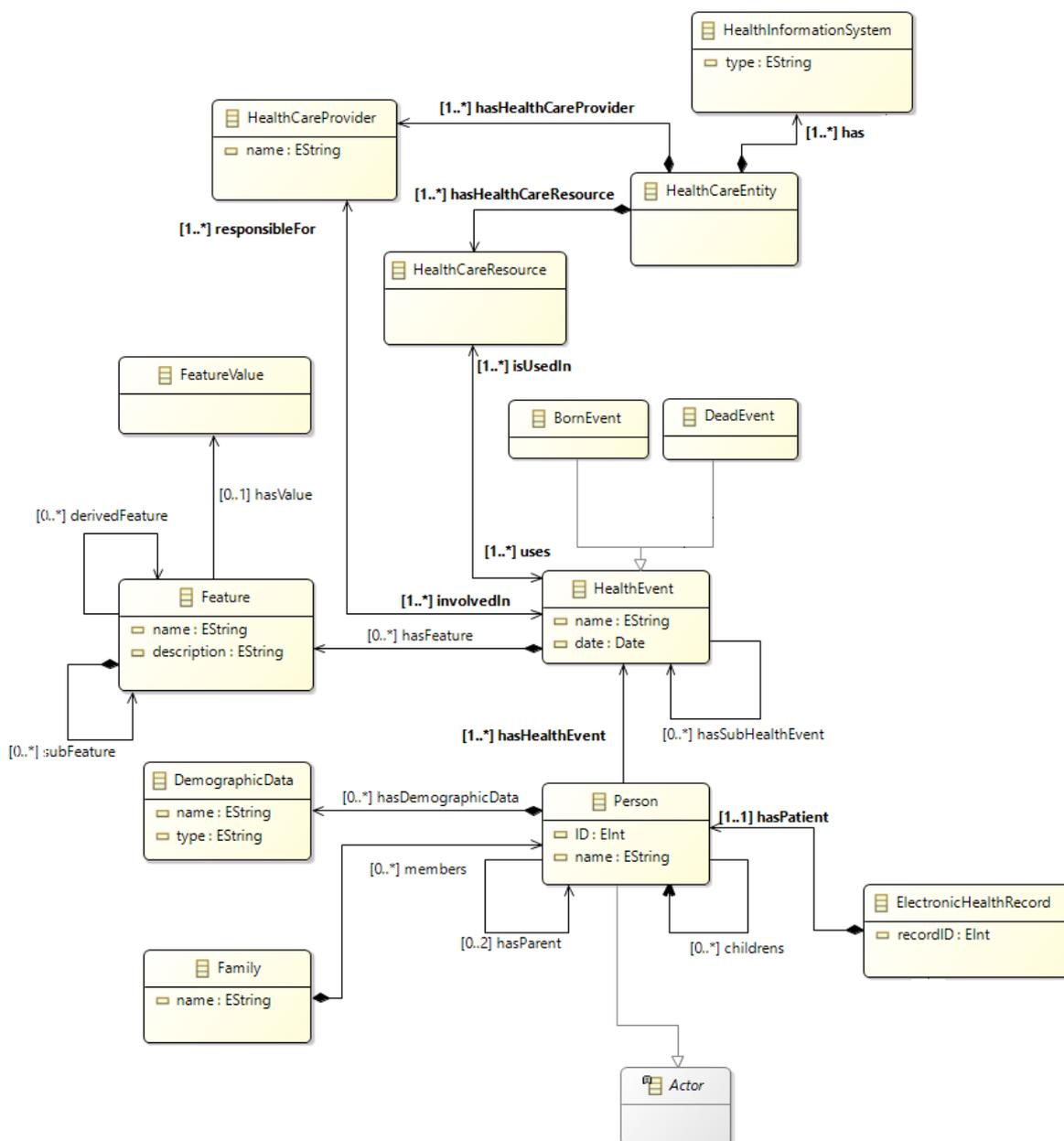


Figura 19: Fragmento do metamodelo relacionado a processos clínicos (Fonte: Autor)

Outras classes que fazem parte dos processos clínicos são *HealthCareProvider* que abstrai aos profissionais da área da saúde que atuam em um procedimento médico (*HealthEvent*). Os recursos que são necessários em um processo de atendimento ao paciente são capturados como *HealthCareResource*, estes recursos podem ser equipamentos médicos ou infraestrutura. As classes *HealthCareProvider* (prestadores de cuidados médicos), *HealthCareResource* (recursos para cuidados médicos) e *HealthInformationSystem* (sistemas de informação de saúde) compõem uma entidade de assistência médica (*HealthCareEntity*).

O prontuário eletrônico do paciente é representado pela classe *ElectronicHealthRecord* (em conformidade à modelo ISO 13606 visto na seção 4.2.2) e está associado a uma pessoa por meio de um identificador único. Este prontuário está composto por várias entidades que são mostradas na Figura 20. A classe abstrata *Record_Component* representa as características comuns de qualquer informação registrada no prontuário. Esta classe possui classes derivadas como *Folder*, *Content* e *Item* que modelam diferentes aspectos da estrutura do PEP.

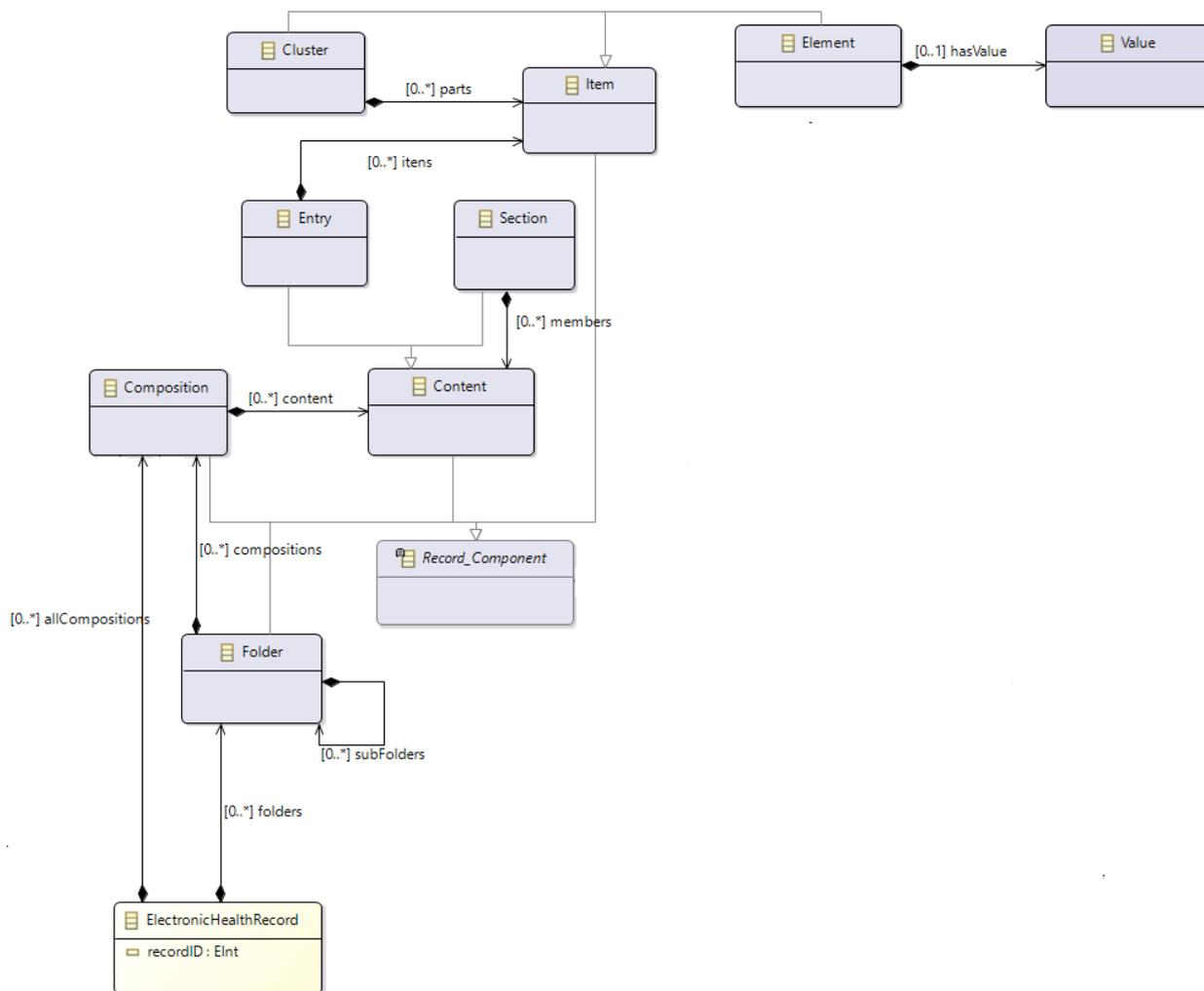


Figura 20: Fragmento do metamodelo relacionado ao PEP (Fonte: Autor)

A classe *Folder* corresponde a um nível de organização do PEP que serve para agrupar a informação em episódios clínicos, períodos em função da organização ou serviços clínicos (Muñoz et al., 2011). Exemplo de instâncias são: obstetrícia, atenção de diabetes, pediatria, atenção primária, etc. Esta classe pode gerar uma hierarquia por meio da relação de composição *subFolder* de modo a corresponder à organização do PEP.

Um documento clínico é modelado por meio da classe *Composition* que agrupa um conjunto de informações geradas dentro do PEP e enviadas por uma instância da entidade *HealthInformationSystem* no âmbito de um evento de saúde (Muñoz et al., 2011). Uma instância de *Composition* está formada por conteúdos representados pela classe *Content* que por sua vez é uma generalização da classe *Section* e *Entry*. A classe *Section* visa agrupar dados do PEP que, dentro de uma composição, se encontram dentro de um mesmo cabeçalho e que correspondem a um único processo de captura de informação ou para facilitar sua compreensão. Exemplos de instâncias são: Razão da visita, antecedentes clínicos, história familiar, alergias, etc. A informação que é registrada em um PEP como resultado de uma ação clínica, uma observação, uma interpretação clínica ou um propósito clínico é modelada pela classe *Entry*. Exemplos de instâncias de *Entry* são: um sintoma, uma observação, uma prova, uma prescrição, uma reação alérgica, etc.

A classe *Cluster* suporta a representação de estruturas de dados complexas necessárias para representar os valores de dados reais em uma observação multipartida, uma declaração clínica, ou uma instrução (Muñoz et al., 2011). Um Cluster é adequado para representar séries temporais, tabelas ou listas de dados. Exemplos de instâncias são: Resultado de um audiograma, interpretação de um encefalograma, tabela de resultados, lista de medicamentos, etc.

A classe *Element* serve como recipiente para os nodos folha da hierarquia do PEP. Cada ocorrência desta classe terá um único tipo de dado, como uma cadeia de texto, um número inteiro ou um elemento multimídia. Exemplos de instâncias desta classe são: Pressão sanguínea sistólica, pulso, nome de um medicamento, peso corporal, etc.

A Figura 21 mostra o fragmento do metamodelo referente à preservação dos dados clínicos gerados nos PEPs. A classe *ContentInformation* (em conformidade à modelo de referência OAIS visto na seção 4.2.3) representa objetos digitais a serem preservadas. Estes objetos digitais são informações extraídas do PEP representadas na classe *Composition* e suas classes derivadas (*Section*, *Entry*, *Cluster*, *Item* e *Element*). Um exemplo de instância da classe

ContentInformation é um laudo médico, este laudo pode ser preservado em um formato determinado (e.g. pdf, xml). Associado a este laudo existem um conjunto de metadados que o descrevem (classe *Metadata*) com seus respectivos valores (*MetadataValue*).

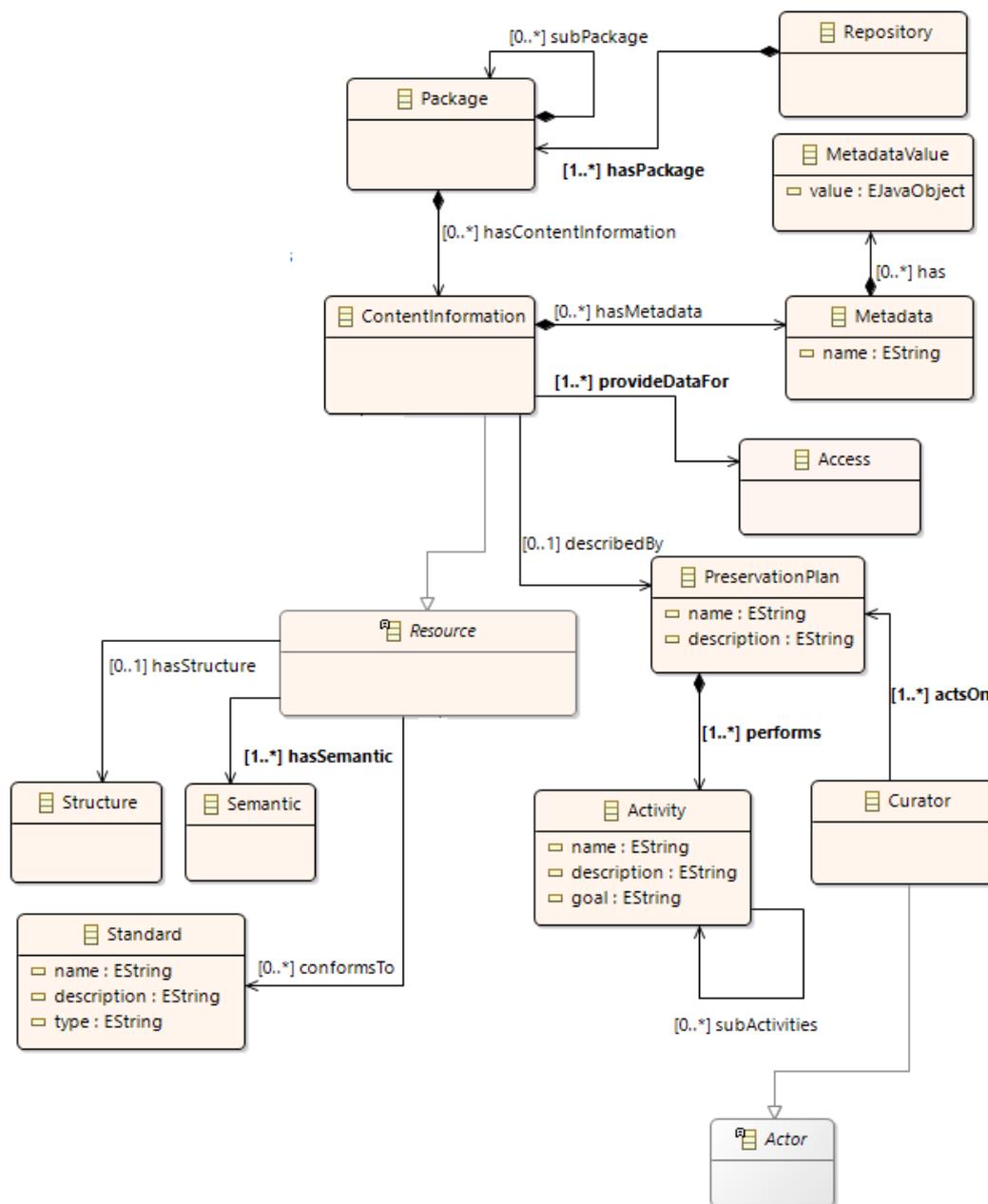


Figura 21: Fragmento do metamodelo relacionado à preservação de dados (Fonte: Autor)

Os valores dos dados a serem preservados são definidos por meio de um plano de preservação realizado por uma equipe de curadoria (classe *Curator*). Instâncias de *Curator* podem ser profissionais da área de saúde, arquivistas, pesquisadores, entidades governamentais, etc. As instâncias de *Curator* definem um conjunto de estratégias e atividades de preservação. As classes *PreservationPlan* e *Activity* visam abstrair este processo. Os objetos digitais

representados por *ContentInformation* são agrupados em pacotes (classe *Package*), que por sua vez são organizados em hierarquias mediante a relação *subPackage* e fazem parte de um repositório (*Repository*).

A classe *Access* representa um software que permite aos consumidores localizar, analisar, encomendar ou recuperar informações clínicas preservadas. Já a classe *Resource* é uma generalização da classe *ContentInformation* e está em conformidade com um padrão (*Standard*) que pode representar por exemplo padrões de interoperabilidade semântica como HL7 CDA. Um recurso também possui uma estrutura (*Structure*) e uma semântica que o define (*Semantic*). Um exemplo de instância de estrutura pode ser uma definição de XML *Schema*. O modelo de referência do HL7 RIM pode ser uma instância de semântica.

