



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

**Modelo de Interoperabilidade Semântica Aplicado ao
Domínio da Saúde: Um Estudo de Caso na
Vigilância Alimentar e Nutricional**

Rodrigo André Cuevas Gaete

Brasília
2012



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

**Modelo de Interoperabilidade Semântica Aplicado ao
Domínio da Saúde: Um Estudo de Caso na
Vigilância Alimentar e Nutricional**

Rodrigo André Cuevas Gaete

Dissertação apresentada como requisito parcial
para conclusão do Mestrado em Informática

Orientadora
Prof.^a Dr.^a Célia Ghedini Ralha

Brasília
2012

Universidade de Brasília – UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Mestrado em Informática

Coordenador: Prof. Dr. Maurício Ayala Rincón

Banca examinadora composta por:

Prof.^a Dr.^a Célia Ghedini Ralha (Orientadora) – CIC/UnB

Prof. Dr. Marcelo Ladeira – CIC/UnB

Prof.^a Dr.^a Marisa Brascher Basilio Medeiros – CIN/UFSC

CIP – Catalogação Internacional na Publicação

Gaete, Rodrigo André Cuevas.

Modelo de Interoperabilidade Semântica Aplicado ao Domínio da Saúde: Um Estudo de Caso na Vigilância Alimentar e Nutricional / Rodrigo André Cuevas Gaete. Brasília : UnB, 2012.

109 p. : il. ; 29,5 cm.

Dissertação (Mestrado) – Universidade de Brasília, Brasília, 2012.

1. Mapeamento Ontológico, 2. Interoperabilidade Semântica, 3. Saúde Pública

CDU 004.9

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro – Asa Norte
CEP 70910-900
Brasília – DF – Brasil

Dedicatória

Eu dedico este trabalho ao povo brasileiro que usa o Sistema Único de Saúde (SUS), sistema este que se apresenta como um dos melhores modelos de sistemas de saúde do mundo, mas que ainda tem muitos desafios a enfrentar.

Agradecimentos

Agradeço primeiramente aos meus pais, Mario e Patrícia, por me fazer chegar até aqui e com muita sabedoria me ensinaram valores como honestidade, humildade, justiça, amor e tantos outros valores que nos ajudam a ser uma pessoa melhor.

Gostaria de agradecer a toda minha família que por muitos momentos precisei me afastar para alcançar o objetivo de concluir este curso. Em especial a Mariana, minha esposa, minha companheira para todas as horas que eu amo tanto e que sempre me apoia nas minhas decisões e me ajuda a enfrentar os problemas da vida.

Ao Ministério da Saúde, por me permitir em vários momentos me afastar das minhas atividades no trabalho para enfrentar mais este desafio.

Ao governo brasileiro, por me garantir acesso a uma universidade pública, gratuita e de qualidade.

Aos professores e mestres que me ajudaram em vários momentos, através de conselhos, discussões e na minha formação como profissional.

Em especial quero agradecer à Profa. Dra. Célia Ghedini Ralha, minha orientadora, pela confiança, dedicação e ousadia em trabalhar com um tema ainda pouco difundido no Brasil.

Resumo

A histórica desarticulação dos Sistemas de Informação em Saúde (SIS) no Brasil, em especial dos sistemas do Sistema Único de Saúde (SUS), tem fragilizado a estrutura de gestão em saúde nas três esferas governamentais. A área de Informática em Saúde em conjunto com a área de Engenharia de Ontologias tem colaborado com técnicas e métodos de representação de conhecimento como instrumentos fundamentais para viabilizar a interoperabilidade semântica dos SIS. Nesta direção, este trabalho apresenta uma abordagem metodológica aplicada ao domínio da saúde de modelagem em dois níveis, o qual utiliza: (i) um modelo de referência, de mais alto nível ontológico, formado por conceitos amplamente consensuados, constituindo uma camada invariante de conceitos; (ii) modelos de conhecimento, chamado de arquétipos, consensuado em nível de aplicação, que busca através de múltiplas camadas de conceitos representar um nível mais flexível. Os arquétipos são utilizados como estruturas de representação de conhecimento que vem substituindo os níveis mais complexos de representação ontológica, os quais estruturam conhecimento em um nível mais próximo das aplicações. O tratamento de níveis variantes de conhecimento são viabilizados pelos arquétipos, possibilitando um mapeamento das ontologias de mais alto nível em ontologias de aplicação. Com a finalidade de avaliar a estrutura metodológica definida foi realizada uma prova de conceito utilizando o Sistema de Informação da Vigilância Alimentar e Nutricional brasileiro, no qual foram representados os conceitos e as informações do sistema. A partir destas estrutura de conceitos foram gerados extratos de dados semanticamente enriquecidos, mostrando um modelo de mapeamento ontológico de SIS. Os resultados alcançados pelos modelos definidos indicam que a modelagem de dois níveis apresenta uma solução viável para integrar SIS através de camadas ontológicas. O mapeamento por arquétipos entre níveis ontológicos reforçam o tratamento dos sistemas do SUS de forma consistente e interpretável.

Palavras-chave: Mapeamento Ontológico, Interoperabilidade Semântica, Saúde Pública

Abstract

Health Information Systems (HIS) in Brazil has been historically disarticulated, particularly the information systems of the Brazilian Unified Health System (UHS). Three government spheres have been weakened in their management health structure because of this disarticulation. Studies in the area of Semantic Web have put the health ontologies as a fundamental part to enable the semantic interoperability of HIS and enable integrated access to information. In this sense we present the methodological approach applied in the health domain, two levels methodology: (i) reference model, the first level, is used a higher level ontological concepts made up of consensual concepts, constituting an invariant layer concepts, (ii) knowledge models, called archetype, the application level, which searches through multiple layers of concepts represent a more flexible level of the same. The archetypes are used as frameworks for knowledge representation that is replacing the more complex levels of ontological representation. In practice, it is clear that the archetypes have been building a knowledge-level closer to the applications, i.e., variant levels of knowledge over the time, thus providing a mapping of ontologies to the highest level in the lower-level ontologies. A case study was conducted to assess the methodological framework defined. The Brazilian Information System for Food and Nutrition Surveillance was used in this work. Where archetypes are defined to represent the concepts and system information. Semantically enriched data extracts are generated from archetypes, showing a model of ontological mapping of HIS. In conclusion, the results achieved in applying the models in the study case, we present evidence that two-level modeling provides a viable solution to join HIS through ontological layers, which enables the mapping between ontological levels archetypes to enhance the treatment of systems of the UHS and consistently interpretable.

Keywords: Ontological Mapping, Semantic Interoperability, Public Health Care

Lista de Figuras

2.1	Classificação de Ontologias	11
2.2	Categorização da Representação de Conhecimento	15
2.3	Os três aspectos da linguagem OIL.	17
2.4	Relação entre as linguagens OWL, RDFS e RDF.	17
3.1	Modelo Simples	19
3.2	Modelo Dual	20
3.3	Classes do padrão openEHR	21
3.4	Processo de Investigação clínica	22
3.5	Classes da Norma CEN/ISO 13606	23
3.6	Estrutura da linguagem ADL	26
3.7	Relação entre modelo de referência e conhecimento	27
3.8	Sistema Único de Saúde	31
3.9	Interfaces da Atenção Primária no Brasil	33
4.1	Relação do RES com a ABS	36
4.2	Associação de conceitos usando arquétipos	37
4.3	Macro-etapas da Proposta	38
4.4	Relação entre os principais modelos de referências	39
4.5	Comparação entre os modelos de referências OpenEHR e CEN/ISO 13606	39
4.6	Exemplo de estruturação do SOAP usando openEHR	40
4.7	Metodologia para desenvolver de arquétipos	41
4.8	Mapa mental do conceito Peso Corporal	42
4.9	Mapa mental do conceito Peso Corporal usando arquétipos openEHR	43
4.10	Extrato RES, openEHR, de SIS	45
4.11	Solicitando extratos openEHR	46
4.12	Tela gerado pelo openEHR-Gen-Framework.	48
5.1	Recorte do Formulário do SISVAN	51
5.2	Recorte Banco de Dados do SISVAN	52
5.3	Mapa Mental do Conceito da VAN	58
5.4	Mapa Mental dos Conceitos de VAN por Fase da Vida	60
5.5	Mapa Mental de Acompanhamento Nutricional no padrão openEHR	63
5.6	Acompanhamento Nutricional no LinkEHR	65
5.7	“Arquetipando” informação	67
B.1	Formulário, frente, Dados Cadastrais	82
B.2	Formulário, frente, Dados de Acompanhamento Nutricional	83

B.3 Formulário, verso, Legendas	83
---	----

Lista de Tabelas

5.1	Tabela de variáveis do SISVAN	53
5.2	Arquétipos já definidos nos repositório do openEHR e do NEHTA	59

Lista de Abreviaturas e Siglas

ABS	Atenção Básica à Saúde
ADL	Archetype Definition Language
CNES	Cadastro Nacional de Estabelecimentos de Saúde
CadSUS	Cadastro Nacional do SUS
CDA	Clinical Document Architecture
CIAP	Classificação Internacional da Atenção Primária
CID	Classificação Internacional de Doenças
CPIISS	Catálogo de Padrões de Interoperabilidade de Informação de Sistemas de Saúde
DATASUS	Departamento de Informação e Informática do SUS
DICOM	Digital Imaging and Communications in Medicine
EHR	Electronic Health Record
HL7	Health Level Seven
IS	Informática em Saúde
LOINC	Logical Observation Identifiers Names and Codes
MS	Ministério da Saúde
OMS	Organização Mundial de Saúde
OWL	Web Ontology Language
PNIIS	Política Nacional de Informação e Informática em Saúde
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
RES	Registro Eletrônico de Saúde
SIS	Sistema de Informação de Saúde
SISVAN	Sistema de Informação de Vigilância Alimentar e Nutricional
Snomed-CT	Systematized Nomenclature of Medicine - Clinical Terms
SOAP	Subjetivo, Objetivo, Avaliação e Plano
SUS	Sistema Único de Saúde
XML	EXtensible Markup Language
XML	EXtensible Markup Language Schema
TIC	Tecnologia de Informação e Comunicação
VAN	Vigilância Alimentar e Nutricional
W3C	World Wide Consortium

Sumário

1	Introdução	1
1.1	Motivação	2
1.2	Apresentação do Problema	4
1.3	Objetivos	4
1.4	Solução Adotada	5
1.5	Contribuições deste trabalho	5
1.6	Apresentação do Documento	6
2	Aspectos Semânticos	7
2.1	Interoperabilidade Semântica	7
2.1.1	Heterogeneidade Sintática e Estrutural	8
2.1.2	Heterogeneidade Semântica	8
2.2	Ontologia	9
2.2.1	Classificação de Ontologias	10
2.2.2	Uso da Ontologia em Sistemas de Informação	11
2.2.3	Problemas no Uso de Ontologias	13
2.2.4	Outras Forma de Representar Conhecimento	13
2.2.5	Meta-Ontologias	16
2.2.6	Web Semântica	16
3	Semântica na Saúde	18
3.1	Modelagem em dois níveis	18
3.1.1	Modelos de Referência	20
3.1.2	Modelo de Conhecimento	24
3.1.3	Relação entre os Modelos de Referência e de Conhecimento	26
3.2	Representação do Conhecimento na Saúde	27
3.2.1	Ontologia no Domínio da Saúde	27
3.2.2	Outros Modelos no Domínio da Saúde	29
3.3	Sistema Único de Saúde	31
3.4	Informática em Saúde	34
3.5	Registro Eletrônico de Saúde	34
4	Proposta	36
4.1	Seleção do Modelo de Referência	38
4.1.1	Comparando Modelos de Referência	38
4.2	Desenvolvimento de Arquétipos	40
4.2.1	Metodologia de Desenvolvimento de Arquétipos	41

4.3	Extraindo dados semanticamente enriquecidos	44
4.3.1	Delimitação do Escopo de Trabalho	45
4.4	Trabalhos Correlatos	46
5	Prova de Conceito usando SISVAN	50
5.1	Seleção dos Elementos	50
5.2	Identificação dos conceitos candidatos	52
5.2.1	Analisando as variáveis	53
5.2.2	Analisando as orientações do SISVAN	54
5.2.3	Conceitos Identificados	56
5.3	Pesquisa dos arquétipos existentes	58
5.3.1	O Problema da definição múltipla da DUM	60
5.3.2	Generalização dos Conceito de Estado de Peso e IMC	61
5.4	Análise e identificação de terminologias	61
5.5	Modelagem dos Arquétipos	62
5.6	Codificação de arquétipos em ADL	63
5.6.1	Ferramentas para edição de ADL	63
5.6.2	Arquétipos modelados	64
5.7	Análise dos Resultados	67
6	Conclusões e Trabalhos Futuros	70
	Referências	72
A	CPIISS - Ministério da Saúde	80
A.1	Catálogo de Padrões de Informação	80
B	Formulário do SISVAN	82
C	Arquétipo do SISVAN em linguagem ADL	84
D	Arquétipo de Estatura em linguagem ADL	89
E	Extrato de Informação do SISVAN em XML	93

Capítulo 1

Introdução

A gestão integrada dos serviços de saúde e a continuidade do cuidado da população requerem um aprimoramento no uso de Tecnologia de Informação e Comunicação (TIC) para registro de saúde de tal forma que permita a interoperabilidade entre os Sistemas de Informação em Saúde (SIS). Considerando o uso de TIC na saúde, existe uma demanda cada vez maior por sistemas abertos, distribuídos, interconectados, interoperáveis e com um alto grau de confiabilidade e segurança a um custo cada vez mais acessível a grandes e pequenas empresas do setor saúde.

A busca por implantação de padrões tem sido bastante importante em setores mais tradicionais da indústria, aumentando a qualidade do processo e diminuindo custo de operação. Existem importantes argumentos, a favor, para considerar aspectos similares no setor da saúde. A área de Informática em Saúde (IS) tradicionalmente é formada por um mercado de software fragmentado, inúmeras aplicações incompatíveis, altos custos no desenvolvimento de soluções desintegradas do SUS, curto ciclo de utilidade da solução somado a problemas de manutenção, entre outros, os quais vem dificultando a interoperabilidade de sistemas em saúde e conseqüentemente a gestão e o cuidado continuado.

Neste sentido os especialistas da área de IS apontam para a adoção de padrões técnicos de desenvolvimento como um elemento necessário e estratégico para modelagem, especificação e implementação de SIS, em especial de sistemas baseados em Registro Eletrônico de Saúde (RES), e que deveria favorecer a produção de soluções mais estáveis, a menor custo e com mais capacidade de interoperar com outras soluções [3].

Vários pesquisadores, Guarino [42], Gómez-Pérez et al. [47], Hakimpour and Geppert [48], Alexiev and et al. [1], Gruber [41], entre outros, apontam a formalização ontológica de aplicações como uma forma de se resolver interoperabilidade e integração entre sistemas. Parte fundamental para possibilitar interoperabilidade entre sistemas, em especial dos SIS, é viabilizar a criação de padrões de informação pela criação de ontologias que representem o domínio da atenção à saúde [11, 53, 62].

Com a evolução das pesquisas e consolidação das áreas de Engenharia de Ontologia e IS, já é possível utilizar padrões de informação para formalizar conhecimento re-utilizável e em camadas, com alto grau de consenso sobre os conceitos na área de atenção à saúde, e com isso promover a interoperabilidade semântica de SIS. Este trabalho pesquisa como viabilizar a interoperabilidade semântica de SIS através do uso de formalização ontológica de aplicação no domínio da saúde.

1.1 Motivação

Depois de vários anos da criação do Ministério da Saúde (MS) em 1953, e muitas reformas dentro da estrutura política do Governo Federal, com a Constituição Federal de 1988 foi criado o Sistema Único de Saúde (SUS). O SUS teve detalhado seu funcionamento no Congresso Nacional através da Lei Orgânica da Saúde, em 1990. Com a criação do SUS, em 1991 criou-se também a Fundação Nacional de Saúde (FUNASA), definindo o Departamento de Informática do SUS (DATASUS) como órgão integrante da estrutura básica da FUNASA [24]. Ao DATASUS, que em 1998 foi incorporado a estrutura do Ministério da Saúde, lhe compete “especificar, desenvolver, implantar e operar sistemas de informação relativos às atividades finalísticas do SUS, em consonância com as diretrizes do órgão setorial” [24].

Atualmente o SUS tem aproximadamente trinta SIS que mantem a estrutura básica de gestão do SUS e dos programas do Governo Federal [23]. Existem ainda centenas de sistemas que mantem informação de projetos, a nível de gestão estadual e de gestão municipal. Os sistemas de informação com foco em saúde, com poucas exceções, são totalmente desarticulados e buscam resolver seus problemas de gestão de informação de forma isolada. Conforme [4], um diagnóstico realizado pela Rede Interagencial de Informações para a Saúde (RIPSA) do MS, datado de 1996, mostra um pouco da origem dos problemas que hoje enfrentamos:

1. A informação não é adequadamente utilizada como um requisito fundamental do processo de decisão-controle aplicada à gestão de políticas e ações de saúde;
2. Os múltiplos sistemas de informação existentes são desarticulados, insuficientes, imprecisos e não contemplam a multicausalidade dos fatores que atuam no binômio saúde e doença;
3. Inexistem processos regulares de análise da situação de saúde e de suas tendências, de avaliação de serviços e de difusão da informação;
4. O planejamento, a organização e a avaliação dos serviços não estão epidemiologicamente sustentados¹.

Apesar de diagnosticado, poucos intentos, ao longo desses mais de vinte anos, foram ajustando essa estrutura que sustenta atualmente o SUS. Uma solução eminente aos problemas causados pela desarticulação dos SIS do SUS são os sistemas de integração de dados. Um sistema de integração de dados é um sistema capaz de interfacear várias fontes de dados de forma única, deixando transparente ao usuário as diferenças das fontes de dados que se deseja acessar, possibilitando a uniformização e o cruzamento dos dados de uma organização [56].

O conceito de integração de dados dentro do SUS é amplamente conhecido e desejado pelos Departamentos e Coordenações do MS e pelas Secretarias Estaduais e Municipais de Saúde Pública, porém ainda existe uma grande fragilidade das estruturas de informática

¹Epidemiologia é o estudo da distribuição e dos determinantes de estados ou eventos relacionados à saúde em populações específicas, e sua aplicação na prevenção e controle dos problemas de saúde, ou seja, a epidemiologia não somente, está preocupada com a incapacidade, doença ou morte, mas, também, com a melhoria dos indicadores de saúde e com maneiras de promover saúde [14].

e informação de atender essa demanda nas três esferas. O tema da integração de dados e sistemas do SUS tem sido amplamente discutido nas Conferências Nacionais de Saúde realizadas pelo MS e também é constantemente citado nos Encontros Epidemiológicos Nacionais. Em especial na XII Conferência Nacional de Saúde (2004) foram aprovadas deliberações importantes na plenária:

“O Conselho Nacional de Saúde deverá definir estratégias, em conjunto com o Ministério da Saúde, para elaborar e implementar políticas articuladas de informação, comunicação, educação permanente e popular em saúde, para as três esferas de governo, garantindo maior visibilidade das diretrizes do SUS, da política de saúde, ações e utilização de recursos, visando ampliar a participação e o controle social e atender as demandas e expectativas sociais, permitindo: Reforçar a democratização da informação e da comunicação, em todos os aspectos, garantir, nas três esferas de governo, com definição de prazos, a compatibilização, interface e modernização dos sistemas de informação do SUS e o aperfeiçoamento da integração e articulação com os sistemas e bases de dados de interesse para a saúde ... ”²

Estas deliberações se refletem na Política Nacional de Informação e Informática em Saúde (PNIIS), criada em 2004, pelo DATASUS e definem várias diretrizes para alcançar esses propósitos [35]. Duas das diretrizes inseridas na PNIIS podem ser destacadas no âmbito da integração de dados:

1. Estabelecer sistema de identificação unívoca de usuários, profissionais e estabelecimentos de saúde que seja progressivamente adotado, aprimorando o processo de integração dos sistemas de informação de saúde e viabilizando o **registro eletrônico de saúde**. O Cartão Nacional de Saúde (CadSUS) - que identifica univocamente usuários e profissionais - e o Cadastro Nacional de Estabelecimentos de Saúde (CNES) - que identifica univocamente os estabelecimentos - são o passo inicial na construção deste novo paradigma.
2. Fortalecer as áreas de informação e informática nas três esferas de governo, apoiando a sua organização e desenvolvimento, através de:
 - criação de mecanismos de articulação, com vistas à integração dos sistemas de informação em saúde;
 - estabelecimento de mecanismos que permitam a manutenção de quadro permanente de profissionais de informação e informática em saúde.

Estas medidas auxiliam fortemente na integração dos sistemas, primeiramente por permitir a criação de estruturas de identificação dos objetos do SUS e também por articular o conhecimento profissional da complexidade que são os sistemas de informação do SUS. Em Setembro de 2010, a Comissão Intergestores Tripartite (CIT)³, aprovou o Catálogo de Padrões de Interoperabilidade de Informação de Sistemas de Saúde (CPIISS), definido pelo Comitê de Informação e Informática em Saúde (CINFO) do MS. O CPIISS

²CONFERÊNCIA NACIONAL DE SAÚDE, 12 ., 2003, Brasília. Relatório Final: Ministério da Saúde, 2004. Conferência Sérgio Arouca.

³CIT é a instância de articulação e pactuação na esfera federal que atua na direção nacional do SUS (MS) integrada por gestores do SUS das três esferas de governo [23].

foi submetido a consulta pública, em Maio 2011, de onde recebeu diversas considerações dos grupos de pesquisas que vem discutindo os padrões de interoperabilidade no Brasil, por exemplo, da Sociedade Brasileira de Informática em Saúde (SBIS), da Associação Brasileira de Normas Técnicas (ABNT) por meio do Grupo de Informática em Saúde, de empresas do ramo de desenvolvimento de sistemas em saúde, dentre outros. Em agosto de 2011, foi publicada a Portaria GM nº 2.073 do MS que:

“Regulamenta o uso de Padrões de Interoperabilidade e Informação em Saúde para sistemas de informação em saúde no âmbito do Sistema Único de Saúde, nos níveis municipal, distrital, estadual e federal, e para os sistemas privados e do setor de saúde suplementar.”

Este processo mostra a clara intenção do MS de definir uma plataforma de interoperabilidade no contexto nacional e busca estreitar o processo de integração de seus próprios sistemas de informação e fomentar um processo de interoperabilidade semântica. No Apêndice A, pode ser vista a lista completa de padrões de interoperabilidade definidos na portaria que define o CPIISS. Durante o texto iremos destacar e explicar alguns dos padrões de informação definidos pelo CPIISS.

1.2 Apresentação do Problema

Problema de interoperabilidade semântica na atenção a saúde é um grande desafio frente ao complexo domínio da saúde. Resolver o problema da interoperabilidade semântica significa enfrentar uma mudança de paradigma na construção de SIS. Apesar de vários esforços, atualmente não existe um modelo de interoperabilidade semântica aplicado no SUS em nível nacional.