A classe *Resource* também está composta por um conjunto de regras de acesso (representado pela classe *Rule*) (Figura 22). Estas regras fazem parte de políticas de acesso que são diretrizes ou regulamentos práticas que prescrevem como uma organização gerencia, protege e distribui informações (conforme visto na seção 4.2.4). As definições de Políticas são representadas pela classe *Policy*. Um exemplo de instância de *Policy* são políticas de segurança que define os níveis de privacidades dos dados clínicos que estão sendo preservados.

Uma regra de acesso (*Rule*) está formado por um indivíduo (*Subject*) que representa uma entidade que solicita permissão de acesso a um recurso. Instâncias da classe *Subject* possuem uma credencial (*Credential*) e cumprem um papel (*Role*), por exemplo: médico, assistente social, familiar, etc. Uma regra também possui características (*RuleFeature*) que podem abstrair um conjunto de ações que pode ser realizado no recurso (e.g. leitura, escrita), metas ou informações de contexto (Li et al., 2015).

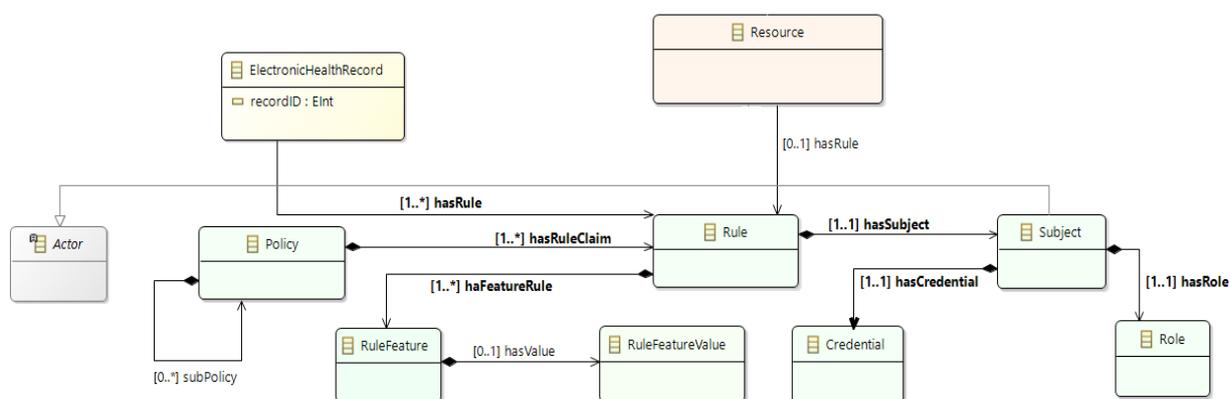


Figura 22: Visão do metamodelo relacionado às políticas de acesso (Fonte: Autor)

As políticas de segurança centrada no paciente e na assistência médica têm sido um tópico de pesquisa popular nos últimos anos (Fatima and Colomo-Palacios, 2018; Bheemanathini, 2018).

5.3 CONSIDERAÇÕES FINAIS

Apesar dos metamodelos serem muito utilizados para descrever diferentes domínios, a sua construção não é uma tarefa trivial (Oliveira, 2019). Para construir o metamodelo proposto foi seguido uma abordagem iterativa. Cada versão obtida do metamodelo era submetido a um processo de validação, cujos detalhes são descritos no Capítulo 6. Caso alguma condição não fosse satisfeita, se repetia o processo de identificação de conceitos, classes, propriedades e relações para que se ajustem melhor a diferentes cenários. Foi justamente por este motivo que se decidiu utilizar modelos já estabelecidos na literatura como padronizações, normativas ISO que serviram como base de construção. Apesar do metamodelo juntar modelos existentes e abstraí-los a um nível superior, se tentou manter a nomenclatura das classes e relações, de modo a facilitar seu uso posterior.

A linguagem utilizada é o UML, pois é suficiente para entender o conhecimento relacionado à preservação de dados clínicos sob um ponto de vista estático (Guedes, 2018). Não foi utilizada uma linguagem temporal para meta modelagem, pois não se pretende instanciar modelos temporais que representem a evolução de um modelo nem históricos de versões de modelos (Gómez and Wimmer, 2018). De acordo à literatura (Bill et al., 2018, Gómez and Wimmer, 2018) modelos temporais podem ser submetidos a consultas temporais de forma a recuperar o conteúdo do modelo em diferentes pontos do tempo, para realizar isto deve-se procurar uma versão do modelo, que geralmente está armazenada em algum sistema de versionamento externo, correspondente ao período desejado.

Finalmente, é importante salientar que o metamodelo considera que o paciente não é responsável direto pelo gerenciamento do PEP. Este é gerenciado por uma entidade externa como uma entidade de saúde ou instituição que seja nomeado pelo governo.

CAPÍTULO 6 – VALIDAÇÃO DO METAMODELO

Neste capítulo é realizado a validação do metamodelo proposto por meio da instanciação de modelos que abstraem diferentes cenários fictícios. Esta etapa corresponde ao penúltimo passo da metodologia empregada nesta tese. Aqui procedemos a validar o metamodelo por meio da instanciação de modelos que devem responder a um conjunto de questões de competência (Gruninger and Fox, 1995). Estes modelos são criados baseados em diferentes cenários que expressam os conceitos específicos do domínio que está sendo estudado. Caso a validação não seja satisfatória o metamodelo é refinado por um processo de retroalimentação.

6.1 INSTANCIAÇÃO DE MODELOS

A Figura 23 ilustra o processo de instanciação de modelos a partir de um nível superior do MOF. Podemos observar que as classes *HealthEvent* e *Feature* que fazem parte do metamodelo proposto pertencem ao nível 2 e são instâncias da classe *EClass* do meta metamodelo *Ecore*. O processo de validação utilizado nesta tese consiste em criar modelos no nível 1. Como podemos ver na Figura 23, a classe *ConsultaRotina* é instância de *HealthEvent* e, por sua vez, as classes *Anamnese* e *PrescriçãoMedica* são instâncias de *Feature*. Uma instância de nível 0 seria um objeto do mundo real, por exemplo um documento que represente uma consulta médica.

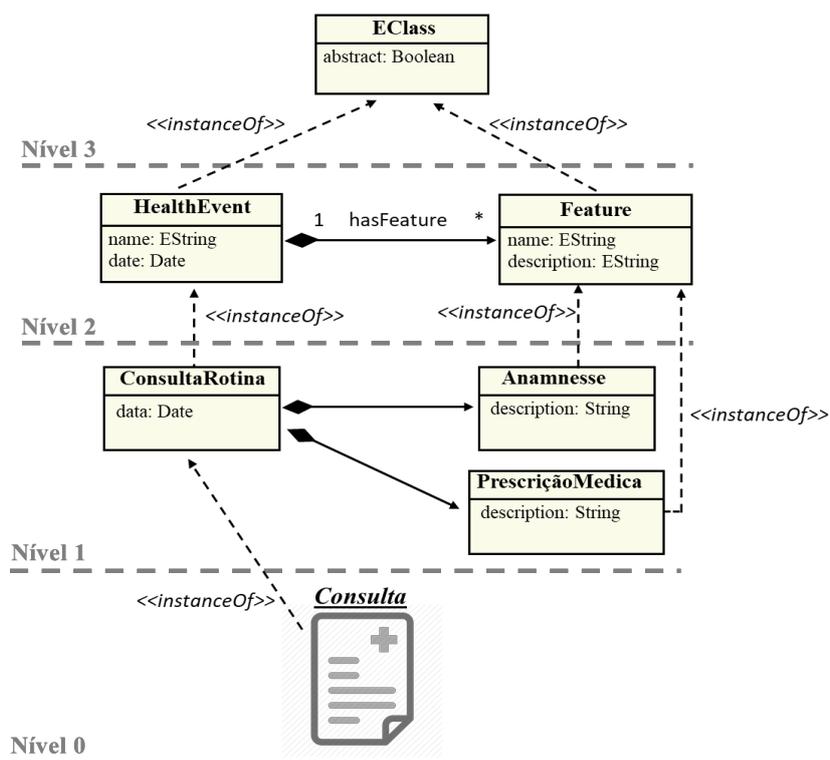


Figura 23: Exemplo de instanciação de modelos. (Fonte: Autor)

6.2 TÉCNICAS DE VALIDAÇÃO DE METAMODELOS

Para validar metamodelos, é comum observar na literatura (Vieira, 2016; Oliveira, 2018) as seguintes técnicas: (1) Construção de ferramentas de modelagem, (2) aplicação em caso de estudo e (3) abordagem híbrida. De acordo a um estudo de revisão da literatura realizada por Vieira (2016), de um total de 97 estudos analisados, apenas 4 estudos utilizaram a primeira técnica, 40 estudos optaram pela aplicação em um estudo de caso e 53 estudos utilizaram a abordagem híbrida, que junta as primeiras, conforme podemos ver na Figura 24.

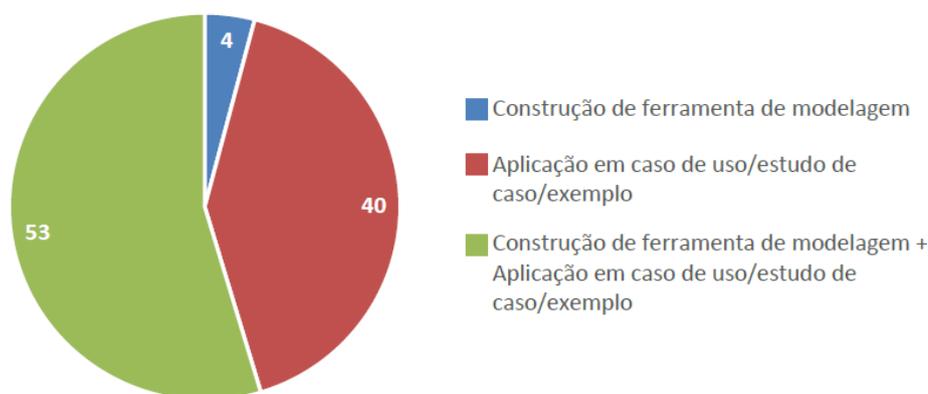


Figura 24: Visão geral das formas de validação de metamodelos. (Fonte: Alves, 2016)

A primeira técnica de validação consiste na construção de uma ferramenta de modelagem que permita instanciar modelos conforme o metamodelo proposto. A técnica por estudo de caso consiste na instanciação de modelos a partir do metamodelo para resolver um ou mais cenários reais ou fictícios. Já a validação híbrida junta as anteriores (Favre, 2004; Steinberg et al., 2008; Durak et al., 2014; López-Fernández et al., 2015). Nesta tese é utilizada a abordagem por estudos de caso para validação.

6.3 DEFINIÇÃO DE QUESTÕES DE COMPETÊNCIA

A elaboração de questões de competência é uma estratégia derivada da área de desenvolvimento de ontologias (Gruninger and Fox, 1995, Cristani and Cuel, 2005). Basicamente consiste na criação de perguntas que auxiliem na especificação dos requisitos da ontologia a ser desenvolvida. Estas questões são então formalizadas de tal maneira que possam ser executadas frente a um conjunto de instâncias do metamodelo, de modo a verificar a sua conformidade em relação às especificações.

Com o objetivo de validar o metamodelo proposto, nesta tese é formulado um conjunto de questões de competência que compreendem cenários fictícios para modelagem de processos

clínicos, modelagem de PEP e a sua preservação, isto em conformidade aos objetivos da tese. As respostas às questões são feitas tanto na forma de explicações textuais quanto na forma de Diagramas de Objetos da UML. O uso de Diagramas de Objetos é muito utilizado na literatura (Carvalho, 2013) e traz a possibilidade de representar cada modelo (e.g., Modelo de PEP) por meio da identificação das classes e das associações necessárias à sua realização. A notação de objetos representa um subconjunto do Diagrama de Classes da UML já que o Metamodelo é por sua vez é um modelo EMF (Steinberg et al., 2008).

A seguir são formuladas as questões de competências que estão diretamente focadas nos problemas que se deseja solucionar, tendo como base os objetivos do metamodelo.

- Q1: O metamodelo permite instanciar modelos que representem processos clínicos?
- Q2: O metamodelo permite instanciar modelos que representem o PEP?
- Q3: O metamodelo permite instanciar modelos que representa a preservação de documentos clínicos?
- Q4: O metamodelo considera a recuperação de documentos clínicos preservados?
- Q5: O metamodelo considera políticas de acesso dos documentos clínicos que foram preservados?

Na sequência definimos três cenários fictícios: Sofia, João e Maria. O cenário de Sofia representa uma visão geral dos eventos de saúde de uma pessoa, desde o nascimento até o óbito. A partir deste cenário genérico, três sub cenários são selecionados para serem modelados visando responder as questões de competência. Já os cenários de João e Maria representam casos mais específicos que o metamodelo também deverá responder.

6.4 CENÁRIO 1

Sofia nasceu no hospital de Niterói (BE – Born Event) com sete meses e meio. Ela estava com o peso abaixo do normal e problemas respiratórios, motivo pelo qual teve que ficar em observação (HE₁ – Health Event) durante uma semana. Durante os primeiros seis meses, Sofia, além de tomar todas as vacinas necessárias, teve problemas de deglutição que foram tratados pelo pediatra do hospital (HE₂).

Quando Sofia cumpriu seis anos de idade, teve que ser levada para um pronto socorro (HE₃) devido a um resfriado que se estava prolongando. Com oito anos de idade sofreu uma

queda onde fraturou o braço e teve que ser encaminhada para uma clínica da cidade (HE₄) onde foi tratada satisfatoriamente.

Após uma vida tranquila, Sofia se casou aos 30 anos e logo engravidou e fez o pré-natal em uma clínica privada da cidade, onde realizava suas consultas de rotina. Após realizar uma viagem, começou a ficar enjoada todos os dias, portanto decidiu antecipar a consulta com sua obstetra (HE₅). Quando cumpriu 37 semanas de gestação nasceu Maria Clara por meio de uma cesárea (HE₆).

Na meia idade, Sofia descobriu um caroço na mama esquerda, preocupada, procurou por uma mastologista em um consultório particular (HE₇). A mastologista solicitou Sofia realizar uma mamografia (HE₈). O resultado foi um fibroadenoma benigno o qual foi extraído com sucesso (HE₉). Após uma consulta ao histórico clínico detectou-se que a avó materna teve câncer de mama, por esse motivo Sofia passou a ser monitorada semestralmente.

Anos depois, Sofia ingressou por emergência (HE₁₀) ao hospital da cidade, pois ela começou a sentir uma súbita e intensa dor de cabeça, seguido de um quadro de vômitos e perda de consciência. Assim, uma consulta rápida (usando técnicas de mineração de dados) ao histórico do paciente no sistema de informação de saúde unificado, os médicos conseguiram ver que Sofia tinha um histórico familiar de doenças vasculares. Considerando essa informação, os especialistas rapidamente identificaram que Sofia estava passando por um rompimento do aneurisma, o que possibilitou o tratamento e salvar a vida dela.

Pouco tempo depois Sofia teve um acidente vascular cerebral (AVC), apenas deu entrada na emergência e sem profissional disponível na área veio ao óbito pela demora no atendimento (DE – Death Event).

Para melhor ilustração podemos resumir o cenário de Sofia na Figura 25 por meio de uma linha temporal, identificando os eventos de saúde (HE - *HealthEvent*). Sofia ao longo da sua vida utilizou diferentes provedores de saúde (rede pública e privada) com diversas infraestruturas. Mas, é desejável que toda informação clínica de um indivíduo, gerada desde a emissão da declaração de nascido vivo²⁰ até o óbito²¹, esteja disponível e seja acessível pelos diferentes tipos de SIS. De acordo com Burns (2011), o acesso a essa informação pode ajudar

²⁰ Segundo a lei 12.662/12, a Declaração de Nascido Vivo é um documento provisório de identificação do recém-nascido, tendo eficácia em todo o território nacional até o registro do nascimento em cartório. Este documento é emitido pelo SINASC (Sistema de Informações sobre Nascidos Vivos)

²¹ http://bvsmis.saude.gov.br/bvs/publicacoes/manual_declaracao_obitos.pdf (acessado em 20/11/2019)

a melhorar a assistência ao paciente, reduzindo custos de cuidados de saúde, contar com um melhor acompanhamento dos resultados, beneficiando a consistência do tratamento e aumentando significativamente a informação disponível para a tomada de decisões (Burns, 2011; Misra et al., 2012; Bunawam and Nordin, 2015). Observe-se que no HE₁₀ o acesso à informação do histórico do paciente possibilitou a detecção precoce de uma doença e facilitou o seu tratamento, isto pode ser possível graças à preservação de PEPs.

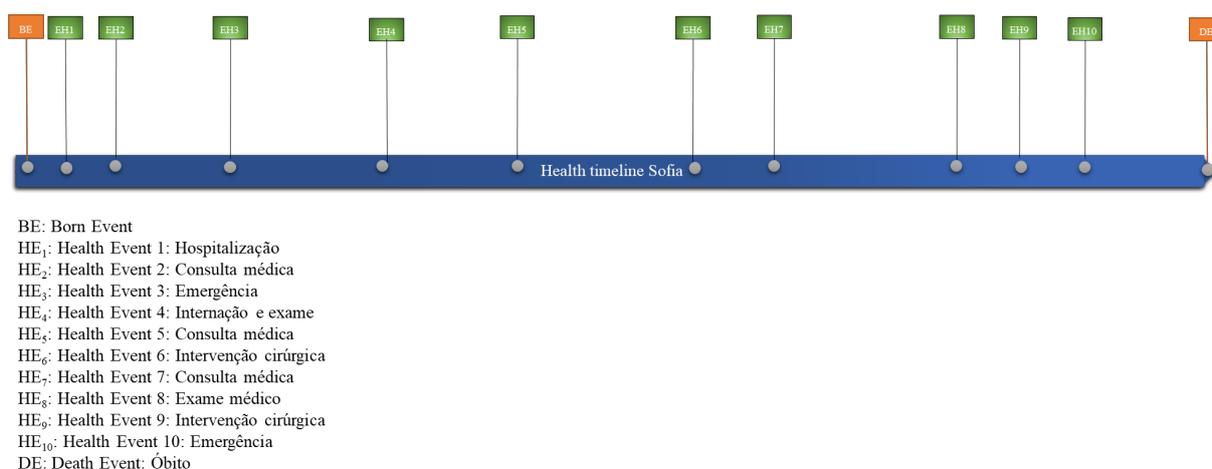


Figura 25: Linha do tempo de saúde de Sofia. (Fonte: Autor)

6.4.1 MODELOS DO CENÁRIO 1

Para responder à questão de competência Q1 consideremos o evento de saúde HE₃ (entrada por emergência) de nosso cenário fictício. A seguir vamos a detalhar esse cenário de modo a facilitar a sua modelagem.

Quando Sofia entrou por emergência por causa um resfriado, passou por triagem, onde uma enfermeira fez um levantamento de dados e realizou algumas medições da temperatura e peso da paciente. Logo na sequência foi encaminhada para uma pediatra. A especialista coletou dados e seguiu o protocolo estabelecido para casos similares e a encaminhou para realização de um exame de raios-x pulmonar. Já com os resultados do exame, a especialista a detectou um quadro de pneumonia e indicou o tratamento adequado.

O metamodelo proposto deve atender ao requisito de representação dos eventos clínicos produzidos em decorrência da emergência inicial do cenário descrito. A Figura 26 mostra uma instância do metamodelo que é gerada a partir da ocorrência do evento de saúde *EmergencyAdmission* que por sua vez gera os sub eventos triagem (classe *Triage*) e consulta

(classe *Consultation*). A consulta médica por sua vez gera o evento de exame médico (classe *MedicalExamination*). Após a conclusão dos eventos clínicos anteriores é gerado o evento de alta (*Discharge*).

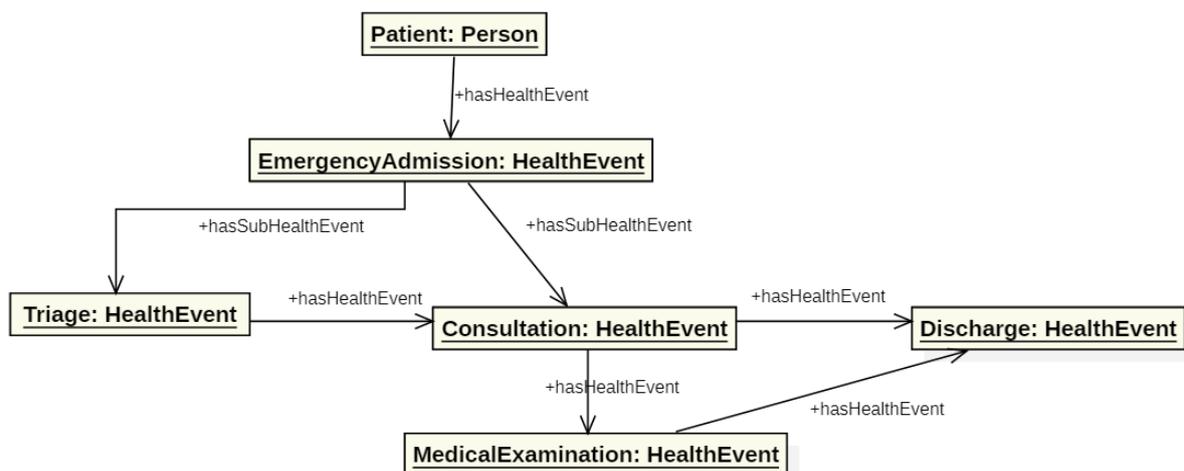


Figura 26: Modelo de eventos clínicos (Fonte: Autor)

Para descrever de melhor forma o processo de atendimento do cenário proposto, consideremos a Figura 27 que apresenta uma sequência de eventos que mostra os dados clínicos que são gerados. Também é possível observar que em cada evento são geradas instâncias da classe *HealthCareProvider* e *HealthCareResource*. Estas representam respectivamente o profissional de saúde, que participa do processo e o recurso que está sendo utilizado no ambiente hospitalar, que por sua vez é uma instância de *HealthCareEntity*.

Observe-se que a Q1 pede para instanciar um modelo que mostre os processos clínicos e um processo clínico é um conjunto de eventos clínicos ou de saúde e a Figura 26 retrata essa instância.

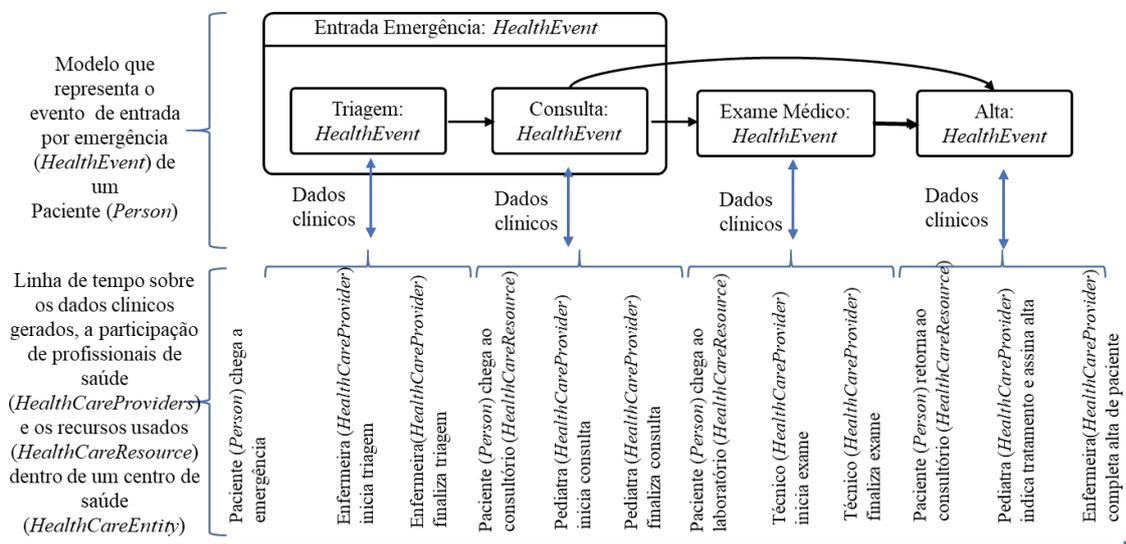


Figura 27: Sequência de eventos de saúde. (Fonte: Autor)

Para responder à questão de competência Q2 vamos detalhar o evento HE₅ do nosso cenário fictício principal como segue.

No dia 05/12/2005 às 13:05, Sofia chega ao serviço de ginecologia da maternidade São Francisco em Niterói para realizar uma revisão. Ela se encontra na semana 28 de gravidez e comenta a sua ginecologista que nos últimos dias se sentia enjoada todo o tempo. Além disso, comenta seus problemas de acidez estomacal. A ginecologista mede sua pressão sanguínea, obtendo 100/60 mmHg. Finalmente, sua médica realiza um exame abdominal no feto, confirmando que se apresenta em posição longitudinal cefálica. Além disso, realiza a medição de seu pulso, que é de 140 pulsações por minuto.

Conforme foi mostrado na Figura 26, cada evento de saúde pode gerar um conjunto de dados clínicos que deverão fazer parte de um documento clínico, que por sua vez, fará parte do PEP do paciente.

Para modelar o cenário de visita de pré-natal em termos de PEP são instanciadas as classes *Folder*, *Composition*, *Entry*, *Section* e *Element* do metamodelo proposto. A Figura 28 mostra o modelo resultante onde é possível observar que em todos os casos as relações são de composição. A classe *Folder* organiza as outras classes em uma hierarquia, por exemplo a classe *BloodPressure* está composto pelos elementos *Diastolic* e *Systolic*. A classe *Folder* por sua vez é parte da classe *ElectronicHealthRecord*.

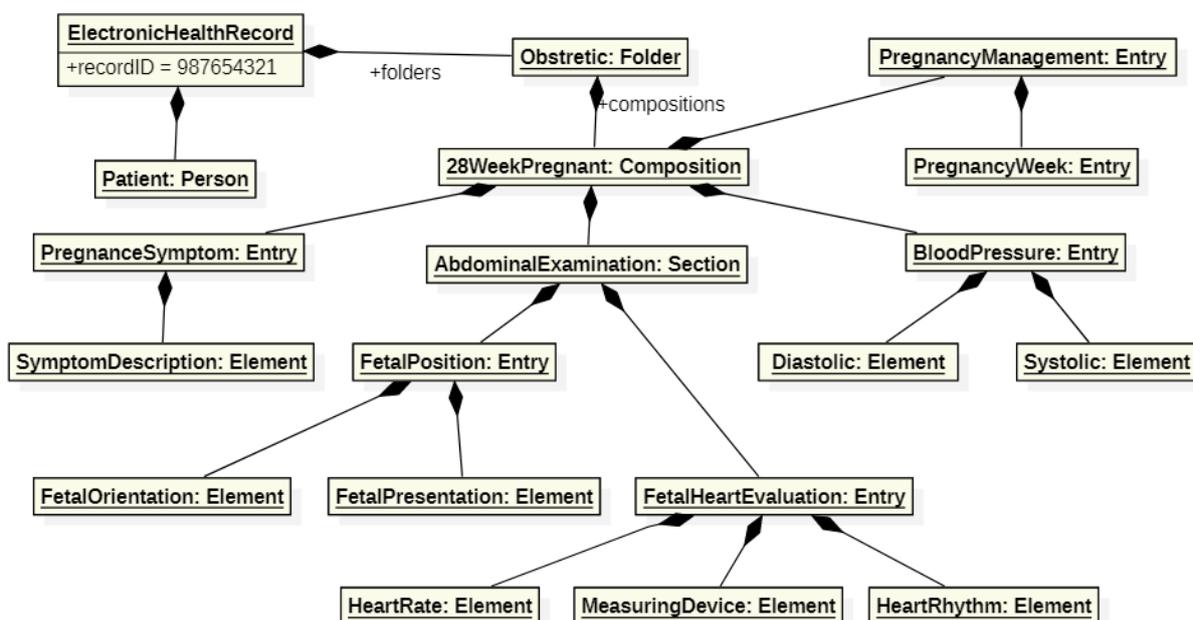


Figura 28: Modelo de prontuário instanciado (Fonte: Autor)

A Figura 29 mostra um exemplo de uma estrutura do documento clínico (visita de pré-natal/consulta médica) que corresponde ao modelo instanciado. É importante mencionar que os valores que são atribuídos às instâncias da classe *Element* devem estar em conformidade com um padrão, como por exemplo o SNOMED para usar um vocabulário médico adequado.

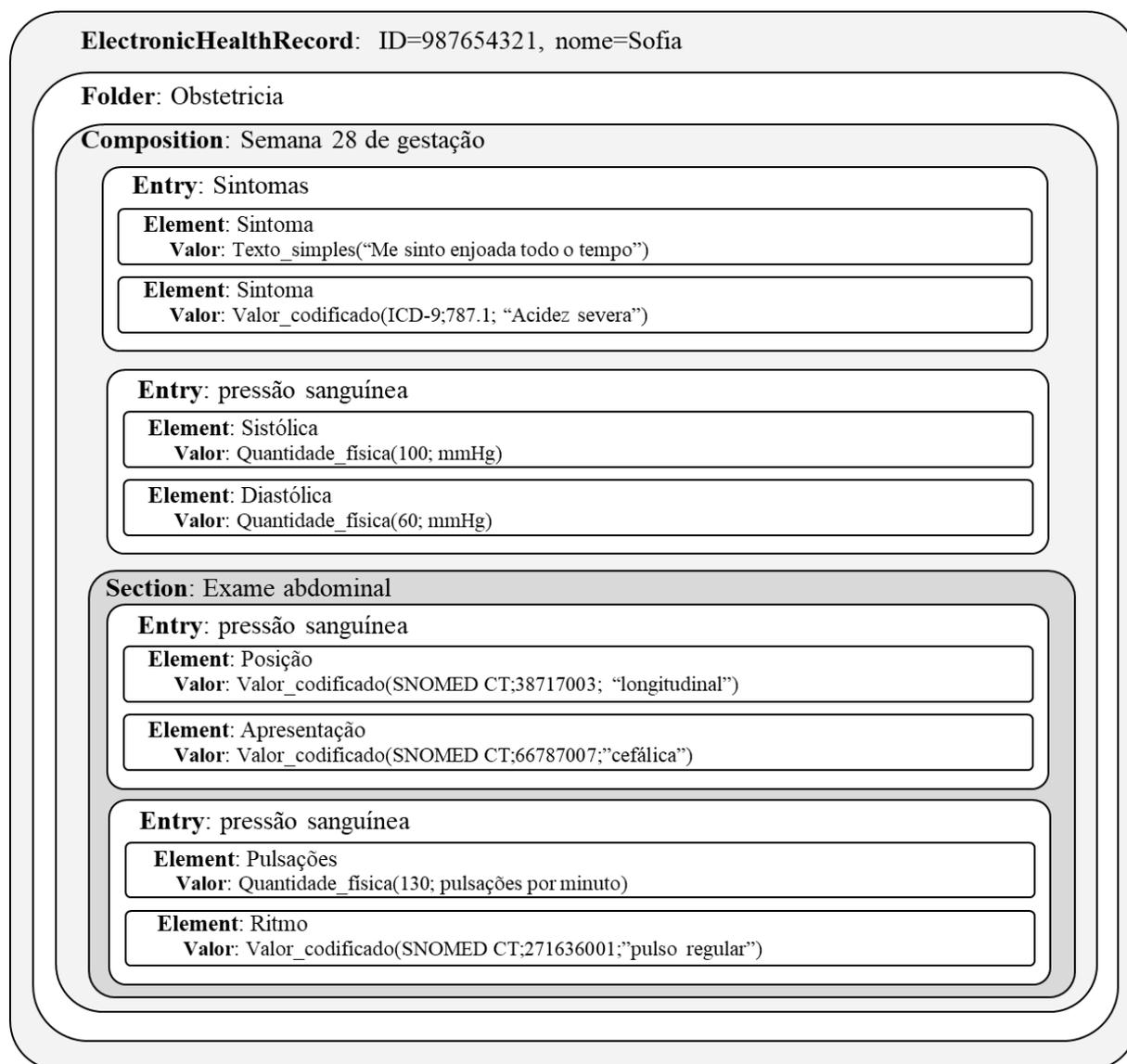


Figura 29: Estrutura de um documento clínico. Adaptado de (ISO13606)

Observe-se que a Q2 é respondida satisfatoriamente pelo modelo da Figura 28, pois ela representa por sua vez uma instância do modelo de referência ISO 13606 que abstrai documentos clínicos que fazem parte do PEP.

Para responder as questões Q3 e Q4 vamos considerar um plano fictício de preservação de dados clínicos baseado no e-ARQBrasil ao qual chamaremos de *eHealthARQBrasil* que será

uma instância da classe *PreservationPlan*. Este plano é inspirado e fundamentado no trabalho de Piconi (2014) que propõe uma ontologia para representar os metadados necessários para assegurar a preservação digital de PEPs.