O MS apresenta o catálogo de padrões do CPIISS na expectativa de apoiar a definição de padrões de modelagem conceitual e de padrões de comunicação para SIS. No entanto, não apresenta políticas claras de como esse processo de padronização deve ser efetivamente realizado. Ademais, existem poucos trabalhos de pesquisa para definir padrões de modelagem conceitual e de integração de SIS.

No SUS, em especial na Atenção Básica, como será visto na Seção 3.3, é importante garantir que o registro de saúde de um paciente, como por exemplo, de um prontuário eletrônico, não se perca, visando a continuidade do cuidado de seus usuários e apoiando os profissionais de saúde em melhores decisões em relação a avaliações e diagnósticos, tanto no cuidado coletivo quanto no cuidado individual.

Para garantir a continuidade do cuidado da população se faz necessário conectar, quando possível, o velho paradigma de desenvolvimento com a nova proposta usando padrões de informação.

1.3 Objetivos

Nessa perspectiva, este trabalho se propõe a estudar uma solução ao problema apresentado, focando nos modelos de interoperabilidade semântica para integração de SIS, buscando aderência dos sistemas atualmente utilizados pelo SUS aos padrões definidos

pelo Ministério da Saúde no CPIISS, e com isso viabilizar a consulta a registro realizados antes da adoção de padrões de informação.

Este trabalho tem o objetivo de descrever um modelo de interoperabilidade semântica para o domínio da saúde usando padrões do CPIISS. Para verificação do modelo de interoperabilidade será desenvolvida uma prova de conceito através com um estudo teórico-prático no Sistema de Informação da Vigilância Alimentar e Nutricional (SISVAN). Este modelo é verificado pela extratificação de informações do SISVAN, o qual é atualmente usado para fazer o monitoramento da Vigilância Alimentar e Nutricional (VAN) no SUS.

Especificamente este trabalho tem o objetivo de:

- apresentar o processo de construção de artefatos ontológicos utilizando técnicas de modelagem em dois níveis, proposta pelo padrão da Fundação openEHR para modelagem de RES, e os conceitos relacionados a Engenharia de Ontologias;
- descrever um método de desenvolvimento de arquétipos com vistas a garantir interoperabilidade semântica;
- definir arquétipos, como artefatos ontológicos, no sub-domínio da VAN;
- gerar extratos de dados semanticamente enriquecidos por meio do conceitos representados em arquétipos.

1.4 Solução Adotada

Tendo em vista os objetivos citados, este trabalho de pesquisa estudou formas adequadas de modelagem conceitual e processo de construção de artefatos ontológicos. Foi descrito um método de desenvolvimento de arquétipos como um artefato ontológico para que estes artefatos possam ser reutilizados em aplicações reais do SUS.

Após estudar formas adequadas de se modelar o conhecimento e compreender as ontologias existentes no domínio da saúde, avançamos no estudo da modelagem em dois níveis usando as especificações do padrão openEHR. Também foi estudada a relação do padrão openEHR com os demais padrões definidos no CPIISS.

A partir desse estudo foram desenvolvidos arquétipos no sub-domínio da VAN, reutilizando arquétipos encontrados em repositórios. Foi descrito um método de desenvolvimento que potencialize interoperabilidade semântica dos SIS pelo uso de padrões de informação.

Na sequência foi realizado um experimento utilizando os arquétipos definidos, através da geração de extratos de dados do SISVAN, observando a capacidade do modelo de conhecimento de representar as informações reais do sistema em produção.

1.5 Contribuições deste trabalho

A contribuição deste trabalho é fazer uma prova de conceito do padrão openEHR, definido pelo CPIISS como padrão para RES, aplicado a SIS nacionais, ou seja, tendo com um dos principais requisitos uma visão mais ampla da aplicação em saúde e com isso permitir a interoperabilidade de SIS.

Este trabalho amplia a proposta metodológica de desenvolvimento de arquétipos proposta por Santos [71], buscando definir um modelo mais completo, sem considerar restrições da aplicação, e em especial buscando a construção de um método que garanta interoperabilidade semântica entre SIS.

Através de um experimentação buscou-se mostrar o uso de outros padrões definidos pelo CPIISS e por meio da geração de extratos de informação do SISVAN, mostramos que é possível enriquecer semanticamente extratos de dados, encapsulando os dados do sistema em extratos de RES restringidos pelo modelo de conhecimento definidos. Ainda foi possível mostrar que os arquétipos oferecem uma estrutura incremental de camadas ontológicas, ampliando a representação do conhecimento em camadas.

1.6 Apresentação do Documento

Os capítulos que seguem estão organizados em cinco capítulos. No Capítulo 2 são apresentados aspectos semânticos e formas de representação de conhecimento em uma visão mais geral da área de ontologia; no Capítulo 3 apresentamos a modelagem em dois níveis e uma discussão de ontologia aplicado ao domínio da saúde; no Capítulo 4 será explanada a proposta do trabalho; na sequência, Capítulo 5, apresentamos uma experimentação realizado na área de Vigilância Alimentar e Nutricional previsto para validar a proposta e tendo como resultado os arquétipos definidos, ainda é feita uma análise dos resultados aplicando os arquétipos na construção de extratos de dados semanticamente enriquecidos; seguido, no Capítulo 6, pelas conclusões provenientes deste trabalho de pesquisa.

Capítulo 2

Aspectos Semânticos

Este capítulo apresenta duas seções principais com a finalidade de apresenta os aspectos semânticos enquanto área de pesquisa. Na Seção 2.1, será apresentado o conceito de interoperabilidade semântica como objeto central para garantir integração de sistemas de informação. Na Seção 2.2, é apresentado o conceito de ontologia e como ela amplia a resolutividade dos problemas de interoperabilidade semântica.

2.1 Interoperabilidade Semântica

Grandes organizações precisam trocar informações entre os diversos sistemas desenvolvidos separadamente. Para que a troca de dados ou informações seja útil, os sistemas da organização devem obter um consenso sobre o significado do que será transmitido, ou seja, a organização deve assegurar a interoperabilidade semântica entre seus sistemas [73].

No entanto, faz-se necessário ter claro que anterior ao processo de tornar um sistema semanticamente interoperável os problemas de heterogeneidade de sistema e informação devem estar superados. Os sistemas que compõem a estrutura de uma média ou grande organização, geralmente, são construídos por pessoas diferentes, em momentos diferentes e portanto com grande chance de utilizarem tecnologias diferentes. Logo, são encontrados modelos de dados diversos, representações distintas para os dados, terminologias conflitantes e ainda questões de sub-utilização da tecnologia pelos mais diversos problemas sócio-econômicos. Os problemas de heterogeneidade podem ser de sistema ou de informação [75].

A heterogeneidade de sistema existe quando se implementam aplicações com diferentes sistemas de representação de dados em baixo nível¹, esta heterogeneidade está ligada ao uso de diferentes modelos e tecnologias. As diferenças podem ser de hardware, de comunicação ou ainda de software. Diferenças de software podem estar associadas ao uso de diferentes sistemas operacionais em um mesmo ambiente ou ainda diferentes sistema de armazenamento de arquivos ou dados, que estão disponíveis e são adotados em uma mesma empresa com o objetivo de facilitar o desenvolvimento dos sistemas assim como a adequação dos requisitos do negócio e que, podem estar distribuídos rodando em diferentes sistemas operacionais em diferentes versões [76].

¹Neste momento, entendemos dados de baixo nível aqueles referentes as trocas de informações entre software e hardware, em geral suportados pelos sistemas operacionais.

A heterogeneidade de informação é subdividida em heterogeneidade sintática, estrutural e semântica.

2.1.1 Heterogeneidade Sintática e Estrutural

A heterogeneidade sintática de informação, ocorre quando sistemas utilizam diferentes linguagens ou representação de dados para definir o mesmo conceito, e.g., é o uso de idiomas distintos na identificação de objetos ou entidades de um sistema.

A heterogeneidade estrutural ou de esquema de informação ocorre quando o mesmo conceito é representado por diferentes construtores do modelo de dados, ou seja, os construtores usados têm diferentes estruturas ou diferentes comportamentos. A forma de armazenar nome de pessoas em um sistema, é um exemplo comum deste tipo de heterogeneidade, alguns sistemas representam o nome por um campo descritivo “Nome Pessoa”, enquanto, outros sistemas representam pelos campos descritivos “Primeiro Nome”, “Nome do Meio” e “Último Nome”. Este tipo de heterogeneidade ainda pode ser relacionado a identidade de objetos, e ocorre quando a identificação de um objeto em sistemas ocorrem de formas diferentes, por exemplo, um funcionário de uma empresa ser identificado no sistema de recursos humanos pelo Número de Identificação do Trabalhador (NIT), enquanto que no sistema de gestão da empresa o mesmo funcionário ser identificado pelo Cadastro de Pessoa Física (CPF). O problema de identidade tem sido solucionado através de ligação ou encadeamento de registro (*Record Linkages*)², utilizando técnicas de pareamento probabilístico por múltiplos registros de identificação, e.g., o nome completo, o nome completo da mãe, a data de nascimento, e outros campos complementares como chaves de identificação [61, 78].

Já no século passado, o surgimento da computação distribuída, da tecnologia de *middleware* e dos padrões nos permitiu superar e apoiar particularmente os problemas de heterogeneidade sintática e estrutural, e nos permitiu abordar questões ao nível da informação [75]. Entretanto, a interoperabilidade semântica requer que sistemas de informação compreendam a semântica da solicitação de um usuário e as informações de cada fonte de informação e portanto ainda é foco de pesquisa soluções para o problema de heterogeneidade semântica.

2.1.2 Heterogeneidade Semântica

A heterogeneidade semântica considera a expressividade das informações nos diversos domínios de conhecimento. Ela se refere as diferenças ou similaridades no significado dos dados de cada base. Segundo Naiman and Ouksel [66], estas diferenças ocorrem por um desacordo entre o significado, a interpretação ou a intenção do uso dos mesmos dados, e podem ser classificadas da seguinte forma:

- semanticamente equivalentes ou sinônimos, termos diferentes são usados para referenciar o mesmo conceito;
- semanticamente não-relacionados ou homônimos, o mesmo termo usado em sistemas distintos para denotar coisas completamente diferentes;

² *Record Linkage* tem o objetivo de relacionar registros pela associação de atributos equivalentes entre as fontes de dados. Esta solução foi proposta por [50] e ampliada por [36]

- semanticamente relacionados, generalização e especificação, conceitos sobrepostos ou conceitualizações diferentes.

Ainda alguns conflitos podem ser detectados por falha na interpretação dos dados:

- conflito de dados, ocorre quando os valores dos dados de mesmo conceito são diferentes em diferentes fontes de dados, por exemplo, o Código de Endereçamento Postal (CEP) do endereço de um paciente ser diferente em cadastros pessoais de dois sistemas diferentes;
- conflitos de distinção, podem ser causados por uma interpretação equivocada dos dados, por exemplo, diferenças temporais da coleta de indicadores, ou seja, um sistema julga que o mês de referência de um indicador é o mês em que o indicador foi coletado, enquanto que outro sistema julga que a referência é a do mês imediatamente anterior;
- conflitos de escala, podem ocorrer quando sistemas de referência diferentes são usados para medir um valor, por exemplo, no armazenamento de um valor monetário usar diferentes moedas;

O problema da heterogeneidade semântica não é de simples resolução, em especial, devido a dimensão do problema de distinguir semanticamente cada conceito, no entanto, diversas áreas como a Ciência da Informação, Engenharia do Conhecimento e Inteligência Artificial vem utilizando representações ontológicas de conhecimento, por meio de ontologias, para resolver esse problema em diversos temas como por exemplo em aplicações relacionadas a gestão do conhecimento, processamento de linguagem natural, comércio eletrônico, integração inteligente de informações, recuperação da informação, integração de banco de dados, bio-informática, entre tantos outros [47].

2.2 Ontologia

Conforme Guizzardi [44], o termo ontologia foi cunhado no século XVII em paralelo pelos filósofos Rudolf Gockel e Jacob Lorhard, porém, foi popularizado nos meios filosóficos só no século XVIII com a publicação, em 1730, do livro *Philosophia prima sive Ontologia* por Christian Wolff.

Na Ciência da Computação Gruber [40] inseriu o termo na Inteligência Artificial e define ontologia como uma descrição dos conceitos e de relacionamentos que podem existir para um agente ou para uma comunidade de agentes. Guarino [42], definiu ontologia como um conjunto de axiomas lógicos concebidos para explicar o significado pretendido de um vocabulário. Atualmente, Gruber [41] define ontologia, na área de sistemas de bancos de dados, como um conjunto de representações primitivas (classes, atributos, propriedades e relacionamentos) com o qual pode ser modelado o domínio de um conhecimento ou discurso.

Diferentes definições fornecem pontos de vista diferentes e complementares de uma mesma realidade. Como as definições apresentadas, ainda podem ser encontradas diversas definições em diferentes áreas, alguns autores fornecem definições que são independentes dos processos seguidos para construir a ontologia e da sua utilização em aplicações, enquanto que outras definições são influenciados por seu processo de desenvolvimento. No

entanto, pode-se dizer que as ontologias visam capturar o conhecimento consensual de forma genérica, e que elas podem ser reutilizadas e compartilhadas entre aplicativos de software e por grupos de pessoas [47].

As ontologias podem ser modeladas com diferentes técnicas de modelagem de conhecimentos e podem ser implementadas em vários tipos de linguagens. Ontologias podem ser altamente informal se forem expressas em linguagem natural; semi-informal, se expressa em uma forma restrita e estruturada da linguagem natural; semi-formal se expressa em uma linguagem artificial e formalmente definido, por exemplo a OWL, e rigorosamente formal, se fornecer meticulosamente termos definidos com semântica formal, teoremas e provas de propriedades, tais como robustez e completude [47].

Neste trabalho, não temos interesse em buscar uma representação que não seja possível de ser processada por computador, tão pouco, uma representação que exija tempo de provas de conceitos. Logo, serão discutidos apenas as ontologias que estejam na faixa do semi-informal e semi-formal.

2.2.1 Classificação de Ontologias

Existem diversas classificações sobre Ontologias, quanto a sua função, ao grau de formalismo, à aplicação, à estrutura ou ainda quanto ao tipo de conteúdo que a mesma representa [2, 47].

van Heijst et al. [85] classificam as ontologias em duas dimensões:

- A primeira dimensão trata da estrutura da conceituação, podendo ser:
 - ontologia terminológica, como um dicionário que define os termos a serem usados para representar conhecimento;
 - ontologia de informação, que especifica a estrutura de registros de banco de dados; e
 - ontologia de modelagem de conhecimento, que especifica a conceituação do conhecimento.
- A segunda dimensão trata da natureza da conceituação e classifica as ontologias como:
 - ontologias de representação, as quais explicam as conceituações que fundamentam formalismos de representação de conhecimento;
 - ontologias de domínio, que expressam conceituações que são específicas para um domínio particular;
 - ontologias genéricas, que são similares às de domínio, porém seus conceitos são genéricos para vários domínios; e
 - ontologias de aplicação, que contêm as definições que são necessárias para modelar o conhecimento requerido por uma particular aplicação.

Em 1998, Guarino [42] sugeriu uma classificação quanto à generalidade da ontologia, conforme Figura 2.1, propondo uma classificação dos tipos de ontologias baseado em níveis de dependência:

- Ontologias de alto nível, descrevem conceitos mais gerais como espaço, tempo, matéria, objeto, evento, ação, etc, que são independentes de um problema particular ou domínio;
- Ontologias de domínio, descreve o vocabulário relacionado a um domínio genérico (como medicina, química, engenharia ou automóveis);
- Ontologias de tarefas, descreve uma tarefa genérica ou atividade (como diagnosticar ou vender);
- Ontologia de aplicação, descrevem conceitos que dependem tanto de um determinado domínio quanto tarefa, e muitas vezes combinam especializações de vários domínios e tarefas. Estes conceitos muitas vezes correspondem a papéis desempenhados por entidades do domínio enquanto executam uma determinada tarefa.

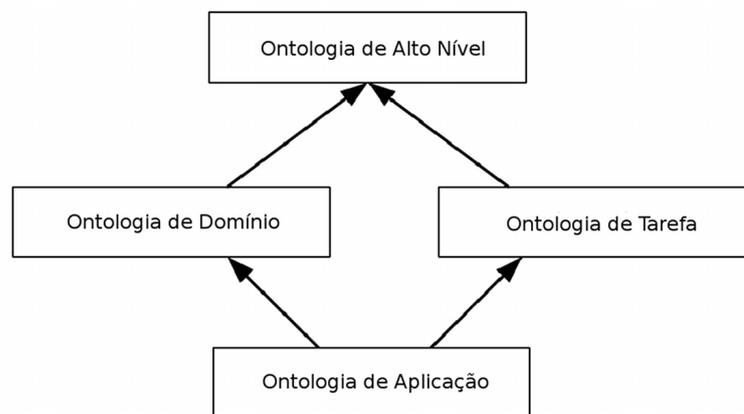


Figura 2.1: Classificação de Ontologias

Em especial, ontologias no nível mais geral de abstração, ou ontologias de fundamentação, tem a preocupação com as categorias que se aplicam às diversas áreas de conhecimento. Essas ontologias são sistemas de categorias filosoficamente bem fundamentados e independentes de domínio. Este tipo de ontologia serve como base para o estabelecimento de consenso e negociação entre humanos [45].

2.2.2 Uso da Ontologia em Sistemas de Informação

As ontologias devem ser construídas sobre um propósito ou aplicação em um determinado domínio. Guarino [42], faz uma discussão quanto ao uso de ontologias em três diferentes tipos de componentes de um SI:

- programas aplicativos,
- recursos de informação, como bancos de dados e / ou bases de conhecimento, e
- interfaces de usuário.

Guarino [42] ainda discute o impacto que uma ontologia pode ter sobre um SI, e distingue duas dimensões ortogonais: uma dimensão temporal e uma dimensão mais estrutural, relativa à forma que uma ontologia pode afetar os componentes de um SI.

Quanto à dimensão temporal, o autor se refere ao uso de ontologias em tempo de desenvolvimento *versus* em tempo de execução:

- Em tempo de desenvolvimento, dois cenários diferentes são apresentados: (i) no primeiro cenário, temos um conjunto de ontologias reutilizáveis à nossa disposição, organizados em uma biblioteca de ontologia contendo ontologias de domínio e tarefa e (ii) no segundo cenário, o grau de reutilização é muito limitado, pois só temos uma ontologia muito genérica, consistindo de distinções entre as entidades básicas do mundo e sobre os tipos de classes e tipos de relações;
- Em tempo de execução, devemos distinguir aqui um SI ontológico-consciente (*ontology-aware*) de um ontológico-guiado (*ontology-driven*): (i) no primeiro caso, um componente do SI apenas tem conhecimento da existência de uma ontologia (possivelmente de acesso remoto) e pode usá-la a qualquer propósito da aplicação quando necessário e (ii) no segundo caso, a ontologia é mais um componente (tipicamente local), colaborando em tempo de execução no objetivo global do SI.

Quanto à dimensão estrutural, o impacto das ontologias nos componentes do SI, onde cada um dos componentes de um SI - programas aplicativos, bancos de dados e interfaces de usuário - pode usar uma ontologia de sua própria maneira:

- Como um componente de banco de dados, o mais evidente é usá-la em conexões com componentes de base de dados. De fato, uma ontologia pode ser comparada a um esquema de um banco de dados. Em tempo de desenvolvimento, uma ontologia pode desempenhar um papel importante na análise de requisitos e fase de modelagem conceitual. Um exemplo de uso de ontologias em tempo de desenvolvimento é a integração de informações, usando um esquema conceitual comum, por exemplo em uma aplicação de *data warehousing* construído pelo mapeamento de esquemas conceituais heterogêneos de uma ontologia de mais alto nível de mesmo domínio. Em tempo de execução, existem muitas maneiras em que ontologias e bases de dados podem cooperar. A disponibilidade de ontologias explícitas para recursos de informação está, por exemplo, no cerne da abordagem de integração de informações baseada em mediação (*mediation-based*);
- Como componente de uma interface do usuário, talvez não tão óbvio, mas, no entanto, muito importante, é sua utilização como componentes de interface do usuário. A ontologia atribui informação semântica sobre as restrições das classes e relacionamentos usados para modelar um determinado domínio e tarefa, e a partir dessa informação gera interfaces baseadas em formulários que verificam a violação das restrições;
- Como componentes de programas de aplicativos de um SI, em tempo de desenvolvimento, pode, em princípio, gerar a parte estática de um programa com a ajuda de uma ontologia, ou ainda, usada para apoiar o desenvolvimento de software orientado a objetos. Em tempo de execução, pode representar explicitamente todo o conhecimento de domínio implicitamente codificado no programa de aplicação, transformando o programa em um sistema baseado em conhecimento. Como é sabido, esta abordagem tem grande vantagem do ponto de vista da facilidade de manutenção, extensibilidade, e flexibilidade.

2.2.3 Problemas no Uso de Ontologias

Segundo O’Leary (1997) apud Guizzardi [43], alguns problemas no uso de ontologias podem ser destacados:

- A escolha de uma ontologia é um processo político, já que nenhuma ontologia pode ser totalmente adequada a todos os indivíduos ou grupos;
- Ontologias não são necessariamente estacionárias, no entanto, poucos trabalhos têm dado ênfase na evolução de ontologias;
- Estender ontologias não é um processo direto, uma vez que ontologias são geralmente estruturadas de maneira precisa e, como resultado são particularmente vulneráveis a questões de extensão, dado o forte relacionamento entre complexidade e precisão das definições;
- A noção de bibliotecas de ontologias sugere uma relativa independência entre diferentes ontologias. A interface entre elas constitui, portanto, um impedimento, especialmente porque cada uma delas é desenvolvida no contexto de um processo político distintos, ontologias desenvolvidas independentemente podem não se integrar efetivamente com outras por vários motivos, desde similaridade de vocabulário até visões conflitantes do mundo.

Todas estas questões são importantes no momento de se construir ou reutilizar uma ontologia, em especial às questões ligadas a localização dos conceitos de uma região para outra (países ou estados, por exemplo), dado que o contexto político pode afetar o uso adequado das ontologias.

2.2.4 Outras Forma de Representar Conhecimento

Neste momento, destacar que existem outras formas de representar conhecimento além da ontologia, é importante para ter claro que as próprias técnicas de representação de conhecimento evoluem no tempo e na forma para ampliar a representação de conhecimento de um certo domínio.

Representação de Conhecimento

A representação do conhecimento pode ser classificada, segundo Brachman (1979) apud [19], de acordo com os tipos de primitivas oferecidas ao usuário, em quatro níveis: lógico, epistemológico, ontológico e conceitual.

O nível lógico é o nível da formalização, entretanto, sem a preocupação com a semântica em termos dos conceitos e de suas relações. Em nível lógico, todo o foco está centrado em uma “sintaxe” que possibilite uma verdadeira ação do pensar. No nível epistemológico, a noção genérica de um conceito é introduzida como uma primitiva de estruturação de conhecimento; ele é o nível da estruturação. O nível ontológico tem por objetivo restringir o número de possibilidades de interpretação do conceito dentro de um dado contexto, a partir de um formalismo que pretende representar o conteúdo do conceito. No nível conceitual, independentemente de um formalismo, os conceitos possuem, a priori, uma interpretação definida. O esqueleto da estrutura dos conceitos em um determinado domínio está dado, e o conhecimento é expresso na forma de uma especificação deste esqueleto.