O e-ARQ Brasil, elaborado pela CTDE do CONARQ, é uma especificação de requisitos a serem cumpridos pela organização produtora/recebedora de documentos arquivísticos digitais, pelo sistema de gestão arquivística e pelos próprios documentos, a fim de garantir sua confiabilidade e autenticidade, assim como sua acessibilidade ao longo do tempo, desde a produção, tramitação e utilização até a sua destinação final, ou seja, arquivamento permanente ou eliminação (CONARQ²², 2011). Embora o e-ARQ Brasil seja de propósito geral e não defina metadados específicos para preservação de PEPs, o trabalho de Piconi (2014) propõe um mapeamento dos metadados do CDA com metadados selecionados do e-ARQ Brasil. Conforme descrito no Capítulo 2, o CDA é um padrão de marcação de documento que especifica a estrutura e a semântica de documentos clínicos com o objetivo de interoperar dados do paciente, podendo conter texto, imagens, sons e outros conteúdos multimídia (HL7, 2019).

Para responder a Q3 consideremos que nosso plano de preservação consiste em usar uma representação CDA dos documentos clínicos como objeto arquivístico a ser preservado. Para isso criamos a entidade *CDADocument* como sendo instância da classe *ContentInformation*. Este documento está descrito por um conjunto de metadados preestabelecidos representado pela classe *PreservationMetadata* que é instância da classe *Metadata*. Os metadados de preservação são dados semânticos que ajudam e documentam a preservação de documentos digitais ao longo do tempo (Sayão, 2013). Além de preservar somente o documento digital (em nosso caso um arquivo CDA no formato XML), é necessário preservar também os metadados que descrevem o ciclo de vida desse documento, como histórico de criação, alteração e contextos tecnológico (hardware e software) desde o momento de sua produção até a sua destinação final, todos estes valores instâncias da classe *MetadataValue*. A Figura 30 mostra o modelo resultante de preservação, sendo uma instância criada em conformidade com nosso metamodelo.

²² Conselho Nacional de Arquivos <http://conarq.arquivonacional.gov.br/documentos-eletronicos-ctde.html> (Acessado em: 25/11/2019)

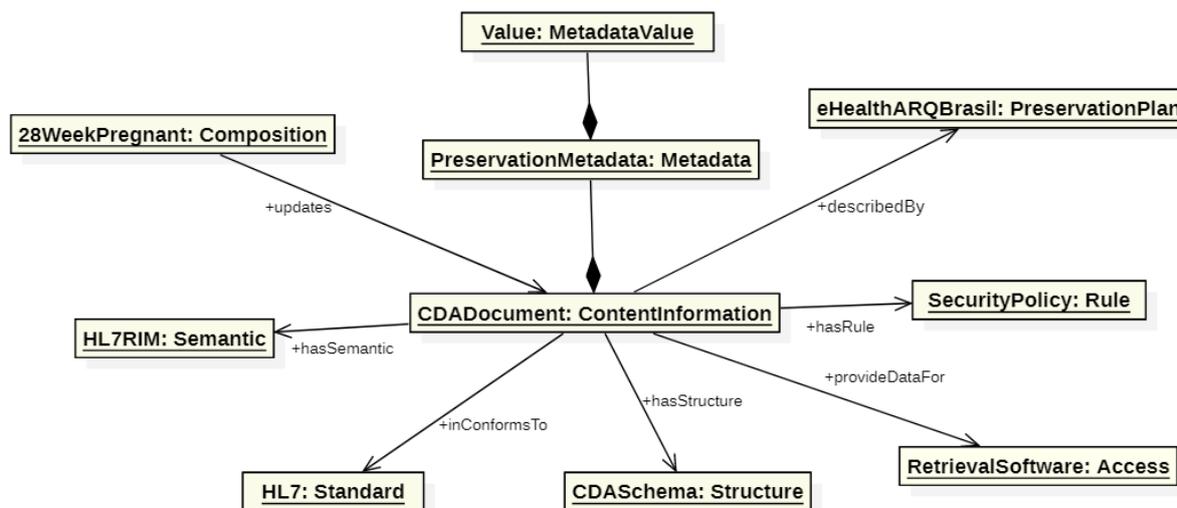


Figura 30: Modelo de preservação de um documento clínico (Fonte: Autor)

Se consideramos nosso cenário fictício da história de Sofia, o evento de saúde consulta na semana 28 de gestação, representado pela classe *28WeekPregnant*, gerará um documento CDA por meio da relação *updates* que corresponde ao processo de ingestão no modelo de referência OAIS. A Figura 31 mostra um fragmento de um documento CDA fictício que ajudará a esclarecer a ideia. Este arquivo está codificado em XML e adota as definições semânticas do HL7 RIM, permitindo assim a reutilização de dados, expressividade e flexibilidade (HL7, 2019). Em nosso modelo, o padrão HL7 seria instância da classe *Standard*, o modelo de referência HL7 RIM seria instância da classe *Semantic* e o esquema que especifica o CDA seria instância da classe *Structure*. Todas estas classes também serão úteis para disponibilizar o documento preservado para sua recuperação ou exibição por um software especializado para esse propósito (instância da classe *Access*) permitindo que os consumidores possam localizar as informações de interesse. Desta forma, o modelo instanciado consegue responder a questão Q4.

É importante salientar que o nível de abstração considerado no metamodelo proposto permite que o padrão de metadados e de estrutura semântica, considerados neste cenário fictício, sejam substituídos por qualquer outro disponível. Os procedimentos adotados são genéricos, sendo possíveis de serem realizados em qualquer outro padrão.

```

<title>Maternidade São Francisco</title>
<effectiveTime value="20000407125632+0200"/>
<confidentialityCode code="R" codeSystem="2.16.840.1.113883.5.25"/>
<languageCode code="pt-BR"/>
<versionNumber value="2"/>
<recordTarget typeCode="RCT">
  <patientRole>
    <id extension="950034253" root="2.16.840.1.113883.2.4.6.3" />
    <id extension="12345" root="2.16.840.1.113883.19.5"/>
    <addr>
      <streetName>Av. Visconde</streetName>
      <houseNumber>123</houseNumber>
      <postalCode>24020000</postalCode>
      <city>Niterói</city>
    </addr>
    <telecom use="HP" value="tel:+3110-23423423"/>
    <patient>
      <name use="L">
        <given>Sofia</given>
        <family>Lima</family>
        <suffix>SF</suffix>
      </name>
      <birthTime value="1975924"/>
    </patient>
  </patientRole>
</recordTarget>

```

Figura 31: Exemplo de um fragmento de arquivo CDA

6.5 CENÁRIO 2

Uma sexta-feira pela noite, João, uma criança (com menos de 12 anos) sofreu um acidente, teve uma perna quebrada e algumas contusões no corpo (indicando que ele pode ter sido abusado). Ele é levado por ambulância para um hospital, e sua mãe o acompanha. Ao chegar ao hospital, ele é levado para a sala de emergência e é atendido por médicos de pronto-socorro.

Inicialmente, o médico de emergência registra uma história, preenche o PEP da criança, atribui uma enfermeira e solicita uma série de exames. Quando o médico de emergência recebe os resultados dos exames, ele suspeita que o filho tenha sido abusado. Em consequência, a polícia e os serviços sociais são informados do possível abuso.

Nesse contexto, a polícia, responsável pela investigação do abuso, solicita acesso aos dados do PEP da criança para fins de investigação. O pedido de acesso à polícia é permitido porque foi validada uma regra de autorização previamente definida no hospital. Mas os policiais precisam cumprir suas obrigações informando ao hospital os resultados da investigação.

Ao mesmo tempo, o assistente social, responsável por ajudar a criança, também solicita acesso aos dados dela, o qual é permitido graças a validação de uma regra também previamente definida.

O cenário descrito acima é uma adaptação de um estudo de caso de privacidade de dados clínicos desenvolvido por Li et al. (2015) e visa responder a questão de competência Q5. Nesse cenário, podemos observar vários atores envolvidos, cada um cumprindo um papel diferente. Cada ator possui requisitos de permissões para acessar os dados do PEP do hospital. Esses atores incluem: (1) a criança que precisa de tratamento; (2) os pais da criança; (3) os médicos de emergência; (4) os enfermeiros; (5) os assistentes sociais que são responsáveis por ajudar a crianças em caso de abusos; e (6) os policiais que são responsáveis por investigar o possível crime de abusos.

6.5.1 MODELO DO CENÁRIO 2

A Figura 32 mostra um modelo gerado a partir do metamodelo proposto que abstrai as entidades descritas no cenário descrito acima. A classe *ElectronicHealthRecord* representa o PEP e tem como parte componente um paciente (instância de classe *Person*). O PEP está associado a um conjunto de regras (*PolicyRule* instância de *Rule*) que é derivada de uma hierarquia de políticas (instâncias da classe *Policy*).

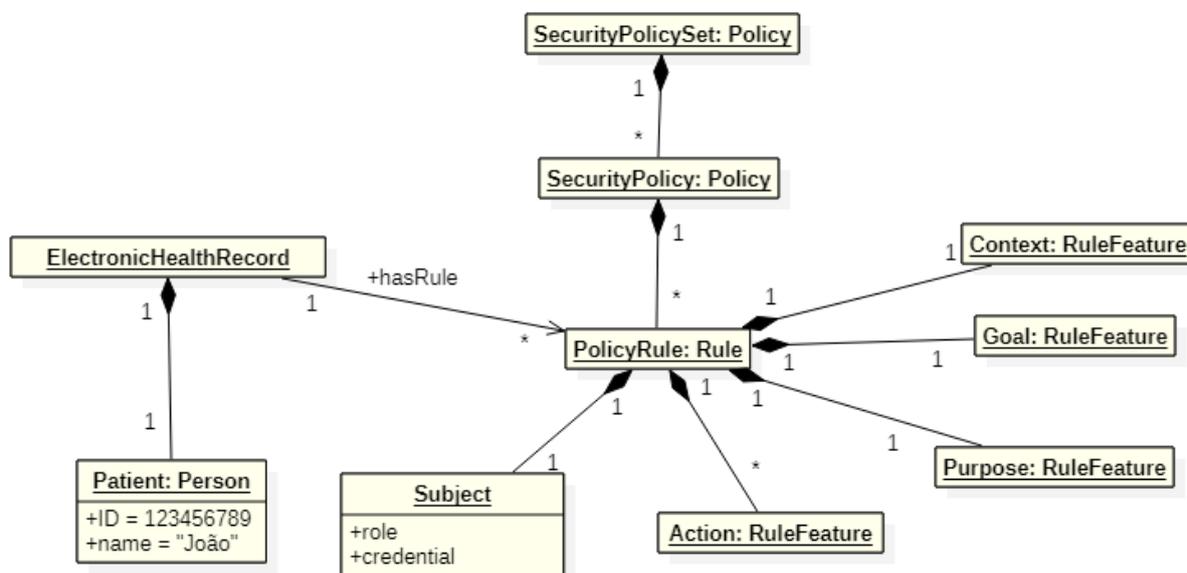


Figura 32: Exemplo de modelo de políticas de acesso a um documento clínico do PEP

A classe *PolicyRule* representa as políticas de segurança que regulam o acesso de diferentes usuários para os dados dos documentos clínicos do PEP de uma instituição de saúde. Estes usuários são modelados pela classe *Subject* e deverão cumprir um papel (*role*) e possuir credenciais (*credential*). Para simplificar, uma regra de política de segurança pode ter várias ações (capturados pela classe *Action* instância de *RuleFeature*) como por exemplo ler, escrever no PEP. Uma regra também deverá ter um propósito (*Purpose*), uma meta (*Goal*) e um contexto (*Context*) e a solicitação de cada acesso de um usuário será avaliada pelo cumprimento de

condições. A Tabela 12 mostra exemplos de regras de políticas de acesso de um PEP, estas regras foram criadas com base nos trabalhos de Ardagna et al. (2010), Jin et al. (2011) e Li et al. (2018).

Tabela 12. Exemplo de regras de políticas de acesso a um PEP. Adaptado de Li et al. (2015)

ID	Regra	Descrição
R1	Subject.role = “Nurse” and Subject.credential = * and Action = {read, writing_report_form} and Purpose = “treatment” and Context = emergency and goal = “authorization” → Permit	Enfermeiras podem ler o PEP de um paciente que não está sob responsabilidade dela para tratamentos em caso de emergência.
R2	Subject.role = “Doctor” and Subject.credential = * and (Action = {read, writing_report_form} or Action = {write, writing_treatment_log}) and Purpose = “treatment” and Context = emergency and goal = “authorization” → Permit	Médicos podem ler/escrever no PEP de um paciente que não está sob responsabilidade dele para tratamentos em caso de emergência.
R3	Subject.role = “Police” and Subject.credential = * and Action = {read, null} and Purpose = “abuseHelping” and goal = “authorization” → Permit	Policial pode ler os dados do PEP de pacientes, mas precisa notificar ao hospital

O modelo instanciado na Figura 32 responde à Q5, visto que a relação *hasRule* atribui um conjunto de regras de acesso ao PEP. Esta mesma relação permite associar políticas de acesso aos recursos abstraídos no metamodelo proposto. Exemplos de recursos são objetos preservados (*ContentInformation*) e qualquer documento clínico (*Composition*). Cabe mencionar que as políticas de acesso descrevem como as informações serão acessadas e recuperadas.

6.6 CENÁRIO 3

Maria está grávida de aproximadamente trinta semanas e realiza um exame de ultrassom obstétrico de rotina em uma clínica particular que fica próximo ao bairro dela e que possui várias unidades no país. Esta gravidez é muito especial para Maria, pois já passou por três abortos espontâneos recorrentes no passado. Em consequência, ela está tomando todos os cuidados recomendados.

Embora Maria haja passado as doze primeiras semanas consideradas as mais críticas, ainda se encontra temerosa, pois já ficou traumatizada devido às experiências passadas. Ela também teme se no futuro a sua filha passará por algo similar devido a uma predisposição genética, por esse motivo ela guarda criteriosamente uma cópia de todos os exames médicos que realiza e sempre solicita informações sobre as políticas de armazenamento de seus exames nas instituições de saúde que visita.

O cenário acima descreve um caso específico de armazenamento e preservação de dados de imagens médicas e visa responder às questões de competência Q3 e Q4. Existem diversas superposições de funções entre o SIS, RIS (*Radiology Information System*) e o PACS (*Picture Archiving and Communication System*) que permitem o gerenciamento de informações em uma entidade de saúde. Estas funções envolvem: a transmissão de imagens, banco de dados online, sistemas que permitem que imagens sejam compartilhados entre profissionais da saúde e que seja feita a visualização local ou remota (Clunie, 2000).

6.6.1 MODELO DO CENÁRIO 3

Consideremos um exemplo de laudo de ultrassonografia obstétrica apresentado na Figura 33 que corresponde ao cenário 3 descrito. Este laudo vem acompanhado de um conjunto de imagens médicas como resultado do exame. Conforme foi visto no cenário 1, o laudo de ultrassonografia pode ser facilmente representado com um modelo similar ao mostrado na Figura 28 e sua preservação seria representado com uma instância da classe *ContentInformation* associado a um conjunto de metadados.

Suponhamos que além do laudo, queremos preservar as imagens médicas. Este tipo de imagens é armazenado geralmente usando o padrão DICOM (Clunie, 2000) que além de guardar os mapas de bits que representam fisicamente as imagens do exame, também guarda as informações associadas a ele. No modelo de dados do DICOM as informações sobre um determinado paciente são organizadas em estudos, que de acordo ao metamodelo proposto corresponderiam a uma instância da classe *HealthEvent*. Um estudo (*Study* instância da classe *HealthEvent*) corresponde a uma internação, uma visita do paciente a uma instituição de saúde, ou a uma sequência de exames e documentos para diagnóstico, tratamento e documentação de uma condição de saúde. Cada estudo, está associado a uma instância de *Composition*, que por sua vez, contém uma ou mais series (representada pela classe *Serie*, instância de *Content*), que são sequências de exames DICOM da mesma modalidade. Em nosso cenário, uma série pode conter uma sequência de imagens de ultrassom referentes ao mesmo estudo. Na Figura 34 podemos ver um modelo de preservação de DICOM instanciado a partir do metamodelo proposto.

Paciente: Maria da Silva	
Data de Nascimento: 11/08/1983	
Data do Exame: 20/10/2019	
ultra-sonografia Obstétrica:	
Feto único em situação longitudinal apresentação cefálica, e com dorso à direita.	
Batimentos cardíacos rítmicos.	
Movimentos fetais espontâneos.	
Biometria Fetal:	
Diâmetro Bi-Parietal	-DBP: 76 mm
Circunferência Cefálica	-CC: 28,4 cm
Circunferência Abdominal	-CA: 25,1 cm
Comprimento Femoral:	-CF: 5,4 cm
Líquido amniótico em Quantidade normal.	
Placenta anterior, grau de maturidade 0/I (0 a III) e espessura média de 2,5 cm.	
O estômago e a bexiga fetal foram bem visualizados.	
Os rins e a coluna vertebral fetal têm aspecto normal.	
Não há evidência de hidrocefalia.	
A placenta situa-se distante do orifício cervical interno.	
Conclusão:	
Gestação de aproximadamente 30 semanas.	
Peso fetal atual aproximado de 1400 g.	
Altura fetal atual aproximada de 36,5 cm.	
Sexo feminino.	
Médico requisitante: Pedro Oliveira	

Figura 33: Exemplo de laudo do exame de ultrassonografia obstétrica em forma de texto.
(Fonte: Bortoluzzi, 2003)

Imagens médicas têm pouca utilidade quando isoladas de seu contexto, mas são informações valiosas quando estão associadas a informações sobre o paciente, laudos, e observações sobre achados. Por esse motivo, o padrão DICOM SR é utilizado para preservação (como instância da classe *ContentInformation*) e ele estabelece como devem ser formados, armazenados e transferidos documentos estruturados que podem representar laudos ou qualquer tipo de observação clínica. Estes documentos contêm informações de contexto, tais como procedimentos que devem ser executados para o sucesso de um tratamento e dados sobre profissionais de saúde envolvidos. Um objeto no padrão DICOM SR pode conter referências embutidas a imagens, arquivos de formato de onda, como eletrocardiogramas e hemogramas, e arquivos de áudio bem como outros documentos no mesmo padrão. O padrão estabelece métodos de referência a coordenadas destas imagens e formatos de onda, de forma integrada ao laudo. Cada documento codifica apenas informações semânticas, além disso, objetos no padrão DICOM fazem uso de terminologia extraída de dicionários de termos médicos, o que evita as

ambiguidades da linguagem natural, facilita o entendimento automatizado do conteúdo, a busca por informações específicas, a comparação de documentos e a tradução do conteúdo. Desta forma a questão de competência Q3 que visa a preservação de documentos clínicos é respondida.

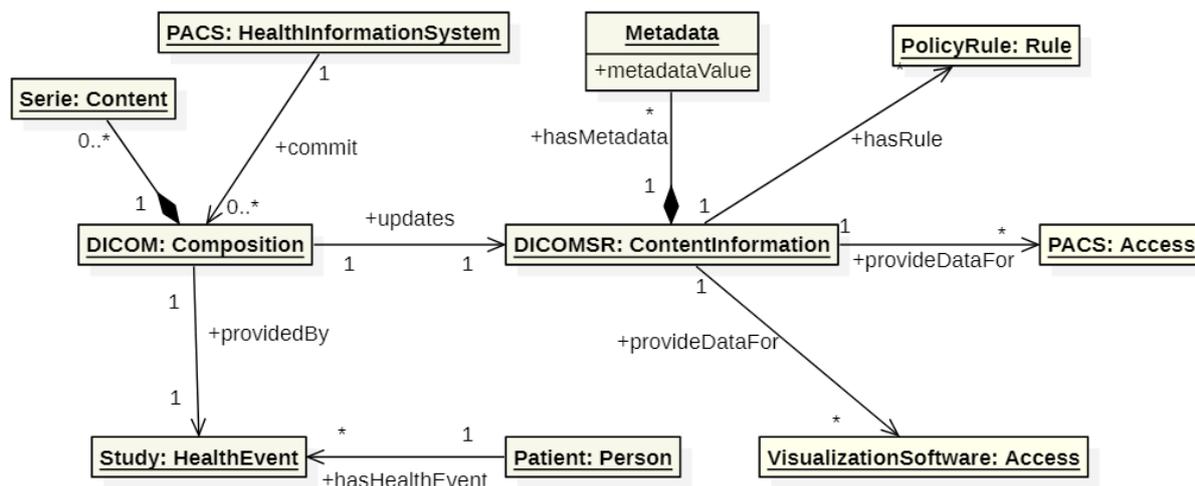


Figura 34: Modelo de preservação de DICOM.

Por fim, atendendo a questão de competência Q4, a recuperação das imagens preservadas pode ser feita por softwares especializados como é o caso do PACS ou qualquer SIS ou aplicativo de visualização de imagens DICOM (instâncias da classe *Access* de acordo com a Figura 34).

6.7 CONSIDERAÇÕES FINAIS

Neste capítulo foram elaboradas questões de competência que permitiram validar o metamodelo quanto a seus objetivos propostos na tese. Para responder estas questões foram criados três cenários a partir da identificação de características de preservação de dados clínicos tratados no Capítulo 4. Estes cenários descrevem em diferentes graus de detalhamento os eventos de saúde de uma pessoa em etapas diferentes da vida (cenário de: Sofia, João e Maria) e casos específicos de geração de documentos clínicos.

Foi verificado que a formulação de cenários, contribuiu a aperfeiçoar o próprio metamodelo. Assim a criação de instâncias a partir dos cenários descritos neste capítulo foi bem sucedido, pois ajudou a consolidar os conceitos necessários para um metamodelo com foco na preservação de dados clínicos. Porém ainda deve ser submetido a outras avaliações que permitam a participação de equipes multidisciplinares de modo a melhorar e expandir sua aplicabilidade.

Desta forma, de acordo à validação apresentada neste capítulo podemos concluir que o metamodelo responde às questões aos quais ele foi submetido, e, portanto, apresenta competência frente a seus objetivos.

CAPÍTULO 7 – CONCLUSÕES E TRABALHOS FUTUROS

7.1 CONCLUSÕES

O objetivo da tese consistiu no desenvolvimento de um metamodelo em um alto nível de abstração que permita representar o conhecimento associado ao armazenamento e preservação de dados clínicos associados ao ciclo de vida de uma pessoa. Para atingir este objetivo principal, foi realizado um estudo sistemático da literatura visando encontrar os elementos conceituais necessários que justifiquem cada decisão tomada durante a elaboração do metamodelo. Com esta identificação, na sequência se definiu um conjunto de requisitos e características que permitiram integrar em um alto nível de abstração os processos clínicos, prontuário eletrônico do paciente e preservação de dados.

Na revisão da literatura, constatou-se que os dados associados ao nascimento e óbito de uma pessoa também devem fazer parte do PEP. Embora a realidade se mostre diferente como o caso no Brasil onde existe sistemas independentes para esses propósitos como o SINAC²³ ou o sistema de informações de mortalidade (SIM²⁴). De acordo com as recomendações da OMS, toda informação clínica deve ser mantida no PEP e ser acessível desde que a entidade interessada tenha as credenciais para isso (Li et al. 2015; Evans, 2016; Fatima and Colomo-Palacios, 2018). A realidade mostra que a unificação do PEP é uma tarefa difícil apesar da existência de padrões de interoperabilidade, pois os dados clínicos encontram-se espalhados em diferentes SIS. Desta forma, a proposta do metamodelo é válido para este contexto.

O metamodelo proposto está baseado e fundamentado em modelos já existentes. O nascimento de uma pessoa, uma ocorrência médica e até o óbito são caracterizados dentro dos processos clínicos como sendo um evento "*EventHealth*" e em consequência um evento pode levar a outros sub eventos. Este entendimento permitiu adaptar modelos sobre processo clínicos (Behnam and Badreddin, 2013). A representação de prontuários eletrônicos foi baseada no padrão ISO 13606. Escolhemos esta norma porque considera que o PEP tem que ser um registro persistente da história clínica de uma pessoa a longo de sua vida e porque é potencialmente distribuída por permitir a recuperação da informação de um indivíduo independente do lugar de atendimento (Santos, 2010; Muñoz et al., 2011). Para a preservação de dados se fez uso da lógica do modelo OAIS. Este modelo visa orientar sistemas dedicados à preservação e

²³ http://bvsmis.saude.gov.br/bvs/publicacoes/funasa/declaracao_nasc_vivo.pdf (Acessado em:30/11/2019)

²⁴ <http://sim.saude.gov.br/default.asp> (Acessado em:30/11/2019)

manutenção do acesso a informações por longo prazo. Além disso foram considerados metamodelos sobre ecossistemas de dados (Oliveira, 2019) e modelos de políticas de segurança de dados clínicos (Li et al., 2015; Fatima and Colomo-Palacios, 2018; Chenthara et al., 2019).

Para validar o metamodelo foram elaboradas questões de competência formuladas em base aos objetivos da pesquisa. As questões de competência foram executadas frente a situações extraídas de três cenários fictícios. O objetivo foi verificar a conformidade da estrutura do metamodelo proposto por médio da instanciação de modelos que representem os cenários de teste. Dessa forma o metamodelo conseguiu responder satisfatoriamente as questões de competência nos cenários propostos. Porém é necessário considerar mais cenários para reforçar sua validade.

A avaliação do metamodelo por parte de especialistas apresentou-se como principal ameaça à validade do metamodelo. Isto é devido ao rigor mais estrito que as leis atuais oferecem para aceder a profissionais da área de saúde e arquivologia, pois implica elaborar projetos para serem submetidos ao comité de ética das respectivas instituições. Portanto, para que a avaliação seja consistente, implica ter uma amostra significativa de usuários que possam utilizar o metamodelo como referência para criar modelos baseado nas atividades particulares de cada especialista.

Por fim, devemos salientar que a preservação do PEP pode não ter implicações na saúde dos pacientes aos quais os documentos se referem, mas as informações que eles contém são recursos valiosos para a pesquisa e tomada de decisões, pois esses documentos possuem informações não apenas relativas à assistência em saúde, mas também informações importantes para a gestão em saúde, para a elaboração de políticas públicas de saúde, para apoiar novas pesquisas, considerando a questão de manter a privacidade dos dados do paciente no que se refere aos aspectos éticos e legais.

7.2 TRABALHOS FUTUROS

O trabalho desenvolvido nesta tese possibilitou identificar trabalhos que podem ser realizados visando a continuidade da pesquisa. A adoção de novas estratégias é necessária. As principais perspectivas em termos de trabalhos futuros são apresentadas em quatro direções: ampliação dos objetivos, aperfeiçoamento da proposta, aprimoramento das avaliações e aplicação da proposta.

Em relação à ampliação de objetivos, embora o metamodelo considere aspectos de privacidade como o controle de políticas de acesso. É necessário explorar outros aspectos de segurança como por exemplo a auditoria.

Em relação ao aperfeiçoamento da tese, há trabalhos futuros que podem ser realizados a fim de aprimorar nosso Metamodelo. Uma possibilidade está na formalização de regras e políticas de privacidade e segurança mais rigorosas. Isto devido a um crescimento constante do uso da nuvem como suporte de armazenamento de preservação de dados clínicos.

Em relação às avaliações realizadas nesta tese, trabalhos adicionais poderiam ser realizados onde as questões de competência desenvolvidas fossem aplicadas em outros cenários. Isso reforçaria a relevância dos resultados obtidos. Outra possibilidade refere-se à participação dos profissionais que atuam na área de arquivologia.

Outro trabalho futuro desta tese consiste em utilizar ferramentas como o GMF²⁵ (*Graphical Modeling Framework*), que também é baseado no EMF, para importar o metamodelo para gerar ferramentas de modelagem com interface gráfica. Isso facilitaria o processo de criação de instâncias para a recuperação e armazenamento de dados clínicos.

Um outro desafio, no âmbito de sistemas de saúde, é utilizar informações de contexto do ambiente. Deve-se considerar a seleção contextualizada em que as informações são inseridas, consultadas ou modificadas em determinada situação do paciente, sendo que ditas informações podem ser preservadas e armazenadas, também, junto à informação clínica do paciente.

²⁵ GMF: <https://www.eclipse.org/modeling/gmp/> Acesso: 31/11/2019

REFERÊNCIAS

Abbott, D. Digital curation and doctoral research. *International Journal of Digital Curation*, v. 10, n. 1, pp.1-17, 2014.

Abouelmehdi, K., Beni-Hessane, A. and Khaloufi, H. Big healthcare data: preserving security and privacy. *Journal of Big Data*, v. 5, n.1, p.1, 2018.

Abrams, S., Cruse, P. and Kunze, J. Preservation is not a place. *International Journal of Digital Curation*, v.4, n. 1, 2009.

Allard, S. DataONE: Facilitating eScience through collaboration. *Journal of eScience Librarianship*, v.1, n. 1, p.3, 2012.

American Library Association. Definitions of digital preservation. In *ALA Annual Conference*, Washington, DC. 2007.
<http://www.ala.org/alcts/resources/preserv/defdigpres0408> Acessado em: 30/11/2019.

Araujo, T. V., Pires, S. R. and Bandiera-Paiva, P. Adoção de padrões para Registro Eletrônico em Saúde no Brasil. *Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, v. 8, n. 4, 2014.

Ardagna, C. A., Di Vimercati, S. D. C., Foresti, S., Grandison, T. W., Jajodia, S. and Samarati, P. Access control for smarter healthcare using policy spaces. *Computers & Security*, v. 29, n. 8, pp. 848-858, 2010.

Askhoj, J., Sugimoto, S. and Nagamori, M. Preserving records in the cloud. *Records management journal*, v.21, n. 03, pp. 175-187, 2011,

Barateiro, J., Antunes, G., Freitas, F. and Borbinha, J. Designing digital preservation solutions: A risk management-based approach. *Int. Digit Curation*, v. 5, n. 1, pp. 4-17, 2010.

Becker, C., Antunes, G., Barateiro, J., Vieira, R. and Borbinha, J. Control objectives for dp: Digital preservation as an integrated part of it governance. *Proceedings of the American Society for Information Science and Technology*, 48(1), pp.1-10. 2011.

Behnam, S. A. and Badreddin, O. Toward a Care Process Metamodel: For business intelligence healthcare monitoring solutions. *5th International Workshop on Software Engineering in Health Care (SEHC)*. 2013.

Bézivin, J., Jouault, F. and Touzet, D. Principles, standards and tools for model engineering. In *Engineering of Complex Computer Systems*, 2005. Proceedings 10th IEEE International Conference on. pp. 28–29, 2005.

Bheemanathini, S. N. A Systematic Review of Blockchain Technology: Privacy Concerns, Security Challenges, and Solutions (Doctoral dissertation, University of Cincinnati). 2019.

Bill, R., Mazak, A., Wimmer, M., Vogel-Heuser, B.: On the need for temporal model repositories. In: *Proceedings of STAF Workshops*, pp. 136–145, 2018

Bote, J., Fernandez-Feijoo, B. and Ruiz, S. The cost of digital preservation: A methodological analysis. *Procedia Technology*, v5, pp.103-111, 2012.

Bote, J., Termens, M. and Gelabert, G. Evaluation of healthcare institutions for long-term preservation of electronic health records. In *International Conference on ENTERprise Information Systems*. Springer, Berlin, Heidelberg, pp. 136-145, 2011.

Bortoluzzi, M. K. Desenvolvimento e Implementação de um editor de documentos estruturados no padrão Dicom Structured Report. *Dissertação de Mestrado, UFSC*, 2003.

Bradley, K., Lei, J. and Blackall, C. Towards an open source repository and preservation system. Paris: UNESCO. v. 30, 2007.

Bunawan, A. A. and Nordin, S. The Challenges in Preserving The Electronic Records Metadata. *International journal of information systems and engineering*, v. 1, n. 1, 2015.

Burns, W. A. Healthcare's Data Tsunami: Why Healthcare Organizations are Drowning in Data and Why They Need Even More If They are Going to Survive the Storm (2011), pp. 56–59, 2011.

Carvalho, S. T. Modelagem de linha de produto de software dinâmica para aplicações ubíquas. Doctoral dissertation, Ph. D. Universidade Federal Fluminense, Niterói, RJ, Brasil, 2013.

CCSDS. Consultative Committee for Space Data Systems. J. Reference model for an open archival information system (OAIS). CCSDS 650.0-B-1, Blue Book. 2012.

Celesti, A., Fazio, M., Romano, A. and Villari, M. A hospital cloud-based archival information system for the efficient management of HL7 big data. In 2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). pp. 406-411. 2016.

Celesti, A., Fazio, M., Romano, A., Bramanti, A., Bramanti, P. and Villari, M. An OAIS-based hospital information system on the cloud: Analysis of a NoSQL column-oriented approach. IEEE journal of biomedical and health informatics, v.22, n.3, pp. 912-918, 2017.

Chaczko, Z., Gordon, L. C. and Bożejko, W. The Metamodel of Heritage Preservation for Medical Big Data. In International Conference on Computer Aided Systems Theory. Springer. pp. 366-371). 2018.

Chenthara, S., Ahmed, K., Wang, H. and Whittaker, F. Security and privacy-preserving challenges of e-Health solutions in cloud computing. IEEE, v. 7, pp.74361 - 74382, 2019.

Clunie, D. A. DICOM structured reporting. PixelMed publishing. 2000.

Cohen, D. J., Keller, S. R., Hayes, G. R., Dorr, D. A., Ash, J. S. and Sittig, D. F. Developing a model for understanding patient collection of observations of daily living: a qualitative meta-synthesis of the Project HealthDesign program. Personal and ubiquitous computing, v.19, n. 1, pp. 91-102, 2015.