Na perspectiva das Ciências da Informação e Computação e da Terminologia, interessamos aqueles níveis em que uma estrutura de conhecimento possa ser sistematizada e representada a partir de contextos específicos. Segundo Newel (1982) apud [19], os níveis epistemológico e ontológico são aqueles que permitem a representação de conhecimento estruturado e formalizado.

No nível epistemológico, como vimos, especificam-se a estrutura dos conceitos e seus inter-relacionamentos. No nível ontológico, avança-se um pouco mais no processo de organização e classificação de um determinado domínio, e acrescenta-se a definição dos conceitos que nele estão inseridos. Enquanto o nível epistemológico é o nível de estruturação, o nível ontológico é o nível de significação.

Na Ciência da Computação, modelos de objetos e de dados são exemplos de representação no nível epistemológico. Contudo, para representar conhecimento, tais representações são ainda pobres (ou ontologicamente leves [42]). Esses formalismos estabelecem apenas significados particulares de estruturação. A ontologia formal é um formalismo classificado no nível ontológico, pois sistematiza conhecimento visando a formalização de definições axiomáticas.

Na Ciência da Informação, a teoria da classificação³ pode ser considerada em um nível de transição entre o nível epistemológico e ontológico: apesar de não pretender chegar à definição dos conceitos de um dado domínio, ela possui um formalismo que possibilita a representação do conhecimento. Já as teorias do conceito⁴ e da terminologia⁵ podem ser classificadas como de um nível ontológico propriamente dito, pois permitem a sistematização de conhecimentos e possuem diretrizes para a elaboração de definições.

Instrumentos de Representação de Conhecimento além da Ontologia

Para ter mais clara a relação das ontologias com os outros instrumentos e modelos de representação de conhecimento, cabe evidenciar dois pontos sobre as definições e classificações de ontologias. O primeiro ponto é a definição feita por Guarino [42] sobre o grau de expressividade, no nível ontológico, dado pela quantidade de axiomas que relacionam os conceitos de um determinado domínio. O segundo ponto, é a classificação, feita por Lassila e McGuinness apud Gómez-Pérez et al. [47], ver Figura 2.2, sobre as técnicas de representação de conhecimento na fronteira do ontológico e não-ontológico. Eles classificam as representações ontológicas de acordo com as informações que o instrumento precisa expressar e a riqueza de sua estrutura interna, observando as seguintes categorias:

³A Teoria da Classificação Facetada (Ranganathan, 1951) é representante de um modelo que se utiliza do método dedutivo para classificar o conhecimento dentro de um contexto. Desta forma, possui mecanismos de representação para trabalhar com metaníveis conceituais, como, por exemplo, as categorias. É a partir delas que os conceitos são ordenados para formar classes de conceitos [26].

⁴Dahlberg [25] define conceito, na Teoria do Conceito, como sendo uma unidade de conhecimento, formada por três elementos: o item de referência ou referente (aquilo que se pretende conceituar), as características (declarações ou predicados verdadeiros sobre o referente) e a forma verbal, que sintetiza as características utilizando um termo ou nome.

⁵Criado pelo austríaco Eugene Wuster na década de 1950, a Teoria Geral da Terminologia (TGT), marco da terminologia moderna, é definida como um instrumento para desambiguação da comunicação científica e técnica, e classifica as relações entre os conceitos em relações lógicas e ontológicas. As relações lógicas são as relações definidas entre conceitos, conhecidas também como relação de semelhança, de similaridade, de abstração ou genérica. As relações ontológicas definem as abstrações das relações estabelecidas entre indivíduos, ou seja, relação ôntica [28].

vocabulários controlados, glossários, dicionários de sinônimos, hierarquias (*is-a*) informais, hierarquias (*is-a*) formais, instâncias formais, quadros (*frames*), restrição de valor, e restrições lógicas gerais.

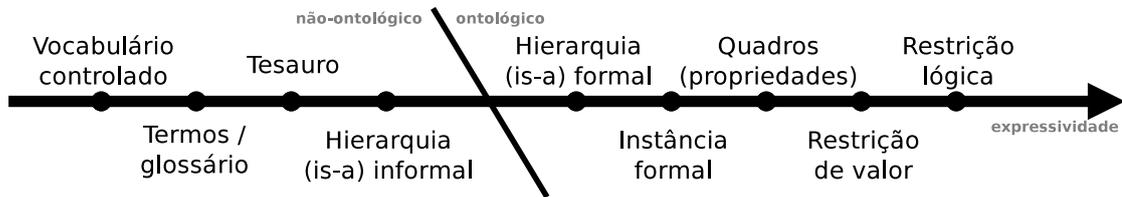


Figura 2.2: Categorização da Representação de Conhecimento

- Vocabulários Controlados (*controlled vocabularies*), são listas finitas de termos;
- Glossários (*glossaries*), são listas de termos com os seus significados associados especificados como instruções de linguagem naturais;
- Tesouros, (*thesauri*), fornecem alguns adicionais semânticos entre os termos. Eles dão informações, tais como relações de sinônimos, mas não fornecem uma hierarquia explícita;
- Hierarquias (*is-a*) Informais (*informal is-a hierarchies*), são tomadas a partir de especificações de termos de hierarquias, no entanto, não estão em uma subclasse estrita ou uma hierarquia *is-a*;
- Hierarquias (*is-a*) Formais (*formal is-a hierarchies*), se B é uma subclasse de A e um objeto é uma instância de B, então o objeto é uma instância de A. Hierarquias de subclasses estritas são necessárias para explorar a propriedade de herança.
- Instâncias Formais (*formal instance*), Hierarquias (*is-a*) Formais que incluem instâncias do domínio;
- Quadros (*frames*), este tipo de representação incluem as classes e as suas propriedades, que podem ser herdadas por classes dos níveis mais baixos da taxonomia formal *is-a*;
- Restrição de Valor (*value restriction*), são as representações que expressam restrição de valor, ou seja, que podem colocar restrições sobre os valores que podem preencher em uma propriedade;
- Restrições Lógicas Gerais (*general logical constraints*), são as representações que expressam restrições lógicas gerais e são as mais expressivas. Podem ser especificadas por restrições de lógica de primeira ordem entre os termos que usam linguagens expressivas de ontologia.

Tendo clara a fronteira entre representação de conhecimento ontológica e não ontológica, e a noção de que se pode aumentar a expressividade de um modelo em relação a outro, em especial pela quantidade de axiomas entre os conceitos, fica evidente que é possível ampliar a expressividade de uma instância de um modelo explicitando as relações entre os termo da instância. Alguns trabalhos na área da saúde, [20, 55, 82], vem mostrado esse processo.

2.2.5 Meta-Ontologias

Conforme Fensel [37], as meta-ontologias, ou seja, as linguagens que definem as ontologias, são baseadas em três paradigmas: (i) o primeiro segue a linha das lógicas de primeira ordem, e.g., CycL e o KIF; (ii) o segundo paradigma, e.g., a Ontolingua e o F-logic, é baseada na abordagem de quadros (*frame-based approach*), modelando as primitivas em lógica de primeira ordem; (iii) o terceiro é baseado no paradigma da lógica descritiva, que descrevem o conhecimento em termos de conceitos e restrições de papéis usadas para derivar automaticamente taxonomias de classificação.

2.2.6 Web Semântica

O projeto da Web Semântica, liderado pela *World Wide Web Consortium* (W3C), tem o objetivo de criar e implantar padrões tecnológicos para permitir, no panorama da Internet como fomentador da troca de dados entre programas e dispositivos, não somente a troca de informação entre pessoas, mas principalmente para estabelecer uma forma de compartilhamento mais significativo de dados entre dispositivos e sistemas de informação de uma maneira geral [80].

O consórcio internacional W3C é a organização responsável pela padronização da Web Semântica, sendo formada por diversas organizações internacionais. De acordo com o W3C, a Web só pode atingir seu pleno potencial se ela se tornar um local onde os dados possam ser compartilhados, processados e entendidos por ferramentas automatizadas da mesma forma que o são pelas pessoas [12].

Na perspectiva da Web Semântica os programas devem ser capazes de compartilhar, processar e entender dados mesmo quando os mesmos são projetados independentemente uns dos outros. O W3C, no ano de 2001, criou um grupo de trabalho chamado Web-Ontology (WebOnt) com o objetivo de especificar uma linguagem de marcação ontológica para a Web Semântica. Esta linguagem recebeu o nome de *Web Ontology Language* (OWL).

Linguagem OWL

A OWL é uma linguagem de marcação semântica para publicar e compartilhar ontologias na Web. Ela foi desenvolvida como uma extensão do RDF e tem como base a linguagem DAML+OIL. Segundo Fensel [37] a linguagem *Ontology Interchange Language* (OIL), unifica três aspectos importantes fornecidos por diferentes comunidades, ver Figura 2.3:

- a semântica formal e suporte a um raciocínio eficiente, como previsto pelo paradigma da lógica descritiva;
- primitivas da modelagem epistemológicas, como previsto pelo paradigma de *Frame*;
- uma proposta de padrão para troca de notações sintáticas, conforme previsto pela comunidade web.

Ao se fundirem à comunidade do *DARPA Agent Markup Language* (DAML), ampliam sua capacidade com as vantagens das linguagens de marcação e outras características importantes para a Web Semântica, além de promover uma linguagem mais limpa [64].

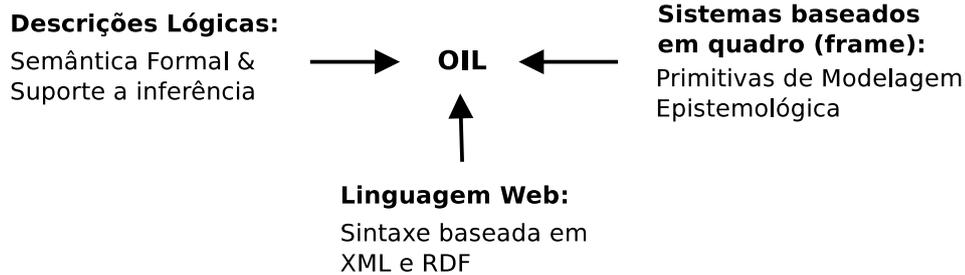


Figura 2.3: Os três aspectos da linguagem OIL.

A OWL possibilita o raciocínio sobre os dados, enquanto o RDF promove a associação e integração desses dados de forma distribuída. A OWL também expande o RDFS, permitindo outros tipos de relacionamento entre classes e ainda estabelece nas subclasses restrições de propriedades herdadas das superclasses, podendo ter tais restrições relações com outras classes [84].

A Web Semântica amplia a forma de representar informação aumentando a capacidade das linguagens na representação de informação e conhecimento. A Figura 2.4 mostra o crescimento da capacidade de representar conhecimento dentro da Web e também a relação entre as linguagens OWL, RDFS e RDF na ampliação do que era apenas representação sintática dos dados [84].

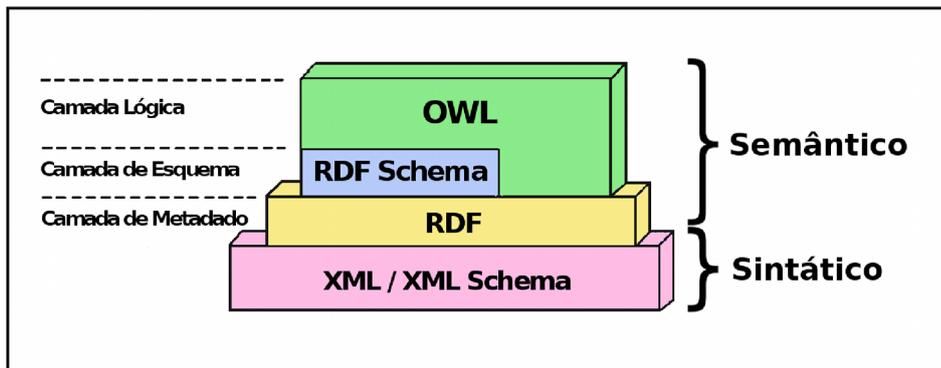


Figura 2.4: Relação entre as linguagens OWL, RDFS e RDF.

A linguagem OWL possibilita um grande aprofundamento da representação semântica dos dados, podendo em alguns casos ser inviável o total relacionamento semântico de alguns domínios de informação. Por esse fato o W3C se preocupou em encapsular a linguagem OWL em três níveis de profundidade chamados: OWL Lite, OWL DL e OWL Full. O encapsulamento permite que as relações semânticas sejam ampliadas conforme a necessidade da aplicação, simplificando o processo de construção da ontologia de um domínio [84].

A OWL Lite é a sub-linguagem mais simples sintaticamente, sendo indicada para usuários que precisam de uma hierarquia de classificação e restrições simples. A OWL DL (*Description Logics*) proporciona expressividade máxima e garantia de completude computacional e decidibilidade. Esta sub-linguagem inclui todos os construtores, porém com algumas restrições. A OWL Full provê expressividade máxima, além da sintaxe ser livre do RDF. Esta sub-linguagem não provê garantia computacional [84].

Capítulo 3

Semântica na Saúde

Este capítulo, amplia a discussão dos aspectos semânticos, observando e restringindo o problema da interoperabilidade semântico no domínio da saúde. Na Seção 3.1, é apresentada a abordagem da modelagem em dois níveis proposta no domínio da saúde para fazer a modelagem conceitual dos SIS. Na Seção 3.2, apresentamos algumas formas de representação de conhecimento no domínio da saúde. Por fim, na Seção 3.3, é apresentado o SUS e os elementos que compõem o SUS, com um destaque especial à Vigilância Alimentar e Nutricional (VAN), por este ser o objeto do estudo de caso deste trabalho.

3.1 Modelagem em dois níveis

Em 1988, a União Europeia estabeleceu a iniciativa *Advanced Informatics in Medicine* (AIM) como parte de seu programa para desenvolvimento de pesquisa e tecnologia. Partindo dessa iniciativa, em 1992, foi criado o projeto de pesquisa *Good European Health Record* (GEHR), posteriormente chamado de *Good Electronic Health Record*, formado por 21 instituições em oito países europeus, e mais tarde continuou com forte participação de grupos de pesquisa da Austrália[86]. Em 1998, dois dos participantes do projeto GEHR original, Sam Heard e Thomas Beale, trabalhando no desenvolvimento de sistema de Registro Eletrônico de Saúde (RES), ver Seção 3.5, na Austrália propuseram o conceito de modelagem de dois níveis, e o uso de arquétipos para separar informação e conhecimento[33].

Desenvolveu-se assim a ideia de uma arquitetura fundamentada na elaboração de dois modelos, em dois níveis diferentes: um modelo de referência (ou informação) e um modelo de conhecimento, onde Beale [11] define a modelagem de dois níveis também conhecida como modelo dual [74], em contraponto à modelagem de desenvolvimento de um nível ou modelo simples.

O modelo de desenvolvimento de um nível, Figura 3.1 [11], é o modelo clássico de desenvolvimento. Nesta abordagem um desenvolvedor por meio de técnicas, como modelagem UML em sistemas orientado a objeto, busca representar um modelo de um determinado domínio. Posteriormente este modelo é implementado em um sistema de software. Para situações onde o número, complexidade e a taxa de mudança de definição de requisitos são pequenos, ou seja, baixa variabilidade espacial e temporal, a abordagem de um único nível pode ser muito mais econômica. No entanto, em grandes e ricos domínios de informações

sujeitos a constantes mudanças, como a atenção à saúde, alta tecnologia de produção e defesa, usar a abordagem clássica pode resultar em consequências negativas, incluindo:

- são usados apenas os requisitos encontrados durante o desenvolvimento atual;
- modelos que tenham conceitos genéricos e de domínio são problemáticos na mesma hierarquia de herança;
- dificuldade de completar os modelos de forma satisfatória;
- problema de adequação semântica;
- dificuldade de manutenção dos requisitos;
- introdução de novos conceitos requer mudanças de software e banco de dados e, normalmente, reconstrução, testes e reimplementação;
- dificuldade em se alcançar interoperabilidade;
- dificuldade em se alcançar padronização no domínio.

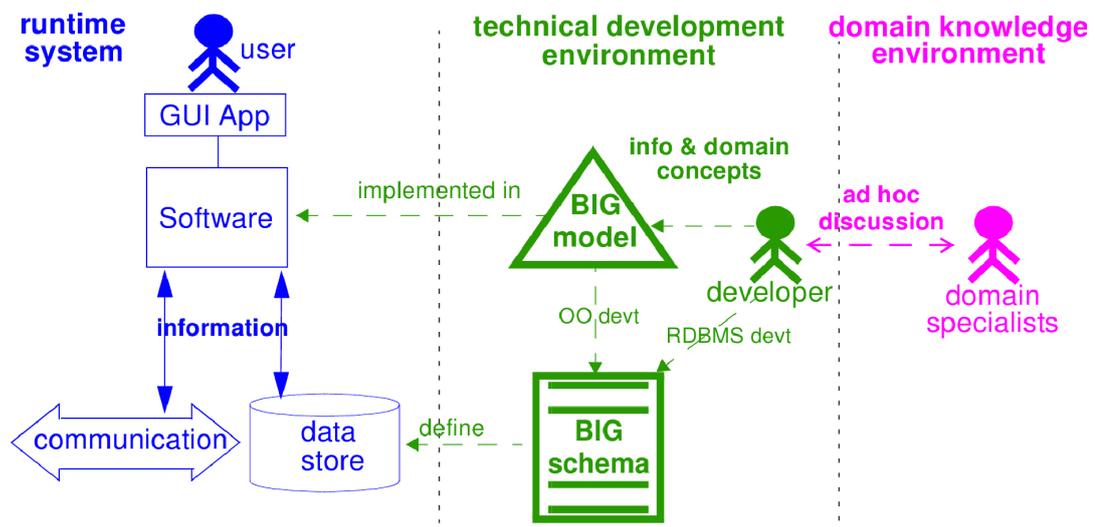


Figura 3.1: Modelo Simples

Segundo Beale [11], uma abordagem alternativa ao modelo simples é o modelo que separa a semântica da informação e conhecimento em dois níveis do modelo. No primeiro nível, localiza-se um modelo de referência, elaborado a partir da identificação de um pequeno conjunto de classes genéricas, suficientes para representar os conceitos relativos ao RES, isto reflete em características estáveis [54]. Nesse modelo, evita-se usar atributos com significados explícitos, deixando as particularidades de conceitos específicos serem representadas no modelo de conhecimento.

No segundo nível, é definido o modelo de conhecimento que especifica, e efetivamente restringe, em uma hierarquia especial de registro as regras de negócios que não são especificadas no modelo de referência. Essas estruturas de dados, desenvolvidas para representar modelos clínicos, são chamadas de arquétipos [11]. Este nível objetiva representar o conhecimento variante, em geral, representam políticas de uso ou **aplicações** do conhecimento invariante.

Desto forma a modelagem de dois níveis, Figura 3.2 [11], permite uma clara divisão de responsabilidades, enquanto a equipe de especialistas em TIC se preocupa com o desenvolvimento dos *frameworks* para tratamento das classes e do modelo de conhecimento, a equipe de especialistas em saúde se concentra com a representação do conhecimento clínico necessário.

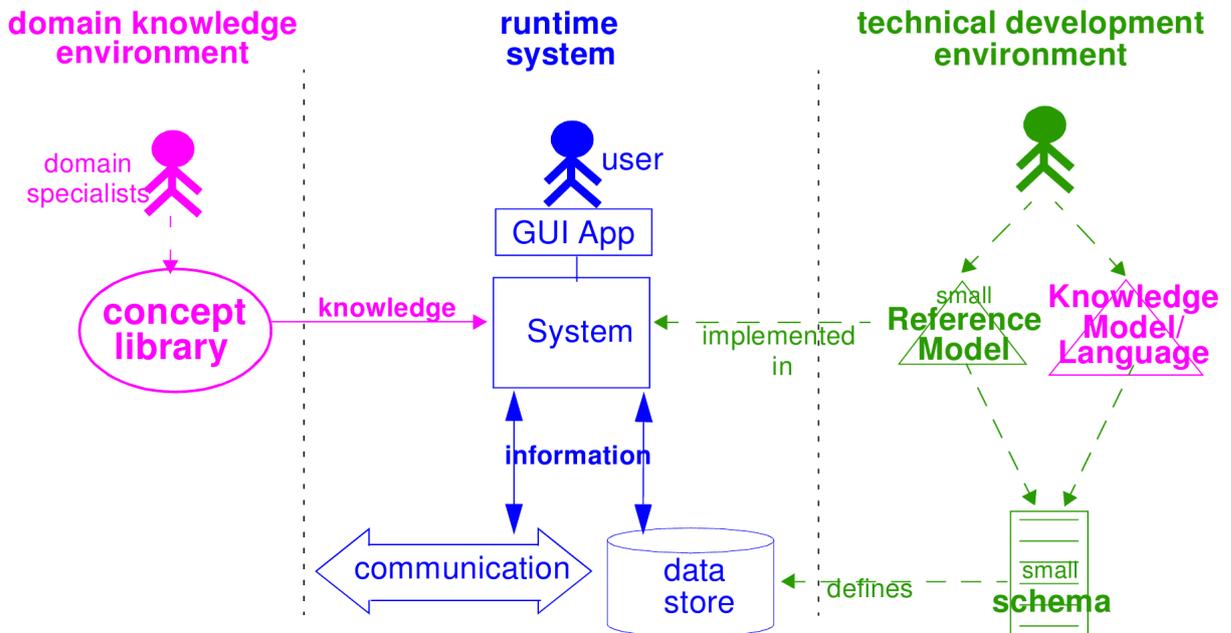


Figura 3.2: Modelo Dual

De acordo com Beale [11], são benefícios da modelagem de dois níveis:

1. software e base de dados dependem somente de modelos de classes pequenos e estáveis, podendo ser desenvolvidos rapidamente, sem aguardar a definição de conceitos, e raramente precisam ser alterados;
2. os modelos técnicos são desenvolvidos por especialistas em TIC, enquanto definições de conceitos são desenvolvidas por especialistas em saúde. Os dois processos são independentes e complementares;
3. os modelos de informação e de conhecimento são compartilhados e, assim, os sistemas podem interoperar no nível de conceitos;
4. os objetos do modelo de conhecimento (arquetipos) auxiliam e facilitam a recuperação de informação nos SIS.

3.1.1 Modelos de Referência

Segundo Santos [71], um modelo de referência constitui-se de:

- a) um conjunto de tipos primitivos;
- b) um conjunto de classes que definem a organização da informação;

- c) um conjunto de classes que descrevem informações contextuais; e
- d) classes que descrevem dados demográficos.

As classes são definidas da forma mais genérica possível, para que o desenvolvimento, a manutenção e a expansão do software impactem menos esse modelo, que deve ser: pequeno, genérico e estável [71], ou seja, conhecimento invariante, e o mais consensual possível, dentro de um domínio. Em geral, um modelo de referência deve ser baseado em **terminologias** e **ontologias** de mais alto nível no domínio, buscando conceitos, formal ou informalmente consolidados.

Ao se optar por trabalhar com o modelagem de dois níveis, inicialmente, é necessário escolher (ou definir) o modelo de referência a ser utilizado. Segundo Schloeffel et al. [72], existem três principais modelos de referência internacionais:

1. openEHR,
2. CEN/ISO 13606.
3. HL7 CDA/RIM,

Cada um destes modelos apresentam características particulares e importantes para uma modelagem mais abstrata dos conceitos relacionados a SIS e portanto serão um pouco mais detalhados.

Padrão openEHR

Segundo Beale et al. [9], o padrão openEHR propõe um modelo de referência genérico que permite representar toda a informação gerada no âmbito da atenção à saúde. O núcleo clínico do modelo de referência é a classe de entrada (*ENTRY*). Observando as classes do padrão openEHR, Figura 3.3 (adaptado de [6]), podemos observar duas classes: *CARE_ENTRY* que corresponde às informações do Cuidado à Saúde e *ADMIN_ENTRY* que corresponde às informações Administrativas.

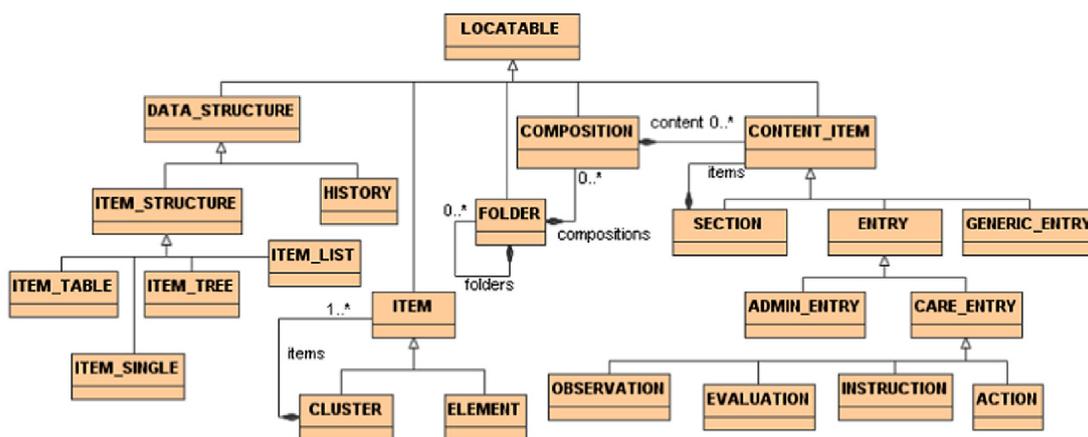


Figura 3.3: Classes do padrão openEHR

As informação de cuidado dizem respeito ao registro clínico de fato. Segundo Beale et al. [9], existem quatro classes genéricas de registro de informações do cuidado à saúde:

- Observação (*OBSERVATION*): para registrar tudo o que puder ser observado, medido ou respondido pelo paciente;
- Avaliação (*EVALUATION*): avaliação, diagnósticos e planos de cuidado;
- Instrução (*INSTRUCTION*): afirmações que podem ser executadas, como receitas de medicamentos, orientações, exames, encaminhamento, entre outros;
- Ações (*ACTION*): informação que se registram como resultado da execução de instruções.

O processo de investigação clínica que corresponde ao uso da tipologia da informação de registros clínicos, apresentado na Figura 3.4 [7], são as classes de entrada de informação do cuidado (*CARE_ENTRY*) e surgem da análise do processo de resolução de problemas clínicos, ou seja, o médico registra as observações (1) sobre o paciente, registra as avaliações (2) realizadas com base nas observações, no seu próprio conhecimento/especialidade e em conhecimento externo como protocolos, cadernos de atenção à saúde ou linhas guias, ou consulta a outros especialistas. Com base nas avaliações, realiza instruções (3) como procedimentos clínicos, prescrição de medicamentos, tratamento terapêuticos, entre outros. E por último, ao realizar as instruções, os profissionais de saúde registram as ações ou serviços realizados (4), retomando o ciclo para saber se o problema foi resolvido [7].



Figura 3.4: Processo de Investigação clínica

Norma CEN/ISO 13606

A norma CEN/ISO 13606 foi desenvolvida pelo Comitê Técnico da ISO de Informática em Saúde (TC 215) a partir de experiência prática obtida durante a implementação do pré-padrão europeu ENV 13606 [54]. Propõe um modelo de referência simplificado e inspirado no modelo de referência proposto pelo padrão openEHR [72]. Seu modelo de referência representa as características globais de componentes de registro eletrônico de saúde e define várias classes genéricas, ver Figura 3.5, podendo ser destacadas [54]:

- COMPOSITION - seção de cuidado clínico, encontro ou documento;
- SECTIONS - títulos clínicos refletindo o fluxo de trabalho ou processo de consulta;
- ENTRY - declarações clínicas sobre “observações”, “avaliações” etc.;
- CLUSTER - estruturas de dados complexos compostos de múltiplas partes;
- ELEMENT - último nível de estrutura onde se encontram os valores de dados;
- FOLDER - organização em pastas das informações de um registro;

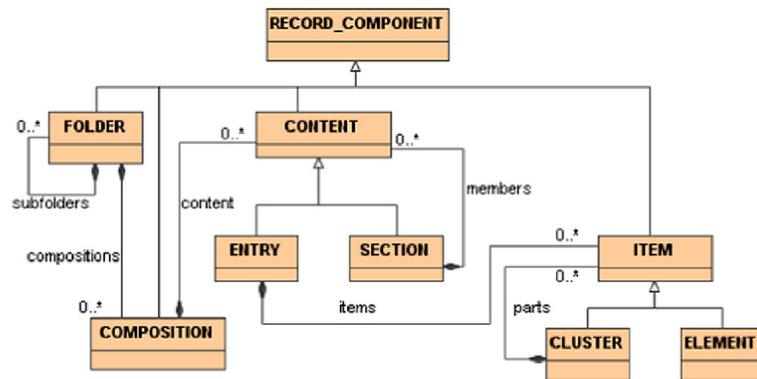


Figura 3.5: Classes da Norma CEN/ISO 13606

Padrão HL7

O padrão ANSI HL7 (*Health Level Seven*), atualmente em sua versão 3 (HL7v3), é um padrão definido na sétima camada, *Open Systems Interconnection* (OSI) padrão ISO, de aplicação, suporta funções como checagem de segurança, identificação de usuários, checagem de disponibilidades, mecanismos de negociação de trocas, e mais importante, estrutura de intercambio de informações. O padrão HL7 define um padrão de troca de mensagens utilizando modelos mais complexos por meio de um Modelo de Referência de Informação (*Reference Information Model - RIM*) chamado de HL7 RIM. Utilizando XML como linguagem de codificação o HL7 RIM é composto por 6 classes: *Act*, *Participation*, *Entity*, *Role*, *ActRelationship*, *RoleLink*, buscando garantir a interoperabilidade semântica que não existia nas versões anteriores do HL7 [32].

Efetivamente, o padrão HL7 utiliza uma estrutura para troca de documentos clínicos, *Clinical Document Architecture* (CDA), que define uma arquitetura em XML, derivada do modelo RIM, para troca de documentos clínicos conhecido como HL7 CDA. Um documento CDA pode ser lido tanto por máquinas quanto por humanos, é constituído por duas partes principais: cabeçalho e corpo. O cabeçalho identifica e classifica o documento e fornece informações sobre a autenticação, o registro de saúde, o paciente, e os fornecedores envolvidos. O corpo contém o relatório clínico e pode ser tanto um *Binary Large Object* (BLOB) não estruturado ou um elemento composto de marcação estruturada [32].

3.1.2 Modelo de Conhecimento

O modelo de conhecimento, representado por arquétipos¹, é a parte que define os conceitos usados na clínica, podendo estes ser mais próximos da aplicação, como um Sumário Clínico de uma Gestante, ou um conceito de mais alto nível como uma observação de pressão arterial ou uma observação do peso do paciente. Os arquétipos habilitam os especialistas de saúde a definirem formalmente o conteúdo clínico; possibilitam maiores chances de conservação do significado dos dados, por meio da especificação bem estruturada e explícita do conteúdo clínico; e simplificam o uso de terminologias clínicas [11].

Os arquétipos podem ser compartilhados, uma vez que são definições representadas de forma padronizada, e usados entre comunidades que compartilham registros de pacientes para definir como dados clínicos locais deveriam ser mapeados consistentemente.

Instrumentos que enriquecem a flexibilidade dos arquétipos são os *templates*, utilizados para construir estruturas que capturam os dados correspondentes a uma tarefa clínica particular e podem combinar dados de diferentes arquétipos ou de outros *templates* [11]. Os arquétipos normalmente são serializados por meio de uma linguagem chamada *Archetype Definition Language* (ADL) [54].

Os modelos de conhecimento formalizados em arquétipos são evidenciados e podem ser consensual entre os especialistas das áreas. Visto as teorias de engenharia de ontologias, temos claro que quanto mais simples e de mais alto nível os conceitos, maior a possibilidade destes obterem consenso e de serem reutilizados para modelar outros conceitos.

Segundo Beale [11], em domínios mais complexos, como o da saúde, é possível identificar níveis de abstração, que podem ser expressos em níveis ontológicos. Na atenção à saúde vários níveis ontológicos são reconhecidos:

- Princípios (*nível 0*): uma ontologia da linguagem e dos princípios da atenção à saúde. Este nível contém o conhecimento dos processos e entidades que constituem os fatos geralmente aceitos do domínio da saúde, questões que são verdadeiras sobre todas as instâncias de uma entidade ou processos.
- Descritivo (*nível 1*): conceitos que são expressos como composições estruturais de elementos do *nível 0*, cada um expresso em uma descrição bem coesa de uma observação, análise de algo no mundo real.
- Organizador (*nível 2*): conceitos criados por profissionais de saúde na tentativa de organizar os itens descritivos. Sua forma é baseada na lógica da investigação e notificação utilizada no domínio, e eles agem como um instrumento de navegação para ambos os autores e leitores de informação. Geralmente a organização dos conceitos é definida de acordo com uma metodologia ou um protocolo, como por exemplo o “registro de saúde orientado ao problema” que dá origem a estrutura SOAP de registro clínico.

¹Segundo Medeiros [65], o termo arquétipo se refere ao primeiro modelo ou imagem de alguma coisa, antigas impressões sobre algo. É um conceito explorado em diversos campos de estudo, como a Filosofia, Psicologia e a Narratologia, em especial na Psicologia, o suíço Carl Gustav Jung, introduz o termo como a forma imaterial à qual os fenômenos psíquicos tendem a se moldar. Ainda anterior a Jung, Plotino na Filosofia, define o termo para designar as ideias como modelos de todas as coisas existentes.

- Temático (*nível 3*): coleções de informações de granulação mais grossa e de mais baixos níveis ontológicos. Inclui conceitos como “história familiar”, “lista de problemas”, “atendimento ao paciente”, “plano de cuidado” ou “gravidez”.

Observando os níveis de abstração e as classes de organização dos conceitos do padrão openEHR, Figura 3.3, podemos fazer uma associação das classes que corresponde a cada nível de abstração, indo do nível mais geral (0) para o nível mais específico (3), ou seja, mais próximo da aplicação:

- Nível 0: classes *ELEMENT* e *CLUSTER*;
- Nível 1: classes *ENTRY*;
- Nível 2: classes *SECTION*;
- Nível 3: classes *COMPOSITION*.

Archetype Definition Language

A linguagem utilizada para formalizar os modelos de conhecimento, ou arquétipos, é a linguagem de definição de arquétipos, ADL, que se baseia em modelos de restrições de entidades de domínio, ou “regras de negócio estruturadas”. Descreve restrições para dados que são instâncias de um modelo de referência e pode ser utilizada em qualquer domínio onde existem modelos de objetos formais que descrevem instâncias de dados [8].

Segundo Beale and Heard [8], a linguagem ADL utiliza três diferentes sintaxes: *Data ADL* (dADL), modo de definição de dados, *Constraint ADL* (cADL), modo de restrições, e a *First-Order Predicate Logic* (FOPL), versão de lógica de predicados de primeira ordem. A Figura 3.6 [8] apresenta a estrutura da linguagem ADL com as seções e suas respectivas sintaxes.

O dADL serve para expressar instâncias de dados baseados em um modelo de referência, e sua estrutura busca ser legível tanto por humanos quanto por computadores. Um documento dADL pode trabalhar com um ou mais objetos de um mesmo modelo de objetos e todos os seus identificadores, como nomes de tipos e nomes de atributos, são oriundos de um modelo de referência.

O cADL é uma linguagem que aplica restrições sobre os dados definidos por modelos de referência orientados a objetos. Podemos destacar sobre suas funcionalidades: *matches* que define uma relação de pertinência; *occurrences* que indica o número de instâncias possíveis de um atributo; *existence* que indica a opcionalidade de um atributo; *cardinality* que define se o atributo é um *container* e as palavras *ordered*, *unordered* e *unique* que definem o tipo de agrupamento lógico usado no *container*. Ainda podem ser incorporadas, ao cADL, expressões regulares para validação de um *data.type text*.

O FOPL é uma linguagem de lógica de predicados de primeira ordem que vem inserida em cláusulas nas seções *Declaration*, instruções declarativas, e *Invariant*, instruções assertivas. Em especial, se utiliza de operadores de comparação e igualdade.

A linguagem ADL tem um sintaxe própria, no entanto pode ser totalmente descrita utilizando a linguagem XML ou ainda OWL, aumentando a integração com as ferramentas da Web Semântica [57, 63, 71].

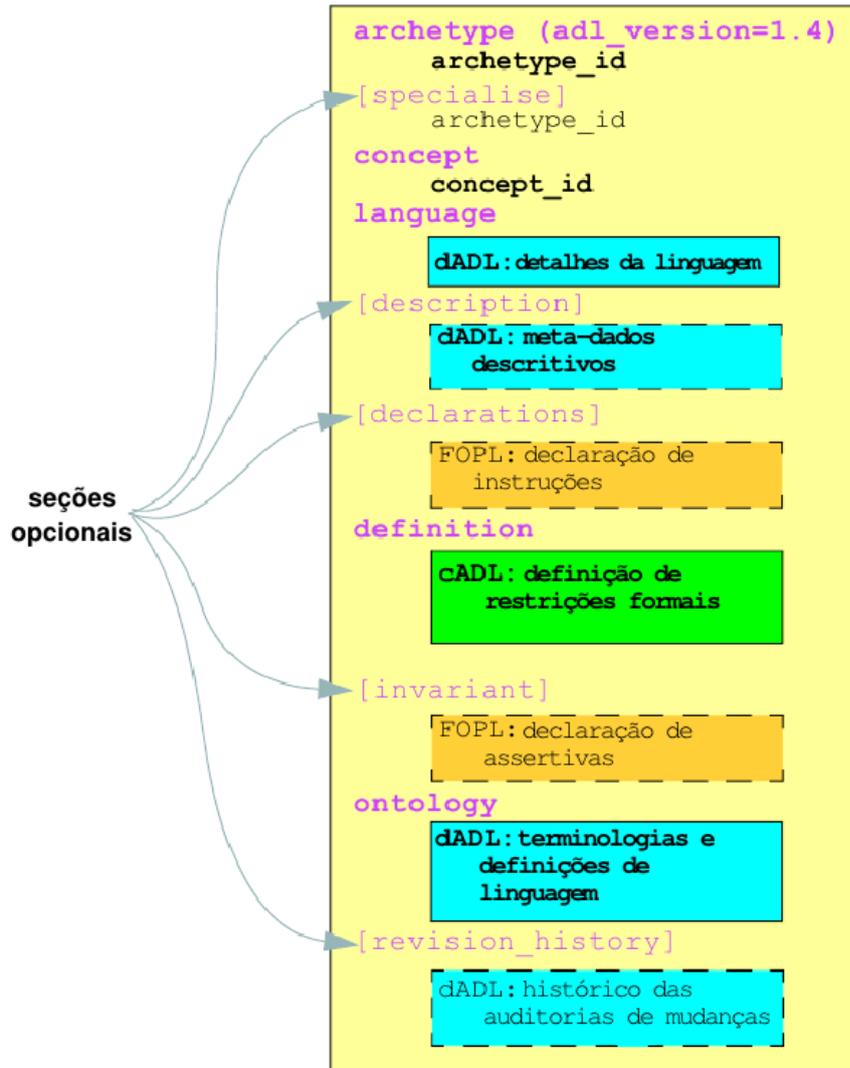


Figura 3.6: Estrutura da linguagem ADL

3.1.3 Relação entre os Modelos de Referência e de Conhecimento

Neste ponto do texto é importante ter clara a relação entre os Modelos de Referência e de Conhecimento, e como esses modelos são instanciados. A Figura 3.7 apresenta a relação entre um modelo de referência que representa a informação e o conhecimento representado através de arquétipos ou linguagens.

Na Figura 3.7 (adaptado de Beale [11]), podemos ver que as informações são instâncias de um modelo de referência, e os arquétipos são instâncias de um modelo de conhecimento. Este, por sua vez, é formalmente vinculado ao modelo de informação. As informações são restringidas em tempo de execução pelos arquétipos. Os conceitos do domínio de saúde são representados na forma de arquétipos.

Ainda ao avaliar a relação da “Informação” e do “Conhecimento”, exibido na Figura 3.7, temos no lado “informação” representada a informação real, ou o que muitas pessoas associam como “dados”. A estrutura e a semântica da informação são definidas por um

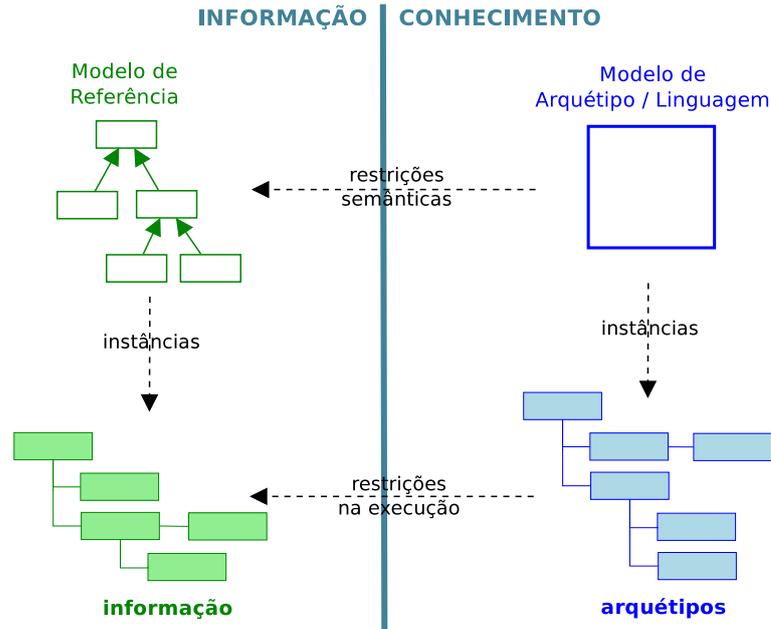


Figura 3.7: Relação entre modelo de referência e conhecimento

modelo de referência, os itens de informação são “instâncias” deste modelo. Sobre o lado do “conhecimento”, representado por arquétipos, isto é, o conhecimento em nível de definições, cuja função é restringir em tempo de execução a estrutura e a semântica da informação. No entanto, em contraste com a relação de classe / instância que restringe os “duros” aspectos de tipos de objetos, atributos de tipos e nomes, e assinaturas de função, os arquétipos restringem os aspectos “leves” de informação, ou seja, multiplicidades, valores de atributos e de nomeação. Em outras palavras, os modelos de arquétipos impõem as restrições semânticas aos modelos de referência das informações, partindo de um nível de conhecimento para tratamento da informação.

3.2 Representação do Conhecimento na Saúde

A representação de conhecimento nas áreas de Saúde e Medicina tem trabalhos apresentados desde os anos 90 [38, 52]. Para poder usar esses conceitos agrupados ou representados em modelos é preciso antes conhecer o conjunto de elementos disponíveis e saber para qual finalidade eles foram desenvolvidos. Portanto serão apresentados alguns modelos e ontologias no domínio da saúde.

3.2.1 Ontologia no Domínio da Saúde

As ontologias no domínio da saúde são desenvolvidas para resolver problemas como de interoperabilidade semântica na transmissão ou compartilhamento de dados clínicos ou de cuidado de um paciente, ou na necessidade de semântica baseados em critérios para fins estatísticos. A comunicação inequívoca de complexos e detalhados conceitos médicos é uma característica crucial nos atuais SIS. Nestes sistemas diversos agentes devem interagir

para compartilhar seus resultados e, para tal devem usar uma terminologia médica com um significado claro e não ambíguo [47].

GALEN

Generalized Architecture for Languages, Encyclopaedias and Nomenclatures in medicine (GALEN), tem o objetivo de desenvolver e difundir modelos de terminologia, métodos, arquiteturas e ferramentas para aplicações clínicas. A partir de 1999 passou a se chamar OpenGALEN, é uma terminologia desenvolvida especialmente para especificar restrições utilizadas em domínios médicos. GALEN foi destinado a ser utilizado com diferentes idiomas e é integrada com diferentes esquemas de codificação. Baseia-se em um modelo de alto nível conhecido como GALEN CODing REference model (CORE model). O GALEN CORE define quatro categorias gerais: *Structures* (estruturas), *Substances* (substâncias), *Processes* (processos) e *Modifiers* (modificadores) [39, 47].

UMLS

A *Unified Medical Language System* (UMLS) foi criada pelo Instituto Nacional de Saúde dos EUA, buscando integrar e distribuir classificação, terminologia e os padrões de codificação, e os recursos associados para promover a criação de sistemas de informação biomédicos mais eficazes e interoperáveis e de serviços, incluindo registros de saúde eletrônicos [47].

A UMLS está estruturada em três partes [47]:

- Metatesauro (*metathesaurus*) contém informação biomédica sobre cada um dos termos incluídos na UMLS. Se um termo aparece em várias fontes, o que é usual, um conceito será criado na UMLS com um nome preferido a ele associado. As informações da fonte original sobre os termos (como fonte de definição, etc) estão ligadas ao conceito e algumas propriedades semânticas também são especificadas, tais como conceitos sinônimos, irmãos e pais, ou as relações entre os termos. Já na edição de 2002, o metatesauro continha cerca de 1,5 milhões de termos.
- Rede Semântica (*semantic network*) é uma ontologia de alto nível de conceitos biomédicos. As relações entre estes conceitos não foram derivadas das fontes biomédicas integrados no metatesauro UMLS, mas criadas como parte do UMLS com o objetivo de fornecer uma estrutura consistente ou de categorização em que os conceitos do metatesauro estão apoiados. Cada conceito do metatesauro está ligado a um conceito ou conceitos da ontologia. Esta ontologia foi desenvolvida, a fim de resolver a heterogeneidade entre as fontes UMLS, o que poderia ser considerado como o resultado da integração das fontes UMLS. Na edição de 2002, a ontologia da rede semântica continha 134 conceitos de alto nível e 54 conceitos de relacionamentos.
- Léxico Especialista contém informação sintática sobre termos biomédicos para ser usado em aplicações de processamento de linguagem natural.