CONARQ. Conselho Nacional de Arquivos. Disponível em: <http://conarq.arquivonacional.gov.br/coletanea-da-legislacao-arquivistica-e-correlata.html>. Acessado em: 30/11/2019.

CFM. Conselho Federal de Medicina. Resolução no 1638, de 10 de julho de 2002. Diário Oficial da União, Brasília, DF, n. 22, s. 1, p. 103, 2002. Disponível em: http://www.portalmedico.org.br/resolucoes/CFM/2002/1638_2002.htm. Acessado em: 30/11/2019.

Cordeiro, L. S., Parga, M. D. F. A. S., dos Santos Barbosa, N. and Menezes, S. C. F. Preservação Digital e a Biblioteconomia. Revista Bibliomar, pp.36-50. 2018.

Corn, M. Archiving the phenome: clinical records deserve long-term preservation. Journal of the American Medical Informatics Association, v. 16, n. 1, pp.1-6, 2009.

Cristani, M. and Cuel, R. A Survey on Ontology Creation Methodologies. *International Journal on Semantic Web and Information Systems*, v. 1, n. 2, pp. 49–69, 2005.

Cruz, J. A. S., Flores, D. and Garcia, O. M. C. O Prontuário Eletrônico de Paciente (PEP) como memória, patrimônio documental e cultural. *Em Questão*, v.17, n. 2, pp. 91-99. 2011.

Cruz, M J. R. and Díez-Carrera, C. Open Archival Information System (OAIS): lights and shadows of a reference model. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, ScienceDirect, v. 30, n. 70, pp. 221-247. 2016.

Cunha, J. D. A. and Lima, M. G.. Preservação digital: o estado da arte. *Questões em Rede*. <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/132>. 2012. Acessado em: 30/11/2019.

DDI. Data Documentation Initiative Specification. Data Documentation Initiative, 2012. <<http://www.ddialliance.org/Specification>>. Acessado em: 30/11/2019.

Dos Santos, V. B. Preservação de documentos arquivísticos digitais. *Ciência da Informação*, Brasília, DF, v. 41, n. 1, 2012.

Doyle, J., Viktor, H. and Paquet, E. Preservation metadata-a framework for 3D data based on the Semantic Web. In *2008 Third International Conference on Digital Information Management*, IEEE, pp. 925-927, 2008.

Dryden, J. The Open Archival Information System Reference Model. *Journal of Archival Organization*, v.7, n 4, pp. 214-217. 2009

Durak, U., Topcu, O., Siegfried, R. and Oguztuzun, H. Scenario development: A model-driven engineering perspective. In: *Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH)*, 2014 International Conference on, pp. 117–124, 2014.

Duranti, L. The long-term preservation of accurate and authentic digital data: the INTERPARES project. *Data Science Journal*, v. 4, pp. 106-118, 2005.

Earl, A. A reference model for computer assisted software engineering environment frameworks. Hewlett-Packard Laboratories, Fort Collins, CO. USA. 1990.

Evans, R. S. Electronic health records: then, now, and in the future. *Yearbook of medical informatics*, v. 25, n. 01, pp. 48-61. 2016.

Eizirik, C. L. and Bassols, A. M. S. O Ciclo da Vida Humana-: Uma Perspectiva Psicodinâmica. *Artmed*. 2013.

Fatoric, S. and Seekamp, E. Securing the future of cultural heritage by identifying barriers to and strategizing solutions for preservation under changing climate conditions. *Special Issue Cultural Heritage Conservation and Sustainability*, v. 9, n. 11, p.2143, 2017.

Fatima, A. and Colomo-Palacios, R. Security aspects in healthcare information systems: A systematic mapping. *Procedia computer science*, v. 138, pp. 12-19, 2018

Favre, J. M. Towards a basic theory to model driven engineering. In 3rd workshop in software model engineering, WISME. pp. 262-271, 2004.

Favre, J. M. and NGuyen, T. Towards a megamodel to model software evolution through transformations. *Electronic Notes in Theoretical Computer Science*, v. 127, n. 3, pp. 59-74. 2005.

Ferreira, M. Introdução à preservação digital: conceitos, estratégias e atuais consensos. Editora Escola de Engenharia Universidade do Minho, Portugal. 2006.

Fettke, P. and Loos, P. Classification of reference models: a methodology and its application. *Information systems and e-business management*, v. 1, n. 1, pp. 35-53, 2003

Fung, B. C., Wang, K., Chen, R. and Yu, P. S. Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys (CSUR)*, v. 42, n. 4, pp. 1-53, 2010.

Galvão, M. C. B. and Ricarte, I. L.M, *Prontuário do paciente*. Editora Guanabara Koogan. 2012.

Gartner, R. and Lavoie, B. *Preservation metadata*. Digital Preservation Coalition. 2013.

Giaretta, D. Introduction to OAIS Concepts and Terminology. In *Advanced Digital Preservation*. Springer, Berlin, Heidelberg. pp. 13-30, 2011.

Gomes, J. R. T. A metainformação na perspetiva do acesso e da preservação: análise da situação atual no CHSJ. Porto, Portugal. Dissertação de mestrado. 2013.

Gómez, A., Cabot, J. and Wimmer, M. TemporalEMF: A Temporal Metamodeling Framework. *Lecture Notes in Computer Science (LNCS)*, v. 11157, pp. 365–381, 2018.

Gonçalves, F., Maçorano, J. P., Gomes, J., David, G. and Cruz Correia, R. Plano de preservação digital como instrumento de gestão de informação clínica. In *Desafios y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas del VII Encuentro Ibérico EDICIC*. 2015.

Goossen, W. T. Detailed clinical models: representing knowledge, data and semantics in healthcare information technology. *Healthcare informatics research*, v. 20, n. 3, pp.163-172, 2014.

Gordon, L. C. C. Digital Data Heritage Preservation (DHP) Modelling and Design (Doctoral dissertation, University of Technology Sydney). 2018.

Graña, M. and Jackowski, K. Electronic health record: A review. In *2015 IEEE international conference on bioinformatics and biomedicine (BIBM)*. pp. 1375-1382, 2015.

Gruninger, M. and Fox, M. S. Methodology for the design and evaluation of ontologies. In *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95*, Montreal, CA, 1995.

Guedes, G. T. UML 2-Uma abordagem prática. Novatec Editora. 2018.

Guenther, R. Metadata to support long-term preservation of digital assets: PREMIS and its use with METS. In *Proceedings of the 2010 Roadmap for Digital Preservation Interoperability Framework Workshop*. pp. 1-6, 2010.

Hevner, A. and Chatterjee, S. Design science research in information systems. In *Design research in information systems*. Springer, Boston, MA. pp. 9-22, 2010.

Hevner, A., March, S. T., Park, J. and Ram, S. Design science research in information systems. *MIS quarterly*, v. 28, n. 1, pp. 75-105, 2004.

HL7. Health Level Seven. Disponível em:
<http://www.hl7.org/implement/standards/index.cfm?ref=nav> . Acessado em: 30/11/2019.

Higgins, S. The DCC curation lifecycle model. *International journal of digital curation*, v. 3, n.1, 2008.

Horiil, S. C., Prior, F. W., Bidgood, W. D., Parisot, C. and Claeys, G. DICOM: an introduction to the standard. In *The meeting material*. 1994.

ISO, International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 24744. *Software Engineering—Metamodel for Development Methodologies*. ISO, Geneva, v.12, 2007.

ISO 14721, *International Standard: Space Data and Information Transfer Systems, Open Archival Information System (OAIS), Reference Model*. ISO, 2012.

ISO 13940, *Health Informatics - System of Concepts to Support Continuity of Care*. ISO, 2016.

Jalali, S. and Wohlin, C. Systematic literature studies: database searches vs. backward snowballing. In *Proceedings of the 2012 ACM-IEEE international symposium on empirical software engineering and measurement*, pp. 29-38, 2012.

Jin, J., Ahn, G. J., Hu, H., Covington, M. J. and Zhang, X. Patient-centric authorization framework for electronic healthcare services. *computers & security*, v. 30, n 2-3, pp. 116-127, 2011.

Joemon, J., Rauber, A., Sebastiani, R. and Frommholz, I. *Research and Advanced Technology for Digital Libraries*. 14th European Conference, ECDL 2010, Glasgow, UK. Springer-Verlag Berlin Heidelberg, 2010.

Kalra, D. Electronic health record standards. *Yearbook of medical informatics*, v. 15, n. 01, pp. 136-144, 2006.

Karadsheh, L. Applying security policies and service level agreement to IaaS service model to enhance security and transition. *computers & security*, v.31, n. 3, pp. 315-326, 2012.

Kiefer, S., Schäfer, M. and Rauch, J. A semantic approach for digital long-term preservation of electronic health documents. In *MIE*. pp. 265-269. 2012.

Kitchenham, B. A. Systematic review in software engineering: where we are and where we should be going. In Proceedings of the 2nd international workshop on Evidential assessment of software technologies. pp. 1-2, 2012.

Kitchenham, B. and Charters, S. Guidelines for performing systematic literature reviews in software engineering version 2.3. EBSE Technical Report EBSE-2007-01, Keele University and University of Durham, v. 45, n. 4, p.1051, 2007.

Korb, J. and Strodl, S. Digital preservation for enterprise content: A Gap-analysis between ECM and OAIS. iPRES 2010 - Proceedings of the 7th International Conference on Preservation of Digital Objects. pp. 221-228, 2010.

Kuroda, T., Kimura, E., Matsumura, Y., Yamashita, Y., Hiramatsu, H. and Kume, N. Simulating cloud environment for HIS backup using secret sharing. Studies in health technology and informatics, v. 192, pp.171-174. 2013.

Laerum H. and Faxvaag, A. Task-oriented evaluation of electronic medical records systems: development and validation of a questionnaire for physicians. BMC medical informatics and decision making, v. 4, n. 1, p.1, 2004.

Lavoie, B. F. The open archival information system (OAIS) reference model: introductory guide. Digital Preservation Coalition. 2014.

Lee, K. H., Slattery, O., Lu, R., Tang, X. and McCrary, V. The state of the art and practice in digital preservation. Journal of research of the National institute of standards and technology, v. 107, n. 1, p. 93, 2002.

Li, S., Zhang, T., Gao, J. and Park, Y. A sticky policy framework for big data security. In 2015 IEEE First International Conference on Big Data Computing Service and Applications. IEEE. pp. 130-137, 2015.

Li, H., Zhu, L., Shen, M., Gao, F., Tao, X. and Liu, S. Blockchain-based data preservation system for medical data. Journal of medical systems, v. 42, n. 8, p.141, 2018.

Limo, J., Kwanya, T. and Chepchirchir, S. Electronic records management strategies by government agencies in Kenya. Record and Library Journal, v. 5, n.1, pp. 62-71, 2019.

Lo, N. W., Wu, C. Y. and Chuang, Y. H. An authentication and authorization mechanism for long-term electronic health records management. *Procedia Computer Science*, v. 111, pp. 145-153, 2017.

López-Fernández, J. J., Cuadrado, J. S., Guerra, E. and De Lara, J. Example-driven meta-model development. *Software & Systems Modeling*, v. 14, n. 4, pp.1323–1347, 2015.

Lozano, R. A Metamodel for clinical data integration: basis for a new EHR model driven by ontologies. Tese (Doutorado). Universidade Autônoma de Barcelona. 2016.

Malhotra, R. and Chug, A. Software maintainability: Systematic literature review and current trends. *International Journal of Software Engineering and Knowledge Engineering*, v. 26, n. 08, pp.1221-1253, 2016.

Malone, J., Brown, A., Lister, A. L., Ison, J., Hull, D., Parkinson, H. and Stevens, R. The Software Ontology (SWO): a resource for reproducibility in biomedical data analysis, curation and digital preservation. *Journal of biomedical semantics*, v. 5, n.1, p.25. 2014.

Márdero A. M. Á. Critérios para a preservação digital da informação científica. Tese (Doutorado), Universidade de Brasília (Brasil). 2008.

Merks, E. and James S. Essential EMF. *DZone Refcardz*, n.39, 2009.

Mihailescu, M., Mihailescu, D. and Carlsson, S. Understanding Healthcare Digitalization: A Critical Realist Approach. In *Proceedings ICIS2017 conference on Digital Libraries*. 2017.

Mildenberger, P., Eichelberg, M. and Martin, E. Introduction to the DICOM standard. *European radiology*, v. 12, n. 4, pp. 920-927, 2002.

Misra, D., Hall, R. H., Payne, S. M. and Thomas, G. R. Digital preservation and knowledge discovery based on documents from an international health science program. In *Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries*. pp. 23-26. 2012.

Muñoz, P., Trigo, J. D., Martínez, I., Muñoz, A., Escayola, J. and García, J. The ISO/EN 13606 standard for the interoperable exchange of electronic health records. *Journal of Healthcare Engineering*, v. 2, n. 1, pp. 1-24, 2011.

Oliveira, M. I. D. S. A metadata curation framework for data ecosystems. Tese (Doutorado), Universidade Federal Pernambuco, Pernambuco, Brasil, 2019.

Oliveira, M. I. D. S. and Loscio, B. F. Metadata Curation Framework for Supporting Data Ecosystems. In Thesis and Dissertation Workshop, Simpósio Brasileiro de Banco de Dados, Uberlândia-MG, 2017.

Oliveira, M. I. S., Oliveira, L. E. R., Batista, M. G. R. and Lóscio, B. F. Towards a meta-model for data ecosystems. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, n. 72, pp. 1-10, 2018

Ou, R., Chen, F., Ren, Y., Liu, Y. and Wang, Q. Digital Continuity Guarantee Method of Urban Construction Archives Based on Risk Management. Journal on Big Data JBD, v.1, n.2, pp. 71-77, 2019.

Park, Y. R., Kim, H., An, E. Y., Kim, H. H., Kim, J. H., Park, R. W., ... and Jung, E. Y.. Establishing semantic interoperability in the course of clinical document exchange using international standard for metadata registry. The Journal of the Korean Medical Association. v. 55, n. 8, pp.729-740, 2012.

Peffer, K., Tuunanen, T., Gengler, C. E., Rossi, M., Hui, W., Virtanen, V., Bragge, J. The design science research process: A model for producing and presenting information systems research. In In: 1st International Conference on Design Science in Information Systems and Technology. pp. 83-106, 2006.

Pennock, M., Digital Preservation: Continued Access to Authentic Digital Assets. Bristol: JISC. 2006.

Piconi, A. C. Ontologia de metadados para a preservação de prontuário eletrônico do paciente (PEP). Campinas, SP. Dissertação. 2014.

Piconi, A. C., Galvão, M. C. B. and Ricarte, I. L. M. Estratégias de preservação digital para prontuários eletrônicos de pacientes. Programa Científico e Resumos em VI EADCA, Universidade de Campinas. 2013.

Pinto, V. B. and Sales, O. M. M. Proposta de aplicabilidade da preservação digital ao prontuário eletrônico do paciente. RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação, Campinas, SP. v.15, n. 2, pp.489-507, 2017.

Rabbi, F., Lamo, Y. and MacCaull, W. A Flexible Metamodelling Approach for Healthcare Systems. In PAHI. pp. 115-128. 2014.

Riley, J. Understanding metadata. What is metadata and what is it for.? NISO How the information world connects, Washington DC, United States, v. 23, 2017.

Roehrs, A., Da Costa, C. A. and Da Rosa Righi, R. OmniPHR: A distributed architecture model to integrate personal health records. Journal of biomedical informatics, v.71, pp.70-81. 2017.

Ross, S. Digital preservation, archival science and methodological foundations for digital libraries. New Review of Information Networking, v. 17, n.1, pp. 43-68, 2012.

Rothenberg, J. Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation. A Report to the Council on Library and Information Resources. Council on Library and Information Resources, 1755 Massachusetts Ave., NW, Washington, DC 20036, 1999.

Rudersdorf, A. Digital preservation ingest can be a CINCH. Library hi tech, v. 30, n. 3, pp. 449-456, 2012.

Rusbridge, C., Burnhill, P., Ross, S., Buneman, P., Giaretta, D., Lyon, L. and Atkinson, M. The digital curation centre: a vision for digital curation. In 2005 IEEE International Symposium on Mass Storage Systems and Technology, pp. 31-41, 2005.

Santana, A. F. L., Alves, C. F. Bpmg—um modelo conceitual para governança em bpm—aplicação numa organização pública. iSys-Revista Brasileira de Sistemas de Informação, v. 9, n. 1, pp. 139–167, 2016.

Santos, M. R., Bax, M. P. and Kalra, D. Building a logical EHR architecture based on ISO 13606 standard and semantic web technologies. In MedInfo, pp. 161-165, 2010.