ON9

Gangemi et al. [38, 39], desenvolveram a biblioteca de ontologias ON9 a partir da integração de ontologias das áreas médicas. Para integrar as ontologias foi utilizada a me-

ontologia ONION, desenvolvida desde 1992 para resolver problemas da heterogeneidade semântica. Em particular, a ontologia *Integrated-Medical-Ontology* (IMO), representa a integração de cinco ontologias médicas: a UMLS, a Snomed, a GMN (*Gabrieli Medical Nomenclature*), a ICD 10 e a GALEN *Core Model*.

3.2.2 Outros Modelos no Domínio da Saúde

Boa parte do conhecimento em saúde está formalizado em modelos menos expressivos que as ontologias, principalmente pelo fato do domínio da saúde ser bastante extenso e ainda por ser um processo natural na representação de conhecimento, dado a quantidade de conceitos envolvidos. Na sequência serão apresentados alguns modelos mais conhecidos.

SNOMED-CT

Systematized Nomenclature of Medicine – Clinical Terms (Snomed-CT) é a terminologia clínica integral, multilíngue e codificada de maior amplitude, precisão e importância desenvolvida no mundo. Snomed-CT é um produto que nasce da fusão entre o *Snomed-RT* (*Reference Terminology*), criada pelo *College of American Pathologists* (CAP) e o *Clinical Terms Version 3* (CTV3), desenvolvida pela *National Health Service* (NHS) do Reino Unido. Esta fusão permitiu a combinação dos termos nos âmbitos da ciência básica, a bioquímica e as especialidades médicas do Snomed-RT com o conteúdo de termos da Atenção Primária à Saúde do CTV3. Atualmente, esta terminologia é mantida e distribuída pela *International Health Terminology Standards Development Organisation* (IHTSDO) [51].

LOINC

O *Logical Observation Identifier, Names and Codes* (LOINC), criada em 1994, pela empresa Regenstrief Institute, surge com o objetivo de uniformizar as informações clínicas e laboratoriais facilitando, assim, a troca e análise de resultados. Trata-se de uma base de dados, livre de licença, composta por nomes e códigos que identificam observações clínicas e laboratoriais [59].

A base é formada por mais de 58.000 termos definidos por uma tupla de 6 partes para a especificação única de cada teste identificado, observação ou medição. As partes que compõem a especificação são [59]:

- Componente, o que é medido, avaliado, ou observado;
- Tipo da propriedade, características do que é medido, como o comprimento, massa, volume, prazo, etc;
- Aspecto de tempo, intervalo de tempo sobre o qual a observação ou medição foi feita;
- Sistema, contexto ou espécime-tipo em que a observação foi feita (exemplo: sangue, urina, ...);
- Tipo de escala, a escala de medida, pode ser quantitativa, ordinal, nominal ou narrativa;
- Tipo de método, procedimento utilizado para fazer a medição ou observação.

DICOM

O *Digital Imaging and Communications in Medicine* (DICOM) é um padrão aberto para representação de conhecimento médico baseado em imagens digitais. Sua especificação está formada por 16 capítulos. Esta especificação define os modelos de dados para as imagens médicas digitais e relatórios radiológicos, as transações e fluxos de informação entre os componentes do sistema, e os serviços que devem fornecer cada componente para poder implementar estes fluxos. O objetivo deste padrão é formalizar o fluxo de trabalho de solicitação e resposta da análises radiológicas [69].

OMS

A Organização Mundial de Saúde (OMS) é a autoridade diretiva e coordenadora das ações sanitárias no sistema das Nações Unidas, esta é a organização responsável por organizar os assuntos sanitários mundiais, coordenar as pesquisas na saúde, estabelecer normas, articular opções de políticas embasadas em evidências, prestar apoio técnico aos países e acompanhar as tendências sanitárias mundiais [29].

A OMS em seus trabalhos vem definindo uma série de classificações que apoiam a área da saúde na unificação dos termos utilizados para identificar tipos de registros, conhecidas como Família de Classificações da OMS, são elas:

- CID - Classificação Internacional de Doenças;
- CIAP - Classificação Internacional de Atenção Primária;
- CIF - Classificação Internacional de Funcionalidade, Incapacidade e Saúde;
- CIIS - Classificação Internacional de Intervenções em Saúde;
- CISP - Classificação Internacional para a Segurança do Paciente;
- CICEL - Classificação Internacional de Causas Externas de Lesões;
- CIPE - Classificação Internacional de Práticas de Enfermagem.

As classificações mais conhecidas no Brasil, são a CID, a CIF e a CIAP, no entanto apenas a CID é intensamente utilizada nos SIS.

Outras classificações e terminologias

No Brasil, segundo analista do DATASUS, existem outras classificações e terminologias que buscam uma identificação de procedimentos realizados em diversas ações e serviços de saúde públicos e privados nos diferentes níveis de atenção à saúde. Abaixo podemos ver algumas dessas representações, que apesar de terem sido criadas, essencialmente, para faturamento dos procedimento de saúde, tem sido usadas como listas de referências em SIS:

- SIGTAP - Tabela Unificada de Procedimentos do SUS;
- TUSS (ANS) - Terminologia Unificada em Saúde Suplementar;
- CBHPM - Classificação Brasileira Hierarquizada de Procedimentos Médicos e

- CBHPO - Classificação Brasileira Hierarquizada de Procedimentos Odontológicos.

Em especial a SIGTAP, desenvolvida pelo Ministério da Saúde, busca unificar em uma única tabela as classificações de procedimentos utilizadas no SUS.

3.3 Sistema Único de Saúde

O SUS é o sistema que organiza o cuidado integral a saúde da população brasileira. Como podemos ver na Figura 3.8, cuidado à saúde da população pode ser organizado em dois eixos: Vigilância em Saúde e Atenção à Saúde.

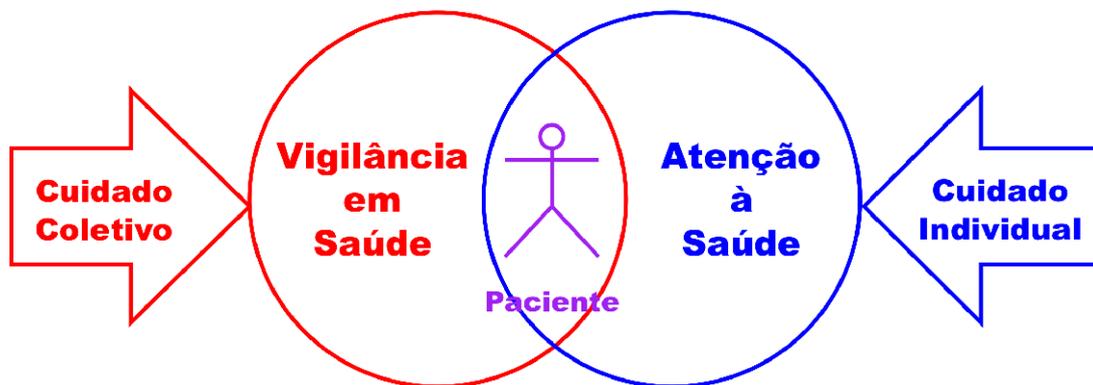


Figura 3.8: Sistema Único de Saúde

Segundo o Ministério da Saúde [23], a Vigilância em Saúde abrange as seguintes atividades: a vigilância das doenças transmissíveis, a vigilância das doenças e agravos não-transmissíveis e dos seus fatores de risco, a vigilância ambiental em saúde e a vigilância da situação de saúde. Atualmente é responsável por acompanhar o comportamento das doenças na sociedade, reunindo informações com objetivo de conhecer, detectar ou prever qualquer mudança que possa ocorrer nos fatores condicionantes do processo saúde-doença, bem como identificar a gravidade de novas doenças à saúde da população.

Ainda segundo o Ministério da Saúde [23], a Atenção à Saúde é tudo que envolve o cuidado com a saúde do ser humano, incluindo as ações e serviços de promoção, prevenção, reabilitação e tratamento de doenças. No SUS, o cuidado com a saúde está ordenado em níveis de atenção, que são a básica, a de média complexidade e a de alta complexidade. Essa estruturação visa à melhor programação e planejamento das ações e serviços do sistema. Não se deve, porém, considerar um desses níveis de atenção mais relevante que outro, porque a atenção à saúde deve ser integral. No entanto, se preconiza que seja a atenção básica que organize o cuidado da população.

Segundo o Departamento de Atenção Básica/MS [27], a Atenção Básica ou Primária à Saúde (ABS) caracteriza-se por um conjunto de ações de saúde, no âmbito individual e coletivo, que abrange a promoção e a proteção da saúde, a prevenção de agravos, o diagnóstico, o tratamento, a reabilitação e a manutenção da saúde. É desenvolvida por meio do exercício de práticas gerenciais e sanitárias democráticas e participativas, sob forma de trabalho em equipe, dirigidas a populações de territórios bem delimitados, pelas quais assume a responsabilidade sanitária, considerando a dinamicidade existente no território

em que vivem essas populações. Utiliza tecnologias de elevada complexidade e baixa densidade, que devem resolver os problemas de saúde de maior frequência e relevância em seu território. É o contato preferencial dos usuários com os sistemas de saúde. Orienta-se pelos princípios da universalidade, da acessibilidade e da coordenação do cuidado, do vínculo, da continuidade, da integralidade, da responsabilização, da humanização, da equidade e da participação social.

Segundo Starfield [81], a ABS, por meio das Unidades Básicas de Saúde (UBS)², devem garantir quatro atributos essenciais:

- A atenção ao primeiro contato deve implicar na acessibilidade e no uso do serviço a cada novo problema ou novo episódio de um problema pelo qual as pessoas buscam atenção à saúde.
- A longitudinalidade pressupõe a existência de uma fonte regular de atenção e seu uso ao longo do tempo. Assim, a UBS deve ser capaz de identificar a população eletiva, bem como os indivíduos dessa população. Além disso, o vínculo da população com sua fonte de atenção deveria ser refletida em fortes laços interpessoais que refletissem a cooperação mútua entre as pessoas e os profissionais de saúde.
- A integralidade implica que a UBS deve fazer arranjos para que o paciente receba todos os tipos de serviços de atenção à saúde, mesmo que alguns possam não ser oferecidos eficientemente dentro delas. Isto inclui o encaminhamento para serviços secundários para consultas, serviços terciários para manejo definitivo de problemas específicos e para serviços de suporte fundamentais, tais como internação domiciliar e outros serviços comunitários.
- A coordenação (integração) da atenção requer alguma forma de continuidade, seja por parte dos profissionais, seja por meio de prontuários médicos, ou ambos, além de reconhecimento de problemas. Este reconhecimento de problemas será facilitado se o mesmo profissional examinar o paciente no acompanhamento ou se houver um prontuário médico que esclareça estes problemas. Assim, tanto a continuidade como o reconhecimento de problemas são necessários para avaliar a coordenação da atenção.

Segundo o Departamento da Atenção Básica/MS [27], a Atenção Básica no Brasil tem a Saúde da Família como estratégia prioritária para sua organização de acordo com os preceitos do SUS e tem como fundamentos:

- I possibilitar o acesso universal e contínuo aos serviços de saúde de qualidade e resolutivos, caracterizados como a porta de entrada preferencial do sistema de saúde, com território adscrito de forma a permitir o planejamento e a programação descentralizada, e em consonância com o princípio da equidade;
- II efetivar a integralidade em seus vários aspectos, a saber: integração de ações programáticas e demanda espontânea; articulação das ações de promoção à saúde, prevenção de agravos, vigilância à saúde, tratamento e reabilitação, trabalho de forma interdisciplinar e em equipe, e coordenação do cuidado na rede de serviços;

²A Unidade Básica de Saúde ou Unidade de Atenção Básica é o estabelecimento de saúde que presta os serviços a população referente a atenção básica ou primária à saúde.

- III desenvolver relações de vínculo e responsabilização entre as equipes e a população adscrita garantindo a continuidade das ações de saúde e a longitudinalidade do cuidado;
- IV valorizar os profissionais de saúde por meio do estímulo e do acompanhamento constante de sua formação e capacitação;
- V realizar avaliação e acompanhamento sistemático dos resultados alcançados, como parte do processo de planejamento e programação; e
- VI estimular a participação popular e o controle social.

Garantir a integralidade e a continuidade do cuidado não são tarefas fáceis de se realizar frente a um complexo sistema que se apresenta. Para ilustrar essa complexidade, em parceria com o Departamento de Atenção Básica (DAB/SAS/MS), foi construído um mapa mental das interfaces da ABS no país, conforme apresentado na Figura 3.9.

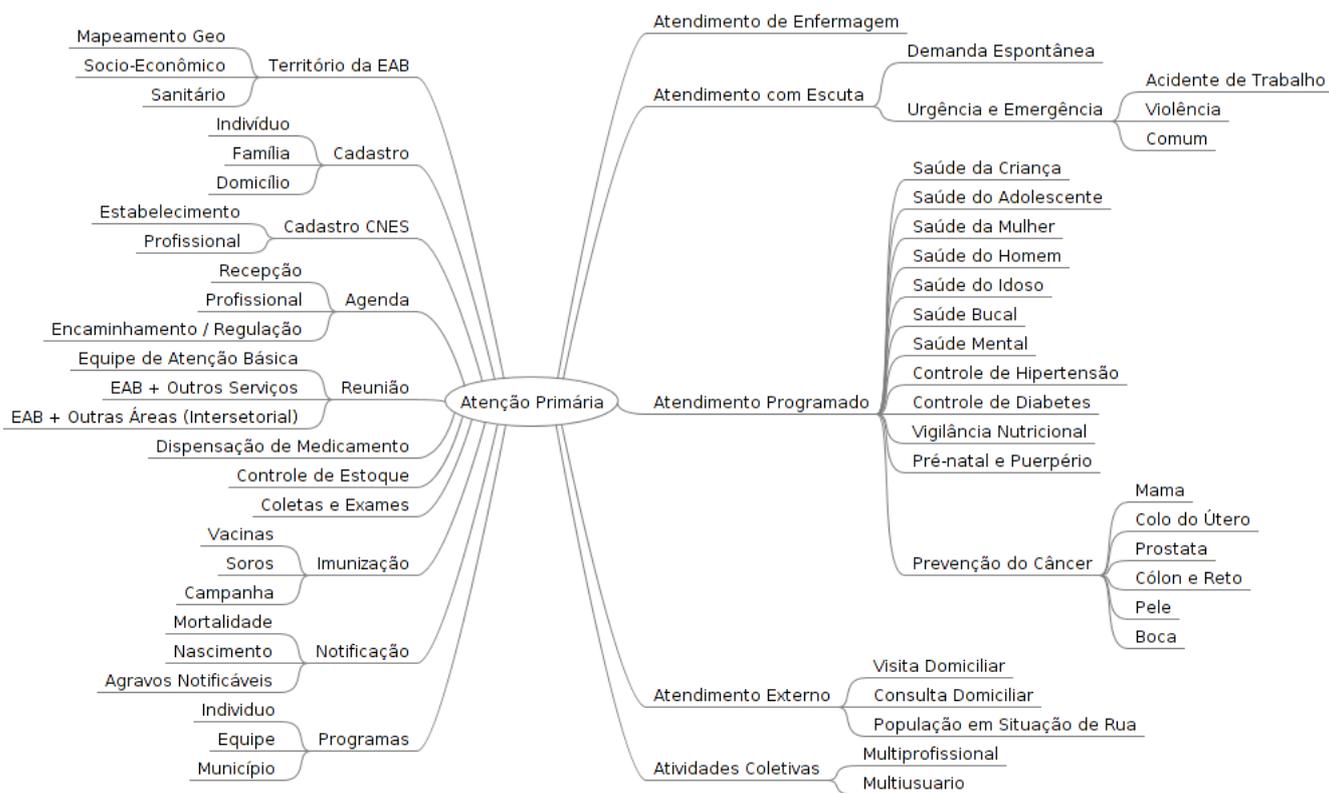


Figura 3.9: Interfaces da Atenção Primária no Brasil

Ter claro a dimensão do problema ajuda a optar por estratégias mais aderentes ao se buscar uma solução para interoperabilidade semântica desses diversos subsistemas que compõem a ABS. A complexidade que se apresenta, observando os subsistemas da ABS, ainda é ampliada se consideramos todos os subsistemas de atenção à saúde.

Muitas características diferentes determinam a forma como os sistemas e subsistemas de saúde operam. A multiplicidade destas características oferece oportunidade para inúmeras permutas e combinações, de forma que nenhum sistema de saúde é parecido

nem atua da mesma forma que outro [81]. Esta variabilidade exige dos profissionais da Informática em Saúde (IS) o estudo de técnicas que permitam o avanço da representação de conhecimento sem criar amarras que impossibilitem aprimoramentos nos protocolos de cuidado de atenção à saúde e adequações para especificidades regionais. Neste ponto é importante avaliar o contexto da IS e, de forma resumida, como a área vem evoluindo.

3.4 Informática em Saúde

Nos Estados Unidos, a primeira definição de Informática Médica apareceu nos anúncios em 1977 para a *Third World Conference on Medical Informatics (MEDINFO 80)* em Tóquio, onde Morris F. Collen, definiu o termo: “Informática Médica como a aplicação da tecnologia de computação para todos os campos da medicina, assistência médica, ensino médico e pesquisa médica.”. No entanto, Shortliffe em Stanford, enfatizou que a informática médica inclui mais do que as aplicações da tecnologia de computadores para a medicina porque alguns pesquisadores começaram a estudar a ciência básica da área de informática médica como sujeito e não como uma ferramenta [22].

No Brasil, o termo “informática médica” foi alterado para “informática em saúde” pela Sociedade Brasileira de Informática em Saúde (SBIS), que resolveu utilizar o termo mais amplo “saúde” ao invés de “médica”, ao contrário do que se faz na Europa, Ásia e nos Estados Unidos, pois o adjetivo “*medical*” é utilizado em sentido tão amplo quanto o “saúde” que adotamos no Brasil [77].

O crescimento da Informática em Saúde (IS) como uma disciplina deve-se, em grande parte: aos avanços nas tecnologias de computação e comunicação, à crescente convicção de que o conhecimento médico e as informações sobre os pacientes (clínicas e/ou sócio-econômicas) são significativamente complexas e de difícil gerenciamento por métodos tradicionais.

A área de informática em saúde, aponta a padronização de Registro Eletrônico de Saúde (RES) como uma forma para se alcançar a interoperabilidade entre os SIS [17, 31, 33, 58]. É uma das poucas áreas a ter conhecimento de domínio consideravelmente exposto através de padrões. Esses padrões oferecem um valor significativo em termos de expressar a semântica de serviços para interoperabilidade de SIS [31].

A primeira questão a ser respondida é sobre o próprio conceito de RES. Apesar de existir um entendimento comum, cabe fazer uma pequena definição de RES e como usamos esse termo neste trabalho.

3.5 Registro Eletrônico de Saúde

O RES, ou *Electronic Health Record (EHR)*, é um registro longitudinal eletrônico de informações de saúde do paciente gerado por um ou mais encontros em qualquer configuração de prestação de cuidados. Incluem-se neste registro dados demográficos dos pacientes, notas de progresso, problemas, medicamentos, sinais vitais, história patológica pregressa, imunizações, dados laboratoriais e relatórios de radiologia. O RES automatiza e simplifica o fluxo de trabalho clínico. O RES tem a capacidade de gerar um registro completo de um encontro clínico com o paciente, bem como apoio a outros cuidados de

atividades relacionadas direta ou indiretamente, incluindo apoio à decisão baseada em evidências, gestão da qualidade, e elaboração de relatórios de resultados [49, 83].

Burton et al. [17] declara que o uso de RES auxilia a transmitir com segurança dados de pacientes entre os médicos, pode ajudar a coordenar a atenção à saúde, em especial o cuidado de condições crônicas, elementos essenciais para se garantir uma ABS de qualidade. O RES ainda pode apoiar as estruturas de governança regional para incentivar a troca de dados de clínicos.

No Capítulo que segue, apresentamos a proposta do trabalho visando um modelo de conhecimento que possibilite a padronização do registro, e ainda uma delimitação do escopo, considerando um RES da ABS.

Capítulo 4

Proposta

A proposta deste trabalho é descrever um modelo de interoperabilidade semântica, baseada em arquétipos para gerar registros clínicos a partir de informações armazenadas em sistemas de RES. Neste trabalho, serão feitas referências a RES de três formas: (i) RES-Completo, que deve atender toda a demanda de atenção à saúde; (ii) RES-Parcial, que é aderente ao RES-Completo, no entanto está preocupado com algum subsistema da atenção à saúde; e (iii) RES-ABS, um RES-Parcial com conceitos de mais alto nível, visando um nível de fundamentação e que se preocupa com as necessidades de um modelo para a ABS.

A relação do RES-Completo com o RES-ABS é análoga à relação entre a atenção à saúde e o nível de atenção básica. Ou seja, como visto na Seção 3.3, assim como o nível da atenção básica deve garantir a continuidade do cuidado de um paciente na atenção à saúde, o RES-ABS deve garantir uma estrutura mais geral, capaz de servir como um nível de fundamentação para a construção do RES-Completo. Ainda podemos dizer que a partir da especialização dos conceitos do RES-ABS pode-se alcançar a complexidade e o detalhamento necessário para atender as necessidades dos outros níveis de atenção, e.g., um RES-Parcial que atenda a especialidade médica da cardiologia. Na Figura 4.1, ilustramos a relação entre RES-Completo, RES-Parcial e RES-ABS.

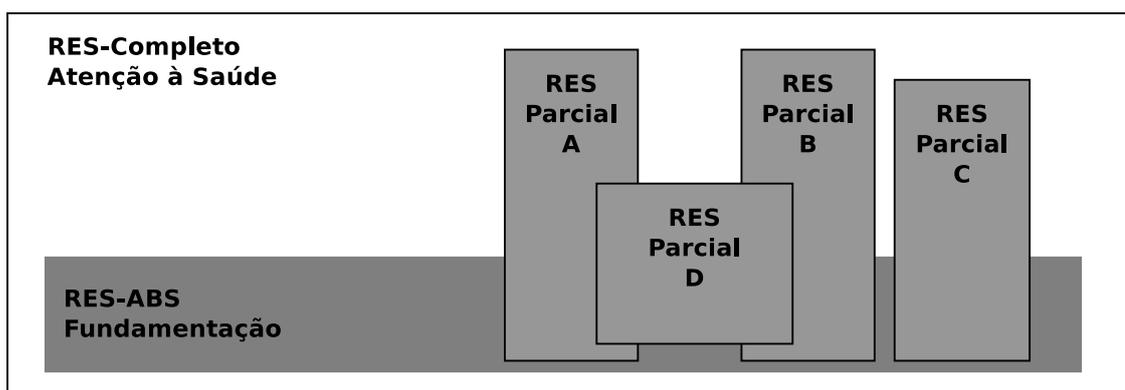


Figura 4.1: Relação do RES com a ABS

Atualmente, pode-se dizer que não existem um RES-Completo que dê conta das necessidades de atenção à saúde no Brasil. Tampouco seria possível, frente a complexidade do próprio domínio da atenção a saúde, construir um modelo que atenda a demanda de

um RES-Completo sem um processo colaborativo, organizado e que possibilite evolução do modelo ao longo do tempo.

Através de um estudo de caso, mostraremos que os arquétipos aumentam a interoperabilidade semântica dos sistemas de RES, pois além de ter um modelo de construção em multi-camadas, ainda permitem uma vinculação (*binding*) com conceitos já definidos em ontologias, terminologias e classificações no domínio da saúde. A Figura 4.2 apresenta a associação de conceitos usando arquétipos, conforme proposto por Bisbal et al. [13].

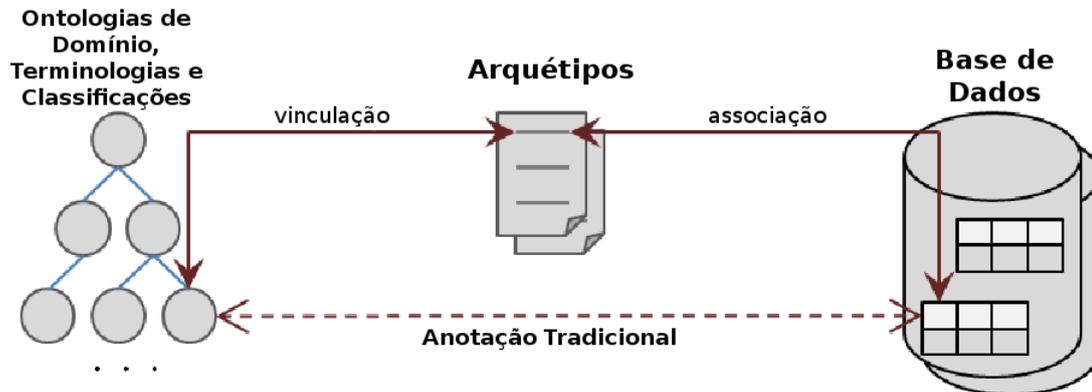


Figura 4.2: Associação de conceitos usando arquétipos

Os arquétipos ainda permitem a definição de uma estrutura, por meio da linguagem ADL, mais próxima dos modelos de dados objeto-relacionais ¹ e portanto é mais fácil de entender as estruturas e de fazer um mapeamento das estruturas de dados para os modelos conceituais. Esta é uma característica importante para se ter resolutividade dos problemas de interoperabilidade semântica.

Ao se pensar em um modelo de interoperabilidade semântica, baseada em arquétipos, é necessário pensar em um processo colaborativo, organizado e que possibilite evolução do modelo ao longo do tempo. Logo, é importante definir um método de desenvolvimento arquétipos capaz de reutilizar os conceitos já definidos e principalmente garantir que a estrutura conceitual geral garanta a interoperabilidade semântica dos SIS em questão. E anterior a definição dos arquétipos é necessário ter claro qual é o Modelo de Referência mais adequado a aplicação desejada, conforme visto na Seção 3.1. Tendo os arquétipos desenvolvidos ainda é importante verificar se o modelo conceitual está de acordo com a aplicação. Em particular, considerando sistemas legados, é necessário verificar se os modelos definidos são capazes de representar as informações contidas no SIS (extrair informações instanciadas a partir dos modelos).

Por conseguinte, esta proposta se concentra em três macro etapas, como visto na Figura 4.3, a saber:

1. Seleção do Modelo de Referência;
2. Desenvolvimento de Arquétipos; e

¹Temos convicção que pelos arquétipos terem características similares as técnicas de modelos objeto-relacionais, amplamente usadas e conhecidas na Ciência da Computação, eles são de mais fácil aplicação do que as técnicas apoiadas em modelos que usam descrição lógica, como OWL.

3. Extrato de Dados.

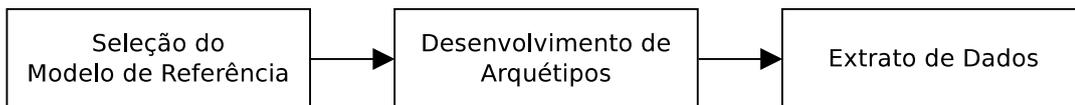


Figura 4.3: Macro-etapas da Proposta

A macro-etapa de seleção do modelo de referência, busca evidenciar características favoráveis ao se trabalhar com o modelo de referência do padrão openEHR, de forma que este seja aderente às necessidade de construção de um modelo que atenda as necessidade do modelo de informação da ABS. A macro-etapa de desenvolvimento de arquétipos visa descrever um método colaborativo e organizado para evolução dos modelos de conhecimentos baseados em arquétipos. E por último a macro-etapa de extrato de dados, descreve uma forma de se usar os modelos de arquétipo como meio de comunicação de tal forma que permita que os sistema novos e legados troquem informações. Para um maior esclarecimento, as macro-etapas são detalhadas nas Seções 4.1, 4.2 e 4.3.

4.1 Seleção do Modelo de Referência

Para se trabalhar com um modelo de dois níveis é necessário selecionar o Modelo de Referência que se pretende utilizar. Este modelo, como já mencionado, irá prover a estrutura mais geral das classes a serem utilizadas para definir os arquétipos e define a parte invariante do modelo. A parte variante, ou seja, os arquétipos, devem estender as classes mais gerais definindo a aplicação que será dada ao modelo.

Nesta Seção, iremos fazer um pequena discussão sobre o Modelo de Referência mais adequado para o contexto da Atenção Básica à Saúde, observando elementos que podem contribuir tanto na padronização da informação, quanto no uso do modelo de RES.

4.1.1 Comparando Modelos de Referência

Conforme descrito na Seção 3.1.1, existem três modelos de referência mais consolidados. Na Figura 4.4 [72], podemos observar uma análise da relação entre esses principais Modelos de Referência, feita por Schloeffel et al. [72] que teve como objetivo selecionar o Modelo de Referência para o compartilhamento de RES na Austrália. Esta avaliação indica, que o padrão openEHR é o mais abrangente, ou seja, apesar de ser um modelo de referência, tem uma visão consensuada mais ampla.

Na Figura 4.5, adaptada de [63], podemos comparar os modelos de referência do padrão openEHR e da norma CEN/ISO 13606, ainda pelo fato de que a norma foi derivada do padrão openEHR em um esforço de produzir um padrão mais geral [72], destacamos as classes da norma CEN/ISO 13606 sob as classes do padrão openEHR e apesar das classes *LOCATABLE*, na norma CEN/ISO, se chamar *RECORD_COMPONENT* e *CONTENT*, na norma CEN/ISO, se chamar *CONTENT_ITEM*, fica claro que a norma CEN/ISO 13606 é um modelo mais geral, e portanto o padrão openEHR é mais rico semanticamente.

Pelo modelo do openEHR ser mais específico, principalmente, por incluir a forma de raciocínio da clínica, conforme visto na Seção 3.1.1, fica mais intuitivo e mais simples,

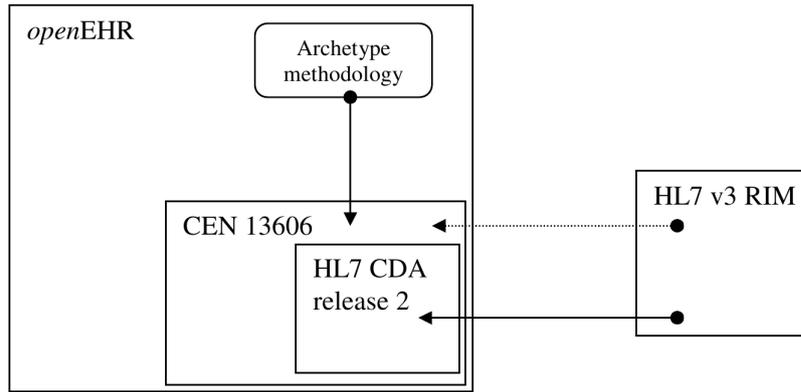


Figura 4.4: Relação entre os principais modelos de referências

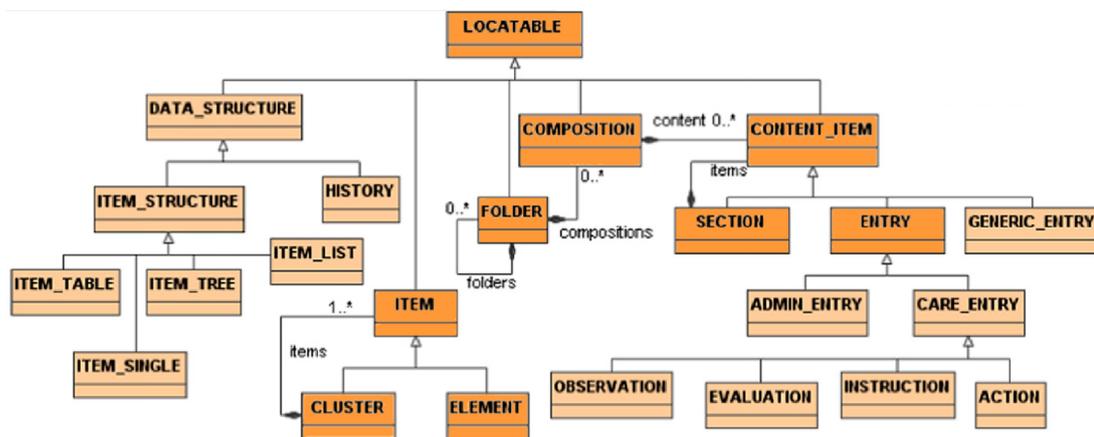


Figura 4.5: Comparação entre os modelos de referências OpenEHR e CEN/ISO 13606

representar o processo do cuidado sobre esse modelo. Uma questão importante a ser destacada no modelo do padrão openEHR é a capacidade, por exemplo, de oferecer um melhor suporte ao Registro Médico Orientado a Problemas (RMOP), proposto por Lawrence Weed em 1968, e registrado através do método SOAP, acrônimo que faz referência a sequência de evolução: Subjetivo, Objetivo, Avaliação e Plano, ou do inglês, *Subjective, Objective, Assessment e Plan*). Este método é apontado como um dos princípios da clínica médica e da medicina de família e comunidade [46], em especial na atenção básica à saúde [81], ver Seção 3.3, por ampliar a organização dos registros clínicos e a orientação do olhar do cuidado sobre o paciente e não apenas sobre a doença.

Na Figura 4.6 [9], podemos ver um exemplo de uso do openEHR na organização das informação da evolução clínica usando o método SOAP, representado por seções (*SECTION*) de informação. O exemplo mostra uma “evolução” do atendimento a um paciente que tem um problema de “diabetes”, onde o profissional de saúde registra uma queixa de “dor de cabeça”, a “pressão arterial” e o “peso” do paciente e diagnostica que o paciente está com “hipertensão arterial” e o instrui a usar “pouco sal” na sua dieta.

Os elementos apresentados, nos fazem ter convicção que o modelo openEHR é o modelo mais apropriado para se representar o conhecimento da ABS. Além disso, o padrão openEHR foi o padrão apresentado no CPIISS, conforme Item 1 do Catálogo de Padrões

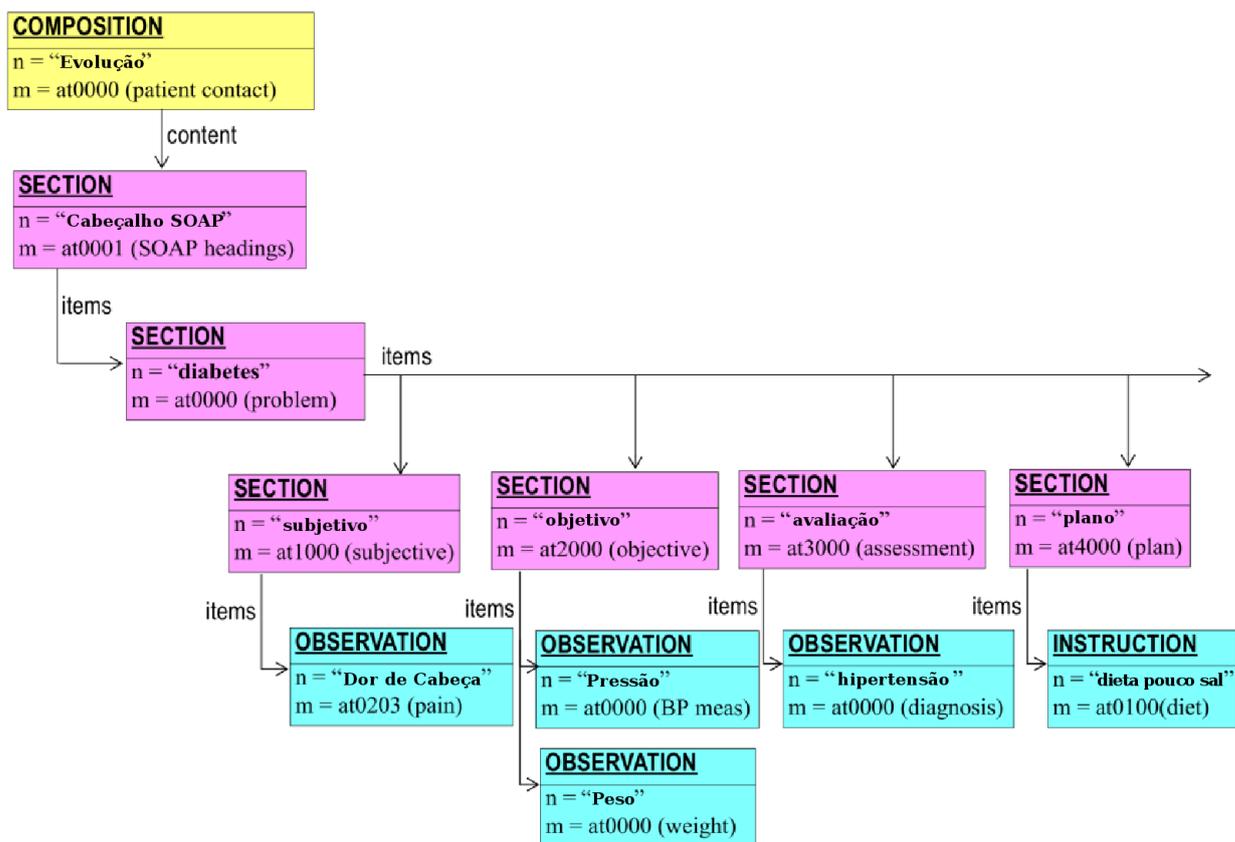


Figura 4.6: Exemplo de estruturação do SOAP usando openEHR

de Informação, ver Apêndice A, como o padrão para definição de Registro Eletrônico de Saúde.

4.2 Desenvolvimento de Arquétipos

Após definir o Modelo de Referência a ser utilizado é necessário desenvolver os arquétipos que farão a representação do conhecimento da aplicação em questão.

Para o desenvolvimento dos arquétipos, foi utilizado uma versão adaptada da metodologia proposta por Santos [71]. No trabalho dele, existiu uma preocupação quanto ao RES ser sumarizado, para diminuir a quantidade de dados a serem transmitidos. Neste trabalho o objetivo é poder representar as informações de um SIS por meio dos arquétipos, logo não existiu a preocupação de se fazer uma análise quanto ao modelo de conhecimento ser sumarizado, tampouco ser otimizado já no nível do modelo. Neste trabalho houve um esforço no sentido de descrever o método de forma mais geral, no entanto, as especificidades do experimento serão discutidas no estudo de caso apresentado no Capítulo 5.

4.2.1 Metodologia de Desenvolvimento de Arquétipos

A metodologia de desenvolvimento de arquétipos, segundo Santos [71], inclui seis etapas² considerando um processo mais geral do desenvolvimento de arquétipos, a saber:

1. Seleção dos elementos de dados;
2. Identificação dos conceitos candidatos;
3. Pesquisa de arquétipos existentes;
4. Modelagem dos arquétipos;
5. Análise e Identificação de terminologias;
6. Codificação de arquétipos.

As etapas do desenvolvimento de arquétipos, não necessariamente precisam ocorrer na ordem apresentada, no entanto existe alguma dependência entre as etapas, que podem ser melhor visualizadas por meio da Figura 4.7.

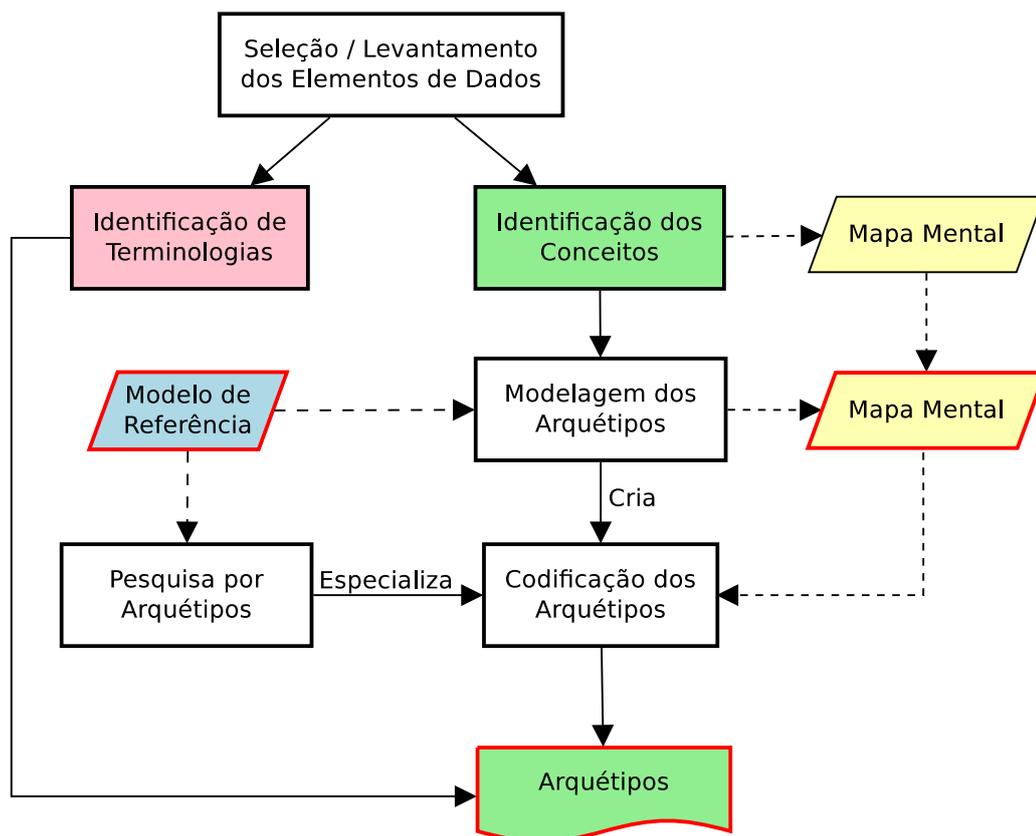


Figura 4.7: Metodologia para desenvolver de arquétipos

²A etapa de análise e identificação de regras de permanência foi suprimida, pois se refere a uma avaliação quanto a persistência de uma informação/elemento em relação a relevância clínica dessa informação e isto depende de uma aplicação ou de uma forma particular de se fazer o cuidado ao paciente.

Na Figura 4.7, podemos ver que além das seis etapas, o método sugere o desenvolvimento de etapas intermediárias, por meio da construção de diagramas do tipo mapa mental, anteriores as etapas de modelagem e codificação dos arquétipos para melhor organizar os conceitos e facilitar a evolução das estruturas de arquétipos. Também podemos ver que as etapas de pesquisa e modelagem dos arquétipos dependem da prévia seleção do modelo de referência dos arquétipos. Para melhor compreender o método de desenvolvimento de arquétipos será apresentada uma breve descrição de cada etapa.

ETAPA 1 - Seleção dos Elementos de Dados

A etapa de **seleção dos elementos**, como o próprio nome refere, seleciona os elementos de dados a serem representados. Os elementos são selecionados com base em alguns princípios claros que buscam restringir o escopo do modelo de conhecimento a ser representado. Um exemplo de princípio, pode ser, somente serão selecionados elementos totalmente consensuados entre os especialistas da área de conhecimento. Os elementos não consensuados ou parcialmente consensuados poderão ser representados como especialização dos conceitos básicos.

Após definir os princípios da seleção é necessário escolher os documentos ou instrumentos de referência do modelo de conhecimento a ser representado. A seleção é feita de forma similar ao levantamento de requisitos de um sistema, apoiado em instrumentos e entrevistas aos especialistas de forma a capturar as informações necessárias para a aplicação. Cada um dos elementos de informação selecionados é submetido a uma análise para a identificação do domínio em relação ao tipo de informação e a possível representação por listas de termos, como por exemplo, a formalização de listas de opções do motivo de alta de um paciente. Na seleção dos elementos é importante evitar informações mais difíceis de se tratar computacionalmente, como p.e., campos de textos, no entanto se percebe que em alguns casos os modelos contemplam informações desse tipo para contemplar lacunas na modelagem dos conceitos.

ETAPA 2 - Identificação dos Conceitos Candidatos

Tendo o conjunto de elementos, a etapa seguinte é fazer a **identificação dos conceitos candidatos**, ou seja, elencar quais elementos de dados serão representados nos modelos de conhecimento. A metodologia sugere a criação de mapas mentais³ (*mind maps*) que representem a organização dos conceitos identificados, como uma forma mais intuitiva de validar os conceitos pelos especialistas da área. Os mapas mentais podem ser produzidos sobre os conceitos mais gerais ou mais específicos.



Figura 4.8: Mapa mental do conceito Peso Corporal

³Mapa Mental, criado por Buzan and Buzan [18], é uma ferramenta de anotação de informações de forma não linear, construído em forma de teia, onde a ideia principal é colocada no centro de uma folha de papel branco, usada na horizontal para proporcionar maior visibilidade às ideias descritas apenas com palavras chaves e ilustradas com imagens, ícones e cores

ETAPA 3 e 4 - Pesquisa e Modelagem dos Arquétipos

Após identificar os conceitos, a próxima etapa é fazer a **modelagem dos arquétipos** de acordo com o modelo de referência escolhido. Nesta etapa, é importante verificar modelos de conhecimentos já definidos nos repositórios, ou seja, fazer uma **pesquisa de arquétipos existente**, como p.e., no repositório da Fundação openEHR⁴, visando uma maior padronização dos conceitos. Ao encontrar conceitos mais gerais, é possível especializar (*specialize*) os conceitos ao invés de criar novos conceitos. Na Figura 4.9 pode ser visto um exemplo de arquétipo, peso corporal, já definido no idioma português do Brasil diagramado por mapa mental pela ferramenta *Clinical Knowledge Manager* (CKM) dos repositórios da Fundação openEHR.