Sassi, R., Bond, R. R., Cairns, A., Finlay, D. D., Guldenring, D., Libretti, G., ... and Cuccia, C. PDF-ECG in clinical practice: A model for long-term preservation of digital 12-lead ECG data. *Journal of electrocardiology*, v. 50, n. 6, pp. 776-780, 2017.

Sayão, L. F. and Sales, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. *Informação & Sociedade: Estudos*, v. 22, n.3, 2013.

Schmidt, D. C. Model-driven engineering. *Computer-IEEE Computer Society*, v. 39, n. 2, p. 25, 2006.

Scott, R. E. E-Records in health—preserving our future. *International journal of medical informatics*, v. 76, n. 5-6, pp. 427-431, 2007.

Seidewitz, E. What models mean. *IEEE software*, v. 20, n. 5, pp.26-32. 2003.

Silva, C. I. Estruturação de Laudos de Tomografia Computadorizada do Crânio com Base em Ontologias. Tese de Doutorado, Instituto de Computação, Universidade Federal Fluminense. 2018.

Steinberg, D., Budinsky, F., Paternostro, M., Merks, E. EMF: Eclipse Modeling Framework, Second Edition. Addison-Wesley Professional, 2008.

Suderman, J. Principle-based concepts for the long-term preservation of digital records. In *Proceedings of the 1st International Digital Preservation Interoperability Framework Symposium*. pp. 1-8, 2010.

Takeda, H., Veerkamp, P., Yoshikawa, H. Modeling design process. *AI magazine* v. 11, n. 4, p. 37, 1990.

Thibodeau, K. Overview of technological approaches to digital preservation and challenges in coming years. *The state of digital preservation: an international perspective*, pp. 4-31, 2002.

Time, R. T. Researching long term digital preservation approaches in the dutch digital preservation testbed (testbed digitale bewaring). *RLG DigiNews*, v. 6, n. 3, 2002.

Vaishnavi, V., Kuechler, W. and Petter, S. Design science research in information systems. v. 20, 2004.

Van Gorp, P., Comuzzi, M., Jahnen, A., Kaymak, U. and Middleton, B. An open platform for personal health record apps with platform-level privacy protection. *Computers in biology and medicine*, v. 51, pp. 14-23, 2014.

Vieira, M., A. Modelagem de Espaços Inteligentes Pessoais e Espaços Inteligentes Fixos no Contexto de Cenários de Computação Ubíqua. Goiânia- GO, Brasil. Dissertação de Mestrado. Instituto de Informática, Universidade Federal de Goiás. p.153, 2016.

Völter, M., Stahl, T., Bettin, J., Haase, A. and Helsen, S. *Model-driven software development: technology, engineering, management*. John Wiley and Sons. 2013.

Voinov, N., Drobintsev, P., Kotlyarov, V. and Nikiforov, I. Distributed OAIS-based digital preservation system with HDFS technology. In *2017 20th Conference of Open Innovations Association (FRUCT)*. IEE, pp. 491-497, 2017.

Walters, T. and Skinner, K. *New roles for new times: Digital curation for preservation*. Association of Research Libraries. 2011.

Waters, D. and Garrett, J. *Preserving Digital Information*. Report of the Task Force on Archiving of Digital Information. The Commission on Preservation and Access, 1400 16th St., NW, Suite 740, Washington, DC 20036-2217. 1996.

Winter, T. *Material Culture as History: Science and the International Ordering of Heritage Preservation*. *A Companion to Public History*, pp.289-300. 2018.

Yina, W. Application of EHR in health care. In *2010 Second International Conference on Multimedia and Information Technology*, v.1, pp. 60-63, 2010.

Zhu, L., Gai, K. and Li, M. *Blockchain-Enabled Cloud Data Preservation Services*. In *Blockchain Technology in Internet of Things*, Springer Nature Switzerland, pp. 43-52, 2019.

APÊNDICE A – ARQUIVO XML ECORE DO METAMODELO

A seguir é disponibilizado o código fonte referente ao metamodelo.

```
<?xml version="1.0" encoding="UTF-8"?>
<ecore:EPackage xmi:version="2.0" xmlns:xmi="http://www.omg.org/XMI"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:ecore="http://www.eclipse.org/emf/2002/Ecore" name="preservationMetamodel"
nsURI="http://www.example.org/preservationMetamodel"
  nsPrefix="preservationMetamodel">
  <eClassifiers xsi:type="ecore:EClass" name="HealthEvent">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore###EString"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="date" eType="ecore:EDataType
http://www.eclipse.org/emf/2003/XMLType###Date"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasSubHealthEvent"
upperBound="-1"
      eType="###HealthEvent"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="responsibleFor"
lowerBound="1"
      upperBound="-1" eType="###HealthCareProvider"
eOpposite="###HealthCareProvider/involvedIn"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="isUsedIn" lowerBound="1"
      upperBound="-1" eType="###HealthCareResource"
eOpposite="###HealthCareResource/uses"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasFeature" upperBound="-1"
      eType="###Feature" containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="DeadEvent" eSuperTypes="###HealthEvent"/>
  <eClassifiers xsi:type="ecore:EClass" name="PreservationPlan">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore###EString"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="description"
eType="ecore:EDataType http://www.eclipse.org/emf/2002/Ecore###EString"/>
```

```

    <eStructuralFeatures xsi:type="ecore:EReference" name="performs" lowerBound="1"
      upperBound="-1" eType=":///Activity" containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Person" eSuperTypes=":///Actor">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="ID" eType="ecore:EDataType
      http://www.eclipse.org/emf/2002/Ecore:///EInt"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasHealthEvent"
      lowerBound="1"
      upperBound="-1" eType=":///HealthEvent"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasParent" upperBound="2"
      eType=":///Person"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
      http://www.eclipse.org/emf/2002/Ecore:///EString"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasDemographicData"
      upperBound="-1"
      eType=":///DemographicData" containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="childrens" upperBound="-1"
      eType=":///Person"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="BornEvent" eSuperTypes=":///HealthEvent"/>
  <eClassifiers xsi:type="ecore:EClass" name="Resource" abstract="true">
    <eStructuralFeatures xsi:type="ecore:EReference" name="conformsTo" upperBound="-1"
      eType=":///Standard"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasSemantic" lowerBound="1"
      upperBound="-1" eType=":///Semantic"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasStructure"
      eType=":///Structure"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasRule" upperBound="-1"
      eType=":///Rule"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Feature">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
      http://www.eclipse.org/emf/2002/Ecore:///EString"/>

```

```

    <eStructuralFeatures xsi:type="ecore:EAttribute" name="description"
eType="ecore:EDataType http://www.eclipse.org/emf/2002/Ecore#//EString"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="subFeature" upperBound="-1"
    eType="#//Feature" containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="derivedFeature" upperBound="-
1"
    eType="#//Feature"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasValue"
eType="#//FeatureValue"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Standard">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="description"
eType="ecore:EDataType http://www.eclipse.org/emf/2002/Ecore#//EString"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="type" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Activity">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="description"
eType="ecore:EDataType http://www.eclipse.org/emf/2002/Ecore#//EString"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="goal" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="subActivities" upperBound="-1"
    eType="#//Activity"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="ContentInformation"
eSuperTypes="#//Resource">
    <eAnnotations source="http://www.eclipse.org/emf/2002/GenModel">
    <details key="documentation" value="Informação a ser preservada&#xD;&#xA;É um
objeto de informação composto por seus dados de conteúdo&#xD;&#xA;AIP&#xD;&#xA;è
tipo um relatório em XML"/>

```

```

</eAnnotations>
  <eStructuralFeatures xsi:type="ecore:EReference" name="describedBy"
eType="#//PreservationPlan"/>
  <eStructuralFeatures xsi:type="ecore:EReference" name="hasMetadata" upperBound="-1"
  eType="#//Metadata" containment="true"/>
  <eStructuralFeatures xsi:type="ecore:EReference" name="provideDataFor"
lowerBound="1"
  upperBound="-1" eType="#//Access"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="HealthInformationSystem">
  <eStructuralFeatures xsi:type="ecore:EAttribute" name="type" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
  <eStructuralFeatures xsi:type="ecore:EReference" name="commit" upperBound="-1"
  eType="#//Composition"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="HealthCareProvider">
  <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
  <eStructuralFeatures xsi:type="ecore:EReference" name="involvedIn" lowerBound="1"
  upperBound="-1" eType="#//HealthEvent"
eOpposite="#//HealthEvent/responsibleFor"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Repository">
  <eStructuralFeatures xsi:type="ecore:EReference" name="hasPackage" lowerBound="1"
  upperBound="-1" eType="#//Package" containment="true"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Package">
  <eStructuralFeatures xsi:type="ecore:EReference" name="subPackage" upperBound="-1"
  eType="#//Package" containment="true"/>
  <eStructuralFeatures xsi:type="ecore:EReference" name="hasContentInformation"
  upperBound="-1" eType="#//ContentInformation" containment="true"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="ElectronicHealthRecord">

```

```

    <eStructuralFeatures xsi:type="ecore:EAttribute" name="recordID"
eType="ecore:EDataType http://www.eclipse.org/emf/2002/Ecore#//EInt"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasPatient" lowerBound="1"
    eType="#//Person" containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="folders" upperBound="-1"
    eType="#//Folder" containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="allCompositions"
upperBound="-1"
    eType="#//Composition" containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasRule" lowerBound="1"
    upperBound="-1" eType="#//Rule"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Semantic"/>
<eClassifiers xsi:type="ecore:EClass" name="Structure"/>
<eClassifiers xsi:type="ecore:EClass" name="Curator" eSuperTypes="#//Actor">
    <eStructuralFeatures xsi:type="ecore:EReference" name="actsOn" lowerBound="1"
    upperBound="-1" eType="#//PreservationPlan"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="DemographicData">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="type" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Metadata">
    <eStructuralFeatures xsi:type="ecore:EReference" name="has" upperBound="-1"
eType="#//MetadataValue"
    containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#//EString"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Record_Component" abstract="true"
eSuperTypes="#//Resource"/>

```

```

<eClassifiers xsi:type="ecore:EClass" name="Content"
eSuperTypes="#//Record_Component"/>
<eClassifiers xsi:type="ecore:EClass" name="Entry" eSuperTypes="#//Content">
  <eStructuralFeatures xsi:type="ecore:EReference" name="itens" upperBound="-1"
    eType="#//Item" containment="true"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Section" eSuperTypes="#//Content">
  <eStructuralFeatures xsi:type="ecore:EReference" name="members" upperBound="-1"
    eType="#//Content" containment="true"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Item"
eSuperTypes="#//Record_Component"/>
<eClassifiers xsi:type="ecore:EClass" name="Cluster" eSuperTypes="#//Item">
  <eStructuralFeatures xsi:type="ecore:EReference" name="parts" upperBound="-1"
    eType="#//Item" containment="true"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Element" eSuperTypes="#//Item">
  <eStructuralFeatures xsi:type="ecore:EReference" name="hasValue" eType="#//Value"
    containment="true"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Folder"
eSuperTypes="#//Record_Component">
  <eStructuralFeatures xsi:type="ecore:EReference" name="subFolders" upperBound="-1"
    eType="#//Folder" containment="true"/>
  <eStructuralFeatures xsi:type="ecore:EReference" name="compositions" upperBound="-1"
    eType="#//Composition" containment="true"/>
</eClassifiers>
<eClassifiers xsi:type="ecore:EClass" name="Composition"
eSuperTypes="#//Record_Component">
  <eStructuralFeatures xsi:type="ecore:EReference" name="contents" upperBound="-1"
    eType="#//Content" containment="true"/>
  <eStructuralFeatures xsi:type="ecore:EReference" name="providedBy" lowerBound="1"
    eType="#//HealthEvent"/>

```

```

    <eStructuralFeatures xsi:type="ecore:EReference" name="updates"
eType="//ContentInformation"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Family">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="name" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#/EString"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="members" upperBound="-1"
      eType="//Person" containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="HealthCareEntity">
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasHealthCareProvider"
      lowerBound="1" upperBound="-1" eType="//HealthCareProvider"
containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasHealthCareResource"
      lowerBound="1" upperBound="-1" eType="//HealthCareResource"
containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="has" lowerBound="1"
upperBound="-1"
      eType="//HealthInformationSystem" containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="HealthCareResource">
    <eStructuralFeatures xsi:type="ecore:EReference" name="uses" lowerBound="1"
upperBound="-1"
      eType="//HealthEvent" eOpposite="//HealthEvent/isUsedIn"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="FeatureValue"/>
  <eClassifiers xsi:type="ecore:EClass" name="MetadataValue">
    <eStructuralFeatures xsi:type="ecore:EAttribute" name="value" eType="ecore:EDataType
http://www.eclipse.org/emf/2002/Ecore#/EJavaObject"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Value"/>
  <eClassifiers xsi:type="ecore:EClass" name="Policy">
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasRuleClaim" lowerBound="1"
upperBound="-1" eType="//Rule" containment="true"/>

```

```

    <eStructuralFeatures xsi:type="ecore:EReference" name="subPolicy" upperBound="-1"
      eType="//Policy" containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Rule">
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasSubject" lowerBound="1"
      eType="//Subject" containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="haFeatureRule"
lowerBound="1"
      upperBound="-1" eType="//RuleFeature" containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Subject" eSuperTypes="//Actor">
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasRole" lowerBound="1"
      eType="//Role" containment="true"/>
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasCredential" lowerBound="1"
      eType="//Credential" containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Role"/>
  <eClassifiers xsi:type="ecore:EClass" name="Credential"/>
  <eClassifiers xsi:type="ecore:EClass" name="RuleFeature">
    <eStructuralFeatures xsi:type="ecore:EReference" name="hasValue"
eType="//RuleFeatureValue"
      containment="true"/>
  </eClassifiers>
  <eClassifiers xsi:type="ecore:EClass" name="Actor" abstract="true"/>
  <eClassifiers xsi:type="ecore:EClass" name="RuleFeatureValue"/>
  <eClassifiers xsi:type="ecore:EClass" name="Access"/>
</ecore:EPackage>

```

APÊNDICE B – ESPECIFICAÇÃO DO METAMODELO

Nas seguintes seções todas as classes, propriedades e relações do metamodelo proposto são definidas.

B.1 CLASSE: HEALTH EVENT

Classe	<i>HealthEvent</i>
Descrição	Um <i>HealthEvent</i> é um elemento no metamodelo que possui um nome e uma data. O <i>HealthEvent</i> refere-se aos eventos temporais de saúde de uma pessoa a longo de sua vida. Por exemplo, consulta médica, triagem por emergência, procedimento cirúrgico, exames de laboratório, entre outros.
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *HealthEvent*

B.1.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Nome atribuído a um evento de saúde, pode estar baseado em um padrão como por exemplo o ICPC.	<i>EString</i>
<i>date</i>	Data de ocorrência do evento de saúde.	<i>Date</i>

B.1.2 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasSubHealthEvent</i>	Permite que um evento de saúde possa ter eventos derivados.	<i>reference</i>	0..*
<i>hasFeature</i>	Um evento de saúde está composto por um conjunto de características que o descrevem.	<i>composition</i>	0..*

B.2 CLASSE: BORN EVENT

Classe	<i>BornEvent</i>
Descrição	Representa o nascimento de uma pessoa que gera documentos únicos como por exemplo. Declaração de nascido vivo e certidão de nascimento.
Sub classe de:	<i>HealthEvent</i>

B.3 CLASSE: DEAD EVENT

Classe	<i>DeadEvent</i>
Descrição	É o evento de falecimento de uma pessoa que gera documentos únicos como por exemplo o relatório de necropsia.
Sub classe de:	<i>HealthEvent</i>

B.4 CLASSE: FEATURE

Classe	<i>Feature</i>
Descrição	Representa um atributo ou aspecto que caracteriza um evento de saúde. A <i>feature</i> pode ou não afetar diretamente ao comportamento do <i>HealthEvent</i>
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *Feature*

B.4.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Refere-se ao vocabulário ou ontologia que explica o significado de uma característica	<i>EString</i>
<i>description</i>	Texto livre que descreve a feature	<i>EString</i>

B.4.2 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>derivedFeature</i>	Refere-se à transformação de uma entidade em outra, ou seja, uma	<i>reference</i>	0..*

	atualização de uma entidade resultando em uma nova ou a construção de uma nova entidade com base em uma entidade preexistente.		
<i>subFeature</i>	Refere-se a que pode existir instâncias <i>Features</i> compostas por outras instâncias de <i>subFeatures</i>	<i>composition</i>	0..*
<i>hasValue</i>	Refere-se ao agrupamento de instâncias do <i>hasValue</i> . Uma instância de <i>Feature</i> pode estar associada de zero ou até uma instância de <i>hasValue</i>	<i>reference</i>	0..1

B.5 CLASSE: FEATUREVALUE

Classe	<i>FeatureValue</i>
Descrição	É o valor de uma <i>Feature</i>
Sub classe de:	<i>EClass</i>