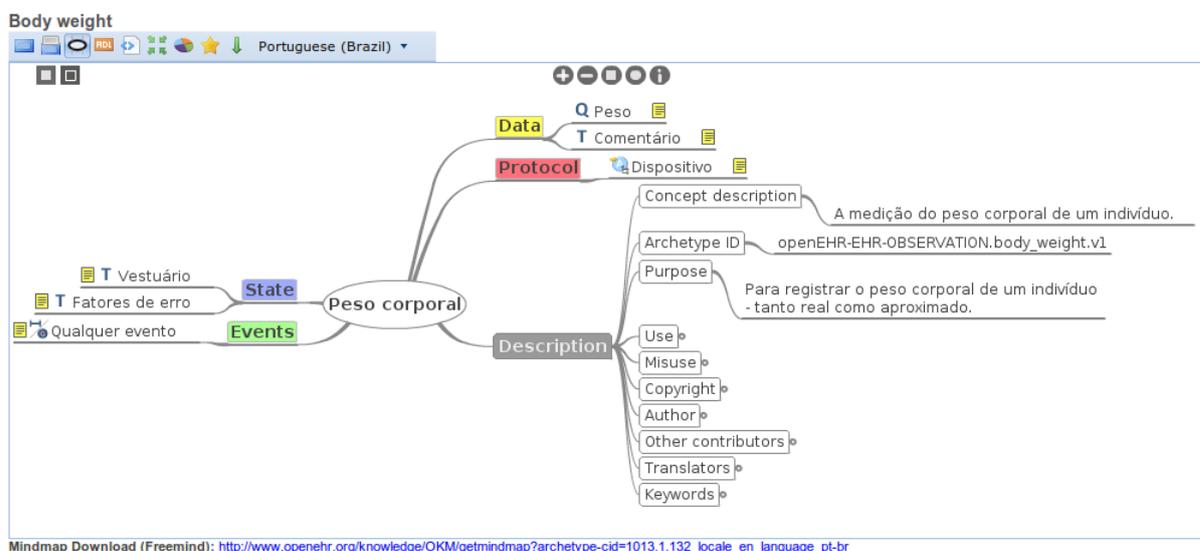


Figura 4.9: Mapa mental do conceito Peso Corporal usando arquétipos openEHR

ETAPA 5 - Análise e Identificação de Terminologias

Após identificar ou criar os arquétipos, a próxima etapa é a **análise e identificação de terminologias**, avaliando as possíveis terminologias descritivas ou listas de termos a serem ligadas ao modelo de conhecimento, considerando um maior controle dos vocabulários utilizados e visando maior interoperabilidade entre os SIS.

Esta etapa pode considerar terminologias existentes e consolidadas ou ainda a necessidade de definir formalmente novos vocabulários considerando contextos mais específicos. Esta etapa tem uma função importante para a interoperabilidade semântica, pois uma boa associação dos termos irá garantir uma vinculação com conceitos bem definidos no domínio da saúde. Usando a estrutura de arquétipos é possível identificar mais de um conceito, e com isso vincular a termos equivalentes ao usado no modelo de conhecimento.

Os arquétipos podem estender a representação dos conceitos usando ligações do tipo *term_binding* e *constraint_binding*, respectivamente, para vincular a um conceito definido e para restringir um conjunto de termo, seja este um conjunto consolidado ou um conjunto

⁴<http://www.openehr.org/knowledge/>

menos consolidado, no entanto consensuado no grupo de trabalho que está produzindo o modelo de conhecimento.

A vinculação dos termos aos arquétipos pode ser relacionada em diferentes terminologias. Por exemplo, a vinculação com a enfermidade *Diabetes Mellitus*, pode ocorrer por meio da vinculação do código CID:E14 *Diabetes mellitus não especificado*, MeSH:C18.452.394.750 *Diabetes Mellitus*, SNOMEDCT:73211009 *Diabetes Mellitus* ou ainda com alguma perda de informação à CIAP, CIAP:T90 *Diabetes não insulino-dependente* ou CIAP:T89 *Diabetes insulino-dependente*. Vale ressaltar a dificuldade em se trabalhar com codificações fechadas como é o caso do Snomed-CT. Apesar da terminologia Snomed-CT atualmente ser mantida por uma organização sem fins lucrativos, a IHTSDO (*International Health Terminology Standards Development Organisation*), é uma terminologia que está sob a licença *UMLS Metathesaurus License* que demanda taxa de pagamento para seu uso.

ETAPA 6 - Codificação dos Arquétipos

Por fim, a última etapa é a **codificação de arquétipos**, a linguagem sugerida pelo método é a linguagem ADL, ver Seção 3.1.2, no entanto pode ser usada qualquer linguagem de representação de conhecimento. Ao selecionar a linguagem é possível escolher um editor que auxilie nessa tarefa, ou ainda usar um editor de texto puro, no entanto os editores específicos das linguagens de representação de conhecimento vem trazendo cada vez mais ferramentas de apoio as boas práticas de desenvolvimento de ontologias, e são altamente recomendados para produzir ontologias de qualidade.

4.3 Extraindo dados semanticamente enriquecidos

Os SIS podem gerar extratos de dados baseados nos arquétipos, ou seja, por meio de algum mecanismo de exportação de dados de registros clínicos encapsulados em extratos de RES restringidos semanticamente pelos arquétipos dos conceitos definidos. Na Figura 4.10, podemos visualizar um esquema que ilustra a relação entre os modelos e os extratos semanticamente enriquecidos. Ainda podemos ver a relação ontológica do modelo de dois níveis, sendo o modelo de referência considerada uma ontologia de domínio e os arquétipos como ontologias de aplicação e vinculados as terminologias e classificações. Os extratos de dados dos SIS são encapsulados em extratos de RES (*EHR_EXTRACT*), restringidos semanticamente pelos arquétipos definidos, ou seja, podemos dizer que os extratos de dados são instâncias das ontologias definidas em duas camadas pelo modelo de dois níveis.

Note pela Figura 4.10 que uma vez definidos os arquétipos, estes podem ser usados como componentes de *middleware* semântico, o que significa que os SIS que já contem um histórico de registros clínicos podem ter seus dados encapsulados em padrões XML aderentes a estrutura estabelecida. Desta forma, mesmo sem obedecer internamente ao padrão escolhido, possam se comunicar fazendo uso dele, por meio de algum mecanismo de comunicação, como p.e. um barramento de comunicação orientado a serviço. Neste trabalho não trataremos as questões da forma de comunicação entre os sistemas, no entanto este problema pode ser resolvido, por exemplo, usando os *profiles* IHE⁵[33].

⁵Os *profiles* IHE fornecem uma linguagem comum para métodos e procedimentos claros para a implementação de normas de comunicação, cuidadosamente documentados, revistos e testados. Eles oferecem

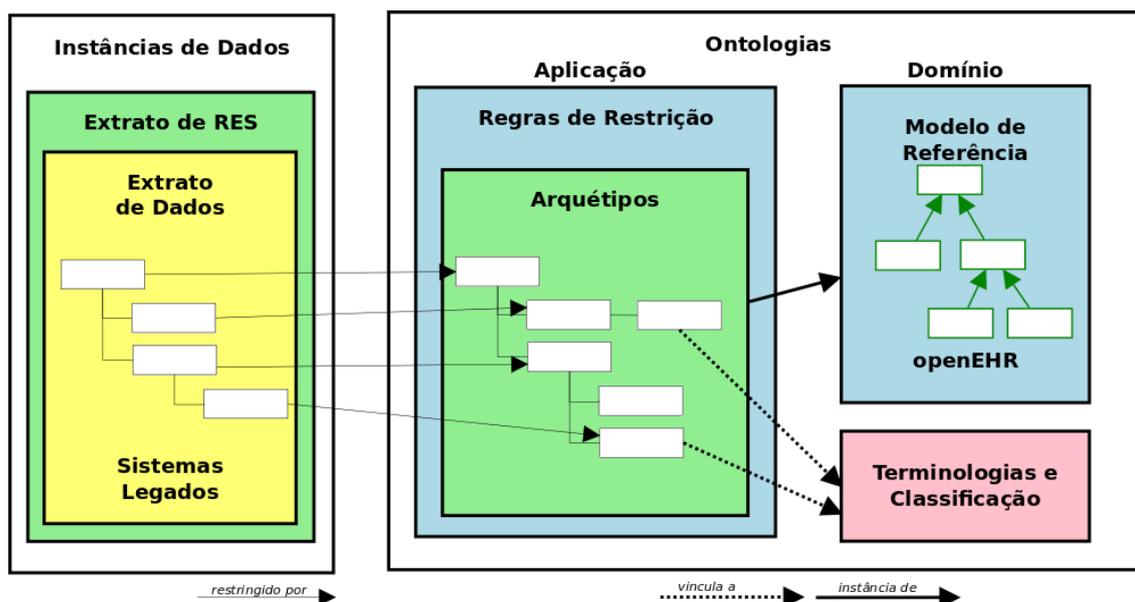


Figura 4.10: Extrato RES, openEHR, de SIS

É possível operacionalizar o uso de extratos openEHR como mostrado na Figura 4.11 (adaptada de [5]). Nesta ilustração, uma solicitação de informações dos registros de um ou mais de um objeto é criado por um sistema solicitante. O registro solicitado pode ser um RES de um paciente, um registro de uma pessoa de um sistema demográfico, ou qualquer outra informação definida pela entidade de mais alto nível. O sistema solicitante recebe a resposta sob a forma de um ou mais extratos. A interação de solicitação / resposta pode ser ativada por um mecanismo de transporte ou eventualmente por outros serviços, como p.e. um *web service*, ou um protocolos simples, ponto-a-ponto, como um serviço SMTP de e-mail.

4.3.1 Delimitação do Escopo de Trabalho

Observando a Figura 3.9, que ilustra a complexidade do SIS da ABS, é importante neste ponto fazer uma delimitação do escopo de trabalho, considerando essencialmente o tempo necessário para se desenvolver um modelo de conhecimento que dê conta de representar toda esse complexidade.

Optou-se então, por fazer um estudo de caso em um subsistema da ABS, e portanto, cabe ressaltar que o modelo de conhecimento desenvolvido deva atender a um RES-parcial e não a um RES-completo, apesar de ter aderência ao modelo completo pelo uso de conceitos comuns e pela adoção do padrão openEHR como modelo de referência.

Vigilância Alimentar e Nutricional

Uma das ações desenvolvidas pela ABS é a vigilância alimentar e nutricional da população, justificado por relevantes mudanças socioeconômicas, geográficas, políticas e tecnológicas

uma ferramenta que reduz o custo, complexidade e ansiedade de implementação de sistemas interoperáveis.

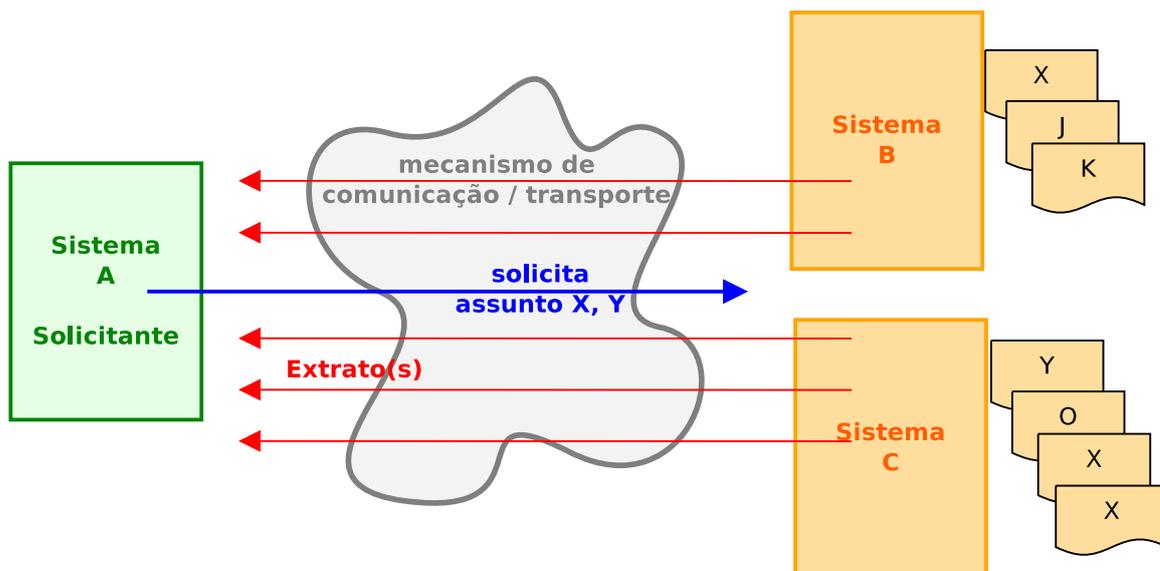


Figura 4.11: Solicitando extratos openEHR

ocorridas nas últimas décadas. Como consequência dessas mudanças, cresce o consumo de alimentos calóricos, com alto teor de açúcares, gorduras, sal e aditivos químicos, que são pobres em nutrientes. E ainda associado ao baixo consumo de frutas, legumes e verduras resultaram em alterações significativas no perfil de morbidade e mortalidade e nos padrões do consumo alimentar e do estilo de vida da população, determinando os fenômenos das denominadas transições demográfica, epidemiológica e nutricional [79].

A Vigilância Alimentar e Nutricional é responsável pela coleta e análise de informações sobre a situação alimentar e nutricional de indivíduos e coletividades, com o propósito de fundamentar medidas destinadas a prevenir ou corrigir problemas detectados ou potenciais. As condições alimentares e nutricionais da população são acompanhadas pelo Sistema de Vigilância Alimentar e Nutricional (SISVAN). Sistema que faz o monitoramento da situação alimentar e nutricional da população por meio da coleta, do processamento e da análise de dados antropométricos [16].

Na Seção 4.4 serão apresentados alguns trabalhos correlatos que apontam diversas possibilidades de ampliação desta proposta. Ainda serão discutidas algumas aplicações feitas usando a técnica de modelagem de dois níveis e o modelo de referência do openEHR.

4.4 Trabalhos Correlatos

Luna et al. [60], grupo ligado ao Hospital Italiano de Buenos Aires, apresentam um trabalho de implementação a utilização de documentos HL7 CDA para transferência de informação clínica detalhada sob um formato padrão. Nesse projeto foram implementados documentos clínicos de vários eventos de saúde e ainda um mecanismo de assinatura digital de validade dos registros por um participante (*Legal Authenticator*) utilizando a infraestrutura de chaves públicas e privadas. Este trabalho traz uma questão importante sobre a segurança da informação, em particular, ao se trabalhar com RES, a questão de segurança deve ser bastante considerada e alinha com a legislação vigente no país referente aos participantes que podem ou não ver determinadas informação sobre um

paciente. O padrão do openEHR, ao contrário do padrão HL7 CDA, já implementa classes e definem propriedades dos objetos que ampliam a segurança da informação, tanto nas características de auditoria quanto nas de controle de acesso.

Santos [70] apresenta um trabalho de integração entre as áreas de engenharia de software e de ontologias, sendo sua principal contribuição a possibilidade de acelerar o processo de geração de ontologias de aplicações por meio de tradução dos modelos de software representados através da linguagem UML (*Unified Model Language*) e serializada em documentos XMI (XML Metadata Interchange) para o OWL. Ainda apresenta um estudo de caso nas áreas médicas, mas especificamente em sistemas de geração de laudos dos exames de Cintilografia Cardíaca (CINT), Eletrocardiograma (ECG), Tomografia Computadorizada (CT) e Ressonância Magnética Nuclear (RMN), para avaliar sua proposta.

O trabalho de Santos [70] traz questões importantes sobre a reutilização da modelagem de uma aplicação usando UML na criação de ontologias de aplicações em saúde, no entanto, sem se preocupar com a modelagem de conhecimento pré-existente, como por exemplo, ontologias ou terminologias em saúde. No domínio da saúde gerar um modelo ontológico “desconectado” pode prejudicar a interoperabilidade semântica com outros SIS, tornando o modelo de conhecimento limitado à aplicação.

Nardon et al. [67] e Neira et al. [68] apresentam um trabalho sobre as questões a serem enfrentadas na construção de aplicações em saúde baseadas em arquétipos do padrão openEHR. Como conclusões apontam que ainda são necessários estudos relativos à melhor forma de representação dos elementos de interface gráfica para a construção de aplicações sólidas e que representem adequadamente os conceitos que são utilizados, também colocam que os custos relativos ao desenvolvimento de aplicações ainda é alto, mas que com o crescimento e apoio da comunidade internacional na implementação de ferramentas que apoiem o desenvolvimento de sistemas reais esse custo deve cair significativamente.

Dias and Freire [30] apresentam arquétipos de representação de dados demográficos baseada em uma análise de diversos sistemas de informação públicos do país e do padrão TISS, bem como diversos padrões internacionais e nacionais para representação dessas informações. O resultado dessa análise demonstra claramente a falta de padronização de conceitos, modelagem e estrutura dos dados e aponta a modelagem em dois níveis do padrão openEHR com uma potente ferramenta para a solução de interpretação e interoperabilidade entre os sistemas de informação de RES. O modelo demográfico do openEHR define conceitos importantes como as classes ROLE (papel) e PARTY_RELATIONSHIP (relação) para pessoas e organizações, dessa forma consegue representar os diversos modos como os atores exercem seus papéis na sociedade e como eles se relacionam entre si. Ainda conclui que esse é um ponto importante a ser considerado na adoção de padrão para RES.

Carrasco and Gutiérrez [21] desenvolveram, junto ao Serviço de Emergência do Hospital Maciel dentro da unidade de docência da Faculdade de Medicina da Universidade da República do Uruguai, o projeto TRAUMAGEN, um sistema de RES com acesso telemático a diagnósticos por imagens digitais, com foco ao cuidado de pacientes em estado gravemente traumáticos. O projeto utilizou múltiplos padrões (openEHR, DICOM, HL7 CDA, HL7 PA, IHE-PDQ e CID10) que atende diferentes serviços de software. No entanto usa o padrão openEHR como núcleo do sistema de RES representando os conceitos clínicos por meio de arquétipos. Ainda propõem um framework de interface de sistema que implementa os modelos de referência do openEHR e compõem as telas através de *templates* e arquétipos. Em especial, o framework desenvolvido nesse trabalho, o openEHR-Gen-

Framework, se mostra como um importante avanço na adoção do padrão openEHR, pois acelera o processo de desenvolvimento de forma muito significativa, a partir do momento que os arquétipos estão definidos. Na Figura 4.12, pode ser visto uma tela de sistema gerada automaticamente a partir de arquétipos e templates definidos na aplicação de Serviço de Emergência.

The screenshot displays a web-based medical interface. At the top, a header bar contains the text 'TRAUMAGEN | Bienvenida: Marta Stewart' on the left, the date and time '2010-04-22 06:11:52' in the center, and language options 'EN' and 'ES' on the right, with a 'salir' button. Below the header, the main content area is divided into two primary sections: 'Resumen de episodio' (Episode Summary) and 'Paciente' (Patient). The 'Resumen de episodio' section includes a start time 'comienzo: 2010-04-22 17:16:00', an observation field 'observaciones: asdfasfdafdsfa', and a red 'T' button. Below this is a navigation bar with tabs for '(A) Vía aérea', '(B) Columna', '(B) Ventilación', '(C) Estado circulatorio', and '(D) Disfunción neurológica'. The main form is split into two columns. The left column, titled 'Evaluación de vía aérea', contains several sub-sections: 'Presenta alteraciones:' with a dropdown menu; 'Trauma sobre vía aérea:' with a sub-section 'Trauma directo sobre vía aérea:' and a dropdown; 'Tipo de trauma:' with a dropdown; 'Dificultad respiratoria:' with a dropdown; and 'Obstrucción de vía aérea:' with a dropdown. The right column, titled 'Acciones sobre vía aérea', contains: 'Máscara de oxígeno:' with a dropdown; 'Apoyo ventilatorio (BVM):' with a dropdown; 'Maniobra de desobstrucción:' with a large text input area; 'Intubación orotraqueal:' with a dropdown; and 'Protocolo de intubación:' with a dropdown. To the right of the main form is a 'Paciente' sidebar with the text 'Paciente no identificado: [identificar paciente](#)' and a list of menu items: 'Lista de episodios', 'Episodio actual', 'Registro clínico', 'Estudios img. (-)', 'Triage (+)', 'Prehospitalario', 'Resumen clínico', 'Evaluación primaria' (highlighted in blue), 'Indicación de paraclínica', 'Evaluación secundaria', 'Diagnósticos', and 'Mover paciente'. At the bottom of the page, a footer reads 'Instituto de Computación, Facultad de Ingeniería, Universidad de la República'.

Figura 4.12: Tela gerado pelo openEHR-Gen-Framework.

Evangelisti et al. [34] apresentam um trabalho de desenvolvimento, no Hospital Sírio Libanês, de um sistema de RES, interoperável, com foco no uso de padrões nacionais e internacionais. O hospital foi motivado, em especial, por custo-eficiência da atenção a saúde, gerenciamento de risco e controle de erros evitáveis, controle de qualidade nos processos assistenciais e de apoio, interoperabilidade com RES de outras instituições. Utilizou o padrão OpenEHR no processo de desenvolvimento do sistema, apoiado na modelagem de dois níveis, aponta como benefícios o engajamento do corpo clínico, a padronização dos conceitos clínicos entre os especialistas de saúde do hospital, e o reflexo disso como incentivo à adoção de diretrizes clínicas. Estruturou e desenvolveu mais de 400 elementos de arquétipos (*element*) organizados em 9 seções (*section*) e 9 telas (*templates*), infelizmente não disponíveis para consulta. Coloca como principais desafios o processo de sistematizar e garantir o processo de especificação de arquétipos (custo/tempo) e a dificuldade de se definir o escopo ideal.

Bisbal et al. [13] apresentam um trabalho sobre mediação semântica usando arquétipos e como os arquétipos oferecem uma estrutura incremental para esses serviços de dados. A idéia apresentado nesse trabalho é essencialmente o uso dos elementos de vinculação dos arquétipos com as terminologias e classificações. Esse trabalho ainda traz elementos na discussão que mostra a potencialidade dos arquétipos na construção incremental de representação de conhecimento. E faz uma comparação com os mecanismos de anotação

de bancos de dados em relação aos mecanismos de *binding* dos arquétipos. Conforme visto na Figura 4.2.

Santos [71] apresenta um trabalho de desenvolvimento de arquétipos usando o padrão ISO 13606 para compor um mecanismo de troca de informação entre SIS. Esse trabalho se dá no contexto do estado de Minas Gerais e se propõem a construir um modelo de conhecimento sumarizado, ao que o autor chama de Sumário Clínico do Paciente, ou seja, se percebe que em todo o processo se teve a preocupação de garantir que o modelo ficasse reduzido e com isso produzir um modelo viável do ponto de vista de armazenamento e transporte de informação entre SIS da Atenção Básica. Nesse trabalho foram gerados 36 arquétipos de entrada e um arquétipo de composição. A grande contribuição desse trabalho foi a riqueza de detalhes na descrição do processo de desenvolvimento dos arquétipos e dos mecanismos de transporte e persistência das informações.