B.6 CLASSE: PERSON

Classe	<i>Person</i>
Descrição	Refere-se a uma especialização da classe <i>Actor</i> que representa a uma pessoa (paciente)
Sub classe de:	<i>Actor</i>

A seguir são descritos os atributos e relações da classe *Person*

B.6.1 Atributos

Atributo	Descrição	Tipo
<i>ID</i>	É a identificação inequívoca de uma instância da Classe <i>Person</i> (ex. documento identidade nacional)	<i>EInt</i>
<i>name</i>	Nome atribuído à pessoa	<i>EString</i>

B.6.2 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasParent</i>	Relação que indica parentesco (pai e mãe)	<i>reference</i>	0..2
<i>childrens</i>	Associa uma pessoa a seus filhos caso tenha.	<i>reference</i>	0..*
<i>hasDemographicData</i>	Uma pessoa está composta por um conjunto de dados demográficos.	<i>composition</i>	0..*
<i>hasHealthEvent</i>	Relaciona uma pessoa a vários eventos de saúde.	<i>reference</i>	1..*

B.7 CLASSE: DEMOGRAPHIC DATA

Classe	<i>DemographicData</i>
Descrição	Refere-se aos dados demográficos de uma pessoa como contato ou residência
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos da classe *DemographicData*

B.7.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	O nome do atributo de um dado demográfico	<i>EString</i>
<i>type</i>	Refere-se a uma categoria de dado demográfico	<i>EString</i>

B.8 CLASSE: FAMILY

Classe	<i>Family</i>
Descrição	Representação de uma instância da Classe <i>Person</i> .
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos da classe *Family*

B.8.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Atributo que descreve a família	<i>EString</i>

B.9 CLASSE: HEALTH CARE PROVIDER

Classe	<i>HealthCareProvider</i>
Descrição	Representa aos profissionais da área da saúde que atuam em um procedimento médico representado pela classe <i>HealthEvent</i>
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *HealthCareProvider*

B.9.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Nome da instância de <i>HealthCareProvider</i>	<i>EString</i>

B.9.2 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>involvedIn</i>	Um provedor de saúde pode estar envolvido em muitos eventos de saúde.	<i>reference</i>	1..*

B.10 CLASSE: HEALTH CARE RESOURCE

Classe	<i>HealthCareResource</i>
Descrição	Refere-se aos recursos de infraestrutura (espaços físicos, equipamentos) que são utilizados em um <i>healthEvent</i>
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *HealthCareResource*

B.10.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>uses</i>	Recursos que podem ser utilizados são equipamentos médicos ou infraestrutura.	<i>reference</i>	1..*

B.11 CLASSE: HEALTH CARE ENTITY

Classe	<i>HealthCareEntity</i>
Descrição	Refere-se a uma entidade de assistência médica ou centros de saúde (ex. Hospitais)
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *HealthCareEntity*

B.11.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasHealthCareResource</i>	Uma instância <i>HealthCareEntity</i> está composto por um conjunto de recursos de infraestrutura.	<i>composition</i>	1..*
<i>hasHealthCareProvider</i>	Uma instância <i>HealthCareEntity</i> está composto por um conjunto de provedores de saúde (profissionais).	<i>composition</i>	1..*
<i>has</i>	Uma instância <i>HealthCareEntity</i> está composto por vários sistemas de informação de saúde representado pela classe <i>HealthInformationSystem</i>	<i>composition</i>	1..*

B.12 CLASSE: HEALTH INFORMATION SYSTEM

Classe	<i>HealthInformationSystem</i>
Descrição	São os sistemas de informação de saúde implantados nas entidades de assistência médica
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *HealthCareInformationSystem*

B.12.1 Atributos

Atributo	Descrição	Tipo
<i>type</i>	Tipo do sistema de informação de saúde	<i>EString</i>

B.12.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>commit</i>	Um sistema de informação de saúde gerencia vários	<i>reference</i>	1..*

B.13 CLASSE: ELECTRONIC HEALTH RECORD

Classe	<i>ElectronicHealthRecord</i>
Descrição	Representa ao prontuário eletrônico do paciente que está associado a uma pessoa representado pela Classe <i>Person</i> .
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *ElectronicHealthRecord*

B.13.1 Atributos

Atributo	Descrição	Tipo
<i>recordID</i>	É o identificador ou número do prontuário	<i>EInt</i>

B.13.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>folders</i>	Um prontuário eletrônico está composto por vários folders (organizadores)	<i>composition</i>	0..*
<i>allCompositions</i>	Um prontuário eletrônico está composto por vários documentos (exames, laudos).	<i>composition</i>	0..*
<i>hasRule</i>	Um prontuário tem um conjunto de políticas de acesso definido por regras	<i>reference</i>	1..*
<i>hasPatient</i>	Um prontuário está associado a uma pessoa.	<i>composition</i>	1..1

B.14 CLASSE: RESOURCE

Classe	<i>Resource</i>
Descrição	Classe que abstrai um recurso que representa a informação que é armazenada como prontuário eletrônico ou preservada.
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Resource*

B.14.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasRule</i>	Um recurso pode ter regras de acesso (políticas)	<i>composition</i>	0..*
<i>hasSemantic</i>	Um recurso está associado a uma ontologia que o descreve.	<i>reference</i>	1..*
<i>hasStructure</i>	Um recurso pode estar associado com	<i>reference</i>	0..1
<i>conformsTo</i>	Um recurso pode estar descrito por um conjunto de padrões.	<i>reference</i>	0..*

B.15 CLASSE: RECORD COMPONENT

Classe	<i>Record_Component</i>
Descrição	Representa as características comuns de qualquer informação registrada no prontuário.
Sub classe de:	<i>Resource</i>

B.16 CLASSE: COMPOSITION

Classe	<i>Composition</i>
Descrição	Representa o conjunto de informações registrado em uma entidade de saúde por um agente como resultado de um único encontro clínico ou sessão de documentação. Exemplo laudo, exames.
Sub classe de:	<i>Record_Component</i>

A seguir são descritas as relações da classe *Composition*

B.16.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>updates</i>	Uma instância de <i>Composition</i> atualiza um objeto digital que será preservado (<i>ContentInformation</i>)	<i>reference</i>	0..1
<i>content</i>	Uma instância de <i>Composition</i> pode estar composto por um conjunto de entradas e seções (derivados da classe <i>Content</i>)	<i>composition</i>	0..*
<i>providedBy</i>	Referência a um evento de saúde	<i>reference</i>	1..1

B.17 CLASSE: FOLDER

Classe	<i>Folder</i>
Descrição	Representa o organizador de mais alto nível do prontuário eletrônico. Divide-se em compartimentos relativos à assistência prestada para um problema ou por uma instituição clínica ou durante um tempo determinado como um episódio.

Sub classe de:	<i>Record_Component</i>
-----------------------	-------------------------

A seguir são descritas as relações da classe *Folder*

B.17.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>compositions</i>	Um folder organiza um conjunto de instâncias da classe <i>composition</i>	<i>composition</i>	0..*
<i>subFolders</i>	Representa uma hierarquia de <i>Folders</i>	<i>composition</i>	0..*

B.18 CLASSE: CONTENT

Classe	<i>Content</i>
Descrição	Representam informações a serem armazenadas no prontuário em um alto nível de abstração, pode se uma seção (<i>Section</i>) ou uma entrada (<i>Entry</i>).
Sub classe de:	<i>Record_Component</i>

B.19 CLASSE: ENTRY

Classe	<i>Entry</i>
Descrição	Representa a informação registrada no prontuário eletrônico como resultado de uma ação clínica, uma observação clínica ou um propósito clínico.
Sub classe de:	<i>Content</i>

A seguir são descritas as relações da classe *Entry*

B.19.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>itens</i>	Uma entrada do prontuário (<i>Entry</i>) pode estar composta por grupos de dados (<i>Cluster</i>) ou elementos (<i>Element</i>).	<i>composition</i>	0..*

B.20 CLASSE: SECTION

Classe	<i>Section</i>
Descrição	Representa um agrupado de dados dentro de uma composição e reflete o fluxo de informação durante um encontro clínico (antecedentes, observações, etc.) o que representa a compreensão futura da composição.
Sub classe de:	<i>Content</i>

A seguir são descritas as relações da classe *Section*

B.20.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>members</i>	Uma seção por sua vez pode estar composta por um conjunto de seções e entradas.	<i>composition</i>	0..*

B.21 CLASSE: ITEM

Classe	<i>Item</i>
Descrição	Representa uma informação armazenada no prontuário eletrônico resultado de uma ação clínica, por exemplo: sintomas, uma observação, reação alérgica, etc.
Sub classe de:	<i>Record_Component</i>

B.22 CLASSE: CLUSTER

Classe	<i>Cluster</i>
Descrição	Representa um organizador de estruturas de dados complexas como dados interconectados, séries temporais ou tabelas. Por exemplo a pressão a arterial está composto por dois elementos: sistólica e diastólica.
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Cluster*

B.22.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>parts</i>	Um <i>cluster</i> pode estar composto por um conjunto de <i>Itens</i>	<i>composition</i>	0..*

B.23 CLASSE: ELEMENT

Classe	<i>Element</i>
Descrição	Representam os nodos, folhas de hierarquia que contém um único valor de um tipo de dados. Por exemplo Sistólica
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Element*

B.23.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasValue</i>	Um elemento está composto por um valor.	<i>composition</i>	0..1

B.24 CLASSE: VALUE

Classe	<i>Value</i>
Descrição	Representa o valor que será atribuído a um elemento.
Sub classe de:	<i>EClass</i>

B.25 CLASSE: PACKAGE

Classe	<i>Package</i>
Descrição	O pacote representa um organizador de objetos digitais preservados
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Package*

B.25.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>subPackage</i>	Permite criar hierarquias de pacotes.	<i>composition</i>	0..*
<i>hasContentInformation</i>	Um pacote pode ter vários objetos digitais.	<i>composition</i>	0..*

B.26 CLASSE: CONTENT INFORMATION

Classe	<i>ContentInformation</i>
Descrição	Objeto digital a ser preservado, por exemplo um laudo, uma imagem médica.
Sub classe de:	<i>Resource</i>

A seguir são descritas as relações da classe *ContentInformation*

B.26.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasMetadata</i>	Um objeto digital pode estar composto por um conjunto de metadados.	<i>composition</i>	0..*
<i>describedBy</i>	Um objeto digital pode estar associado a um plano de preservação (<i>PreservationPlan</i>).	<i>reference</i>	0..1
<i>provideDataFor</i>	Um objeto digital é disponibilizado para sua recuperação e consultas por aplicativos de software externos.	<i>reference</i>	1..*

B.27 CLASSE: REPOSITORY

Classe	<i>Repository</i>
Descrição	Repositório de objetos digitais que devem ser preservados
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Repository*

B.27.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasPackage</i>	Um repositório está formado por um conjunto de pacotes.	<i>composition</i>	1..*

B.28 CLASSE: METADATA

Classe	<i>Metadata</i>
Descrição	Representa os metadados associados a um objeto digital.
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *Metadata*

B.28.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Nome do metadado.	<i>EString</i>

B.28.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>has</i>	Metadados estão associados a valores.	<i>composition</i>	0..*

B.29 CLASSE: METADATA VALUE

Classe	<i>MetadataValue</i>
Descrição	É o valor do metadado, ou seja, a unidade mínima de preservação.
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *MetadataValue*

B.29.1 Atributos

Atributo	Descrição	Tipo
<i>value</i>	Valor atribuído a um metadado.	<i>EJavaObject</i>

B.30 CLASSE: PRESERVATION PLAN

Classe	<i>PreservationPlan</i>
Descrição	Representa o plano de preservação de dados clínicos definido por um curador.
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *PreservationPlan*

B.30.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Nome do plano de preservação	<i>EString</i>
<i>description</i>	Descrição do plano de preservação	<i>EString</i>

B.30.2 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>performs</i>	Um plano de preservação está composto por um conjunto de atividades a serem executadas.	<i>composition</i>	1..*

B.31 CLASSE: CURATOR

Classe	<i>Curator</i>
Descrição	Profissional ou conjunto de profissionais responsáveis por criar planos de preservação de dados clínicos.
Sub classe de:	<i>Actor</i>

A seguir são descritas as relações da classe *Curator*

B.31.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>actsOn</i>	Um curador está associado a um ou vários planos de preservação de dados clínicos.	<i>reference</i>	1..*

B.32 CLASSE: ACTIVITY

Classe	<i>Activity</i>
Descrição	Tarefa ou atividade a ser executada como consequência de um plano de preservação.
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos e relações da classe *Activity*

B.32.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Nome atribuído a uma atividade.	<i>EString</i>
<i>description</i>	Descrição da atividade.	<i>EString</i>
<i>goal</i>	Objetivo da atividade.	<i>EString</i>

B.32.2 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>subActivities</i>	Permite criar hierarquias de atividades.	<i>reference</i>	0..*

B.33 CLASSE: STANDARD

Classe	<i>Standard</i>
Descrição	Representa o padrão ou norma ao qual está um recurso está atrelado. Este padrão é um documento formado por um conjunto de especificações, procedimentos e diretrizes. Exemplo: DICOM
Sub classe de:	<i>EClass</i>

A seguir são descritos os atributos classe *Standard*

B.33.1 Atributos

Atributo	Descrição	Tipo
<i>name</i>	Nome do padrão.	<i>EString</i>
<i>description</i>	Documento ou ontologia que descreve do padrão.	<i>EString</i>
<i>type</i>	Tipo de padrão.	<i>EString</i>

B.34 CLASSE: STRUCTURE

Classe	<i>Structure</i>
Descrição	Estrutura de um recurso que transmite significado de como este é organizado.
Sub classe de:	<i>EClass</i>

B.35 CLASSE: SEMANTIC

Classe	<i>Semantic</i>
Descrição	Indica o significado de linguagem usada na descrição de um recurso.
Sub classe de:	<i>EClass</i>

B.36 CLASSE: POLICY

Classe	<i>Policy</i>
Descrição	Define categorias de políticas de acesso a recursos.
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Policy*

B.36.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>subPolicy</i>	Permite criar hierarquias de políticas de acesso	<i>composition</i>	0..*

B.37 CLASSE: RULE FEATURE

Classe	<i>RuleFeature</i>
Descrição	Define uma característica de uma regra, por exemplo uma ação, objetivo, contexto, etc.
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *RuleFeature*

B.37.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasValue</i>	Valor de uma instância de <i>RuleFeature</i>	<i>composition</i>	0..1

B.38 CLASSE: RULE

Classe	<i>Rule</i>
Descrição	Representa uma política de acesso no formato de uma regra
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Rule*

B.38.1 Relações

Relação	Descrição	Tipo	Cardinalidade
<i>hasFeatureRule</i>	Uma regra está composta por um conjunto de características.	<i>composition</i>	1..*
<i>hasSubject</i>	Uma regra está composta por um indivíduo (<i>Subject</i>).	<i>composition</i>	1..1

B.39 CLASSE: SUBJECT

Classe	<i>Subject</i>
Descrição	Entidade que solicita permissões para acessar um recurso e realizar ações sobre ele.
Sub classe de:	<i>EClass</i>

A seguir são descritas as relações da classe *Subject*

B.39.1 Relação

Relação	Descrição	Tipo	Cardinalidade
<i>hasCredential</i>	Permite atribuir uma credencial a uma instância de <i>Subjet</i>	<i>composition</i>	1..1

<i>hasRole</i>	Permite atribuir um papel a uma instância de <i>Subject</i>	<i>composition</i>	1..1
----------------	---	--------------------	------

B.40 CLASSE: RULE FEATURE VALUE

Classe	<i>RuleFeatureValue</i>
Descrição	Representa o valor de regra de política de acesso
Sub classe de:	<i>EClass</i>

B.41 CLASSE: CREDENTIAL

Classe	<i>Credential</i>
Descrição	Representa uma credencial de acesso que uma instância de <i>Subject</i> pode ter.
Sub classe de:	<i>EClass</i>

B.42 CLASSE: ROLE

Classe	<i>Role</i>
Descrição	Representa o papel que uma instância de <i>Subject</i> pode ter, por exemplo: médico, enfermeira.
Sub classe de:	<i>EClass</i>

B.43 CLASSE: ACTOR

Classe	<i>Actor</i>
Descrição	Classe que abstrai os diferentes atores no metamodelo de preservação de dados clínicos. Exemplo: paciente, curador, especialistas, etc.
Sub classe de:	<i>EClass</i>

B.44 CLASSE: ACCESS

Classe	<i>Access</i>
Descrição	Classe que abstrai os diferentes SIS ou aplicativos usados para consultas, procurar ou exibir os documentos clínicos preservados.
Sub classe de:	<i>EClass</i>