Neste foi apresentada a proposta de desenvolvimento de arquétipos, bem como trabalhos correlatos que apontam o uso de arquétipos e os benefícios advindos desta abordagem.

Capítulo 5

Prova de Conceito usando SISVAN

O SIS responsável por fazer vigilância alimentar e nutricional no SUS é o Sistema de Vigilância Alimentar e Nutricional (SISVAN), ligado à Coordenação Geral de Alimentação e Nutrição (CGAN) do Departamento de Atenção Básica (DAB/SAS) do MS, é um sistema que utiliza dados individualizados de acompanhamento clínico, portanto é minimamente compatível com um sistema de RES. Tem uma estrutura de informação de registro clínico mais simplificada, frente às outras sub-áreas da atenção básica.

Atualmente o SISVAN recebe por ano cerca de 15 milhões de registros de acompanhamento nutricional e consumo alimentar da população brasileira, em particular de crianças e gestantes. E ainda é integrado com os dados de acompanhamento de mais de 13 milhões de famílias do Programa Bolsa Família¹ o que lhe confere uma boa massa de informações.

Dadas características apresentam esse sistema como uma boa opção para se fazer um estudo de caso, tanto pela sua relativa simplicidade quanto pelo conjuntos de informações que nós permite, futuramente, avaliar melhor a proposta deste trabalho em testes de escala.

5.1 Seleção dos Elementos

O SISVAN está implementado em linguagem estruturada e o modelo de dados é apoiado por um banco entidade-relacional. Apesar de ter, conceitualmente, usado as metodologia propostas pela Engenharia de Software, a documentação relacionada ao desenvolvimento do sistema está desatualizada, sem casos de uso ou regras de negócios bem definidas. Logo todo o conhecimento das regras de negócios e conceitos utilizados no processo de desenvolvimento estão centrados na experiência dos desenvolvedores responsáveis pelo sistema e na estrutura do banco de dados. Os conceitos utilizados no desenvolvimento do sistemas estão baseado na norma técnica de *Orientações para a coleta e análise de dados antropométricos em serviços de saúde* do SISVAN publicado no ano de 2004 pelo Ministério da Saúde e revisado em 2011, no entanto, algumas das orientações ainda não foram consolidadas no software do SISVAN. Também foi utilizado o documento de Protocolos do Sistema de Vigilância Alimentar e Nutricional - SISVAN na assistência à saúde [15].

Logo a etapa de seleção dos elementos foi feita utilizando seis instrumentos:

¹O Programa Bolsa Família foi criado para apoiar as famílias mais pobres e garantir a elas o direito à alimentação e o acesso à educação e à saúde. O programa visa a inclusão social dessa faixa da população brasileira, por meio da transferência de renda e da garantia de acesso a serviços essenciais.

- Orientações para a coleta de dados do SISVAN [16];
- Protocolos do SISVAN [15];
- Formulário de coleta de dados;
- Documentos de definição do banco de dados;
- Entrevistas com técnicos da CGAN/DAB.
- Entrevistas com analistas de sistemas.

O primeiro passo foi analisar o formulário de coleta de dados, junto com os documento de orientações para coleta e o documento de protocolos do SISVAN, a fim de perceber os dados que alimentam o sistema, destacando os dados essenciais de acompanhamento nutricional. A Figura 5.1 é um recorte do Formulário (Apêndice B) com os conceitos essenciais para o acompanhamento nutricional. Pode-se perceber que o formulário é organizado em faixa etária (ciclo de vida) da população, considerando informações diferentes para cada faixa e ainda um conjunto especial de informações para acompanhamento da mulher gestante. No formulário ainda é possível registrar doenças, deficiências e intercorrências associado a pessoa acompanhada de forma a verificar se existe algo que esteja relacionada a perda/ganho de peso e que afete sua avaliação nutricional.

ACOMPANHAMENTO NUTRICIONAL**						
Data do acompanhamento*: / /						
Criança (< 10 anos)	Peso (em kg)*:	Altura (em cm)*:	Estado nutricional:		Tipo de alimentação ^(d) ***:	Peso ao nascer (em gramas):
			Peso por idade:	Altura por idade:		
Adolescente (≥ 10 e < 20 anos)	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:		Altura por idade:	IMC por idade:
Adulto (≥ 20 e < 60 anos)	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:	Circunferência da cintura (em cm):	Risco aumentado: Sim Não	
Idoso (≥ 60 anos)	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:			
Gestante	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:	Peso pré-gestacional (em kg):	Data da última menstruação*: / /	
Doenças*:		Deficiências e/ou intercorrências*:			Tipo de Acompanhamento*:	
Anemia falciforme Diabetes mellitus Doenças cardiovasculares Hipertensão Arterial Sistêmica Osteoporose Outras doenças Sem doenças		Anemia ferropriva DDI (Distúrbio por Deficiência de Iodo) Diarréia Infecções intestinais virais IRA (Infecção Respiratória Aguda) Hipovitaminose A Outras deficiências e/ou intercorrências Sem deficiências e/ou intercorrências			Atendimento na Atenção Básica Chamada Nutricional Saúde na Escola _____	

Figura 5.1: Recorte do Formulário do SISVAN

O segundo passo foi analisar a documentação do banco de dados com o apoio do analista de sistema responsável pelo SISVAN, a fim de capturar as estruturas usadas para representar os conceitos de acompanhamento nutricional e da identificação do registro através do paciente, do profissional e da unidade responsável pelo acompanhamento. Na Figura 5.2 pode ser visualizado um recorte do diagrama entidade-relacional das tabelas que definem a estrutura dos dados de acompanhamento nutricional.

Analisando o recorte das tabelas, com 82 variáveis, foi constatado que apenas 37 dessas variáveis são realmente significativas para a descrição dos conceitos, as outras variáveis estão distribuídas entre variáveis de controle dos registros e variáveis calculadas que buscam melhorar o desempenho nas consultas de séries históricas dos dados. As

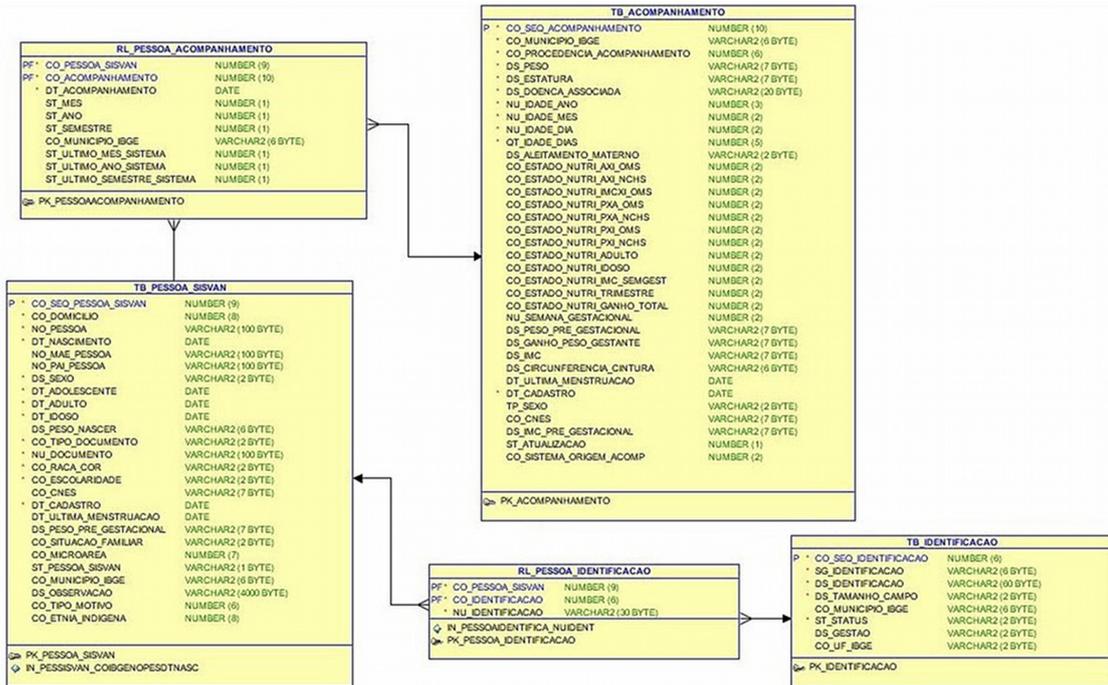


Figura 5.2: Recorte Banco de Dados do SISVAN

variáveis destacadas podem ser vistas na Tabela 5.1 e estão divididas em três grupos de informações:

- Registro Demográfico (localização do registro),
- Registro da Pessoa Acompanhada, e
- Registro de Saúde Geral e de Gestante.

Os grupos dividem as informações nos registros demográficos, cujas variáveis identificam o local onde as informações foram coletadas, os registros da pessoa acompanhada, os quais identificam a pessoa e também agregam informações demográficas que apoiam as análises estatísticas dos dados, e por último os registros de saúde das pessoas acompanhadas, com algumas informações específicas para gestante.

5.2 Identificação dos conceitos candidatos

Os conceitos candidatos vinculados ao sub-domínio da Atenção à Saúde, em particular os dados de acompanhamento nutricional vinculados às variáveis do SISVAN, são os que registram os dados de saúde, ou seja, os conceitos identificados, na Tabela 5.1, pelos grupos: **Registro** e **Gestante**.

As outras variáveis (Demográfico e Pessoa) registram os dados demográficos do acompanhamento nutricional, isto é, os dados de identificação do paciente, do profissional de

Tabela 5.1: Tabela de variáveis do SISVAN

ID	Grupo	Descrição	ID	Grupo	Descrição
1	Demográf.	Código do CNES	20	Registro	Altura
2	Demográf.	Código do Município	21	Registro	Circunfer. da Cintura
3	Demográf.	Código da Micro-área	22	Registro	Peso
4	Demográf.	Código do Domicílio	23	Registro	Peso ao Nascer
5	Pessoa	Código da Pessoa	24	Registro	Estado Nutricional (EN) do Adulto
6	Pessoa	Escolaridade	25	Registro	EN Altura por Idade
7	Pessoa	Etnia Indígena	26	Registro	EN do Idoso
8	Pessoa	Nome do Paciente	27	Registro	EN IMC por Idade
9	Pessoa	Nome da Mãe	28	Registro	EN Peso por Altura
10	Pessoa	Nome do Pai	29	Registro	EN Peso por Idade
11	Pessoa	Raça / Cor	30	Registro	Doenças Associadas
12	Pessoa	Sexo	31	Registro	Deficiências e/ou Intercorrências
13	Pessoa	Código de Identificação	32	Gestante	Data da Última Menstruação
14	Pessoa	Sigla da Identificação	33	Gestante	Peso Pré-gestacional
15	Pessoa	Número da Identificação	34	Gestante	IMC por Semana Gestacional
16	Pessoa	Situação Familiar/Conjugal	35	Gestante	Ganho de Peso Trimestre
17	Registro	Procedência do Acompanhamento	36	Gestante	Ganho de Peso Total Durante a gestação
18	Registro	Data do Acompanhamento	37	Gestante	Aleitamento Materno
19	Registro	Data de Nascimento			

saúde² e do estabelecimento são menos relevantes para nosso estudo de caso, portanto não serão utilizadas no modelo. Para mais informações, sobre questões relacionadas aos dados demográficos, o trabalho de Dias and Freire [30] pode ser consultado.

5.2.1 Analisando as variáveis

A variável (17) de “Procedência do Acompanhamento” foi criada com base nas necessidades de gestão de políticas públicas no Brasil, sendo definida pela CGAN para identificar a origem da informação de acompanhamento. Esta variável registra no instrumento de coleta as seguintes informações:

- Atendimento na Atenção Básica: Atendimento ambulatorial de acompanhamento do paciente na rede de atenção básica à saúde;
- Chamada Nutricional: Investigação nutricional de uma determinada população ou território;
- Saúde na Escola: Programa de triagem de alunos em escolas municipais e estaduais.

²Apesar de não termos uma vinculação direta ao profissional, podemos inferir a informação do Agente Comunitário de Saúde e da Equipe de Atenção Básica associadas ao registro por meio da informação da Micro-área.

Apesar destas informações serem relevantes para a gestão da VAN no Brasil, não são informações de saúde, portanto acreditamos que não deveriam estar organizadas dessa forma. Isto é, estas são informações auxiliares que podem vir por meio das informações do estabelecimento de saúde ou do tipo de registro.

A procedência do registro de acompanhamento é uma informação que está associada a um registro de saúde e portanto é representado pelos modelos de referência do openEHR, porém por informações auxiliares dos modelos de referência. Da mesma forma a variável (18) "Data de Acompanhamento" é contemplado pela própria estrutura do registro eletrônico como "Data do Registro".

Ainda analisando o conjunto de variáveis de registros de saúde, é possível perceber que a variável (33) "Peso pré-gestacional" identifica o "Peso" de um paciente do "Sexo" feminino anterior ao tempo de uma gestação, estimado pela "Data da Última Menstruação" (DUM). Da mesma forma a variável (23) "Peso ao nascer" é o "Peso" de um paciente na "Data de Nascimento", porém a informação é coletada em gramas (g) e não em quilograma (kg) como na variável relacionada o conceito de "Peso".

As especificações do openEHR já contempla as duas situações apresentadas no parágrafo anterior. Uma situação é tratada pelo conceito de evento, ou seja, o "Peso ao nascer" ou o "Peso pré-gestacional" é o "Peso" do paciente no evento de nascimento e no instante pré-gestacional, respectivamente, localizados no tempo pelas datas de nascimento e DUM. A outra situação identificada é em relação as unidades de medida adequadas ao contexto da coleta da informação, representadas nas especificações do openEHR por estruturas de *Data Types* distintos[10].

5.2.2 Analisando as orientações do SISVAN

Nas orientações do SISVAN [16], além dos conceitos associados às variáveis do banco de dados, ainda são discutidos os conceitos relacionados aos índices antropométricos e os parâmetros para VAN, definidos pela Organização Mundial da Saúde (OMS) para determinar a Classificação do Estado Nutricional do paciente/usuário da saúde, sendo eles:

- Peso por Idade
- Peso por Estatura
- Estatura por Idade
- Índice de Massa Corporal (IMC) - $Peso/Estatura^2$
- IMC por idade
- IMC por semana gestacional

A classificação do estado nutricional, conceito definido nas orientações como "Diagnóstico Nutricional" ou diagnóstico antropométrico, se refere aos vários tipos de classificações realizadas a partir dos índices antropométricos, cujas classificações são apresentadas da seguinte forma:

Peso por Idade :

- Muito baixo peso para a idade
- Baixo peso para a idade
- Peso adequado para idade
- Peso elevado para idade

Peso por Estatura :

- Magreza acentuada
- Magreza
- Eutrofia
- Risco de sobrepeso
- Sobrepeso
- Obesidade

Estatura por Idade :

- Muito baixa estatura para a idade
- Baixa estatura para a idade
- Estatura adequada para a idade

IMC :

- Baixo peso
- Adequado ou Eutrófico
- Sobrepeso
- Obesidade

IMC por Idade :

- Magreza acentuada
- Magreza
- Eutrofia
- Risco de sobrepeso
- Sobrepeso
- Obesidade

IMC por Semana Gestacional :

- Baixo peso
- Adequado ou Eutrófico
- Sobrepeso
- Obesidade

Na descrição de conceitos no documento de orientação o conceito “Estado Nutricional” é definido como o resultado do equilíbrio entre o consumo de nutrientes e o gasto energético do organismo para suprir as necessidades nutricionais, sendo classificado de acordo a três tipos de manifestação orgânica:

Adequação Nutricional (Eutrofia) : manifestação produzida pelo equilíbrio entre o consumo e as necessidades nutricionais.

Carência Nutricional : situação em que deficiências gerais ou específicas de energia e nutrientes resultam na instalação de processos orgânicos adversos à saúde.

Distúrbio Nutricional : problemas relacionados ao consumo inadequado de alimentos, tanto por escassez quanto por excesso, como a desnutrição e a obesidade.

Uma questão importante a ser destacada aqui, é o fato do conceito “Estado Nutricional”, associado às variáveis de 24 a 29 e de 34 a 36, da Tabela 5.1, terem diferentes interpretações de acordo com o documento de orientações e da forma como ele é utilizado no sistema do SISVAN. No sistema do SISVAN, as classificações associadas ao conceito de “Diagnóstico Nutricional” são aplicadas ao conceito de “Estado Nutricional”, o que é simples de perceber após uma rápida leitura nas normas e orientações do SISVAN. Este conflito semântico pode trazer alguns transtornos ao serem comparados em outros contextos, pois apesar do “Estado Nutricional”, no SISVAN, derivar do “Diagnóstico Nutricional” estes são conceitos diferentes.

Por último podemos destacar o conceito de **Aleitamento Materno** que identifica a forma de alimentação da criança conforme as seguintes categorias:

- Exclusivo: quando a criança recebe somente leite materno, diretamente da mama ou extraído, e nenhum outro alimento líquido ou sólido, com exceção de gotas ou xaropes de vitaminas, minerais e/ou medicamentos;
- Predominante: quando o lactente recebe, além do leite materno, água ou bebidas à base de água, como sucos de frutas e chás;
- Alimentação Complementar: recebe, além do leite materno, alimentos sólidos e semi-sólidos, incluindo o leite não-humano;
- Não recebe leite materno: a criança já mantém uma alimentação com alimentos sólidos e semi-sólidos, sem o aleitamento materno.

Tipo de Alimentação Infantil está associada às orientações da OMS quanto a forma de alimentação das crianças menores de 10 anos. A OMS e o MS recomendam que o bebê seja amamentado exclusivamente até os 6 meses de vida. Após os 6 meses e no mínimo até os 2 anos de vida, deve-se começar a alimentação complementar com a introdução de novos alimentos, sem abandonar o aleitamento materno.

5.2.3 Conceitos Identificados

Feitas as análises, podemos destacar os conceitos identificados, relacionados com a VAN, em três grupos como segue:

Demográficos (conceitos auxiliares) :

- Sexo,
- Data de Nascimento,
- Data da Última Menstruação (DUM, ou *LMP*).

Antropométricos :

- Circunferência da Cintura,
- Altura,
- Peso,
- Peso Pré-gestacional (evento),
- Peso Ao Nascer (evento).

Gerais :

- Estado Nutricional,
- Diagnóstico Nutricional,
- Doenças Associadas a Pessoa,
- Deficiências (nutricionais),
- Intercorrências (clínica),
- Aleitamento Materno.

Os grupos de conceitos visam separar os tipos de informações usadas no modelo. O primeiro grupo são os dados demográficos, no entanto apenas aqueles conceitos auxiliares utilizados para aumentar a semântica de conceitos relacionados. O segundo grupo são os dados antropométricos usados para fazer a avaliação nutricional. E o terceiro grupo, são conceitos mais gerais. Feita a identificação dos conceitos é possível sugerir a organização destes usando uma técnica de mapa mental, como visto na Figura 5.3, para formar o conceito de Acompanhamento Nutricional.

Segundo as orientações do SISVAN [16], o acompanhamento nutricional ainda está associado às fases da vida do indivíduo acompanhado, categorizados por meio da idade, da seguintes forma:

- Criança: menor de 10 anos de idade;
- Adolescente: maior ou igual a 10 anos e menor que 20 anos de idade;
- Adulto: maior ou igual a 20 anos e menor que 60 anos de idade;
- Idoso: maior ou igual a 60 anos de idade;
- Gestante: pessoa do sexo feminino com idade maior que 10 anos e menor que 60 anos de idade.

Ou seja, existem diferentes formas de se fazer o diagnóstico nutricional de um paciente/usuário da saúde, e estas formas estão relacionadas às fases de vida do mesmo. Na Figura 5.4, podemos visualizar outro mapa mental do conceito de acompanhamento nutricional reorganizando os itens de acordo com as fases da vida, no entanto com alguns conceitos repetidos conforme sua aplicação na fase de vida.



Figura 5.3: Mapa Mental do Conceito da VAN

5.3 Pesquisa dos arquétipos existentes

Após identificar os conceitos, é necessário representá-los em um formato de arquétipos. Para arquétipos usando o padrão openEHR, ou seja, o modelo de referência utilizado nesse padrão, já existem diversos conceitos definidos, portanto é importante ver a possibilidade de reutilizar os conceitos já definidos. Os arquétipos foram pesquisados em duas fontes diferentes:

- openEHR Foundation³;
- NEHTA - National E-Health Transition Authority, Austrália⁴;

A pesquisa sobre os conceitos foi realizada nos repositório conforme resultado mostrado na Tabela 5.2. A tabela está organizada em ID, conceito, arquétipo encontrado sobre o conceito, idioma disponível para os arquétipos encontrados e para cada repositório a versão em que o arquétipo está disponível. Alguns arquétipos ainda não estão disponíveis para o idioma Português do Brasil (pt-br), para tanto será necessário fazer uma tradução/localização (*localization*) desses conceitos para ter um melhor entendimento da relação dos conceitos usados no SISVAN e para uma melhor interação com os especialista que farão a validação do modelo de informação.

É notável que existem arquétipos representando conceitos diferentes dos conceitos identificados, isso se dá pelo fato de que alguns conceitos estão relacionados e possuem elementos internos (ELEMENT e CLUSTER) do arquétipo que representam os conceitos desejados.

³www.openehr.org/ckm/, acessado em 18/12/2011

⁴<http://dcm.nehta.org.au/ckm/>, acessado em 20/12/2011

ID	Conceito	Arquétipo	Idioma	openEHR	NEHTA
01	Circunferência da Cintura	OBSERVATION waist_hip	en	Draft	–
02	Altura	OBSERVATION height	pt-br	Published	Review Suspended
03	Peso	OBSERVATION body_weight	pt-br	Published	Review Suspended
04	Peso ao Nascer	OBSERVATION body_weight-birth	en	Draft	–
05	Índice de Massa Corporea (IMC)	OBSERVATION body_mass_index	pt-br	Published	Team review
06	Data da Última Menstruação	ELEMENT date_of_lmp in EVALUATION pregnancy.v1	en	–	Draft
07	Data do primeiro dia do último ciclo menstrual	ELEMENT last_normal_menstrual _period.v1 in EVA- LUATION preg- nancy_summary.v1	en	–	Team review
08	Data do primeiro dia do último ciclo menstrual (LNMP)	ELEMENT last_normal_menstrual _period.v1	en	Draft	–
09	Alimentação (Infantil)	OBSERVATION infant_feeding	en	Draft	–
10	Aleitamento Materno	CLUSTER breast_feeding in OBSERVATION infant_feeding	en	Draft	–
11	Estado Nutricional	ELEMENT nutrition_status in EVALUATION nutrition_summary	en	–	Review Suspended
12	Estado de Peso	ELEMENT weight_status in EVALUATION nutrition_summary	en	–	Review Suspended
13	Doenças Associadas	EVALUATION problem-diagnosis	pt-br	Draft	Team review
14	Intercorrências	EVALUATION problem-diagnosis	pt-br	Draft	Team review
15	Deficiências	EVALUATION problem-diagnosis	pt-br	Draft	Team review

Tabela 5.2: Arquétipos já definidos nos repositório do openEHR e do NEHTA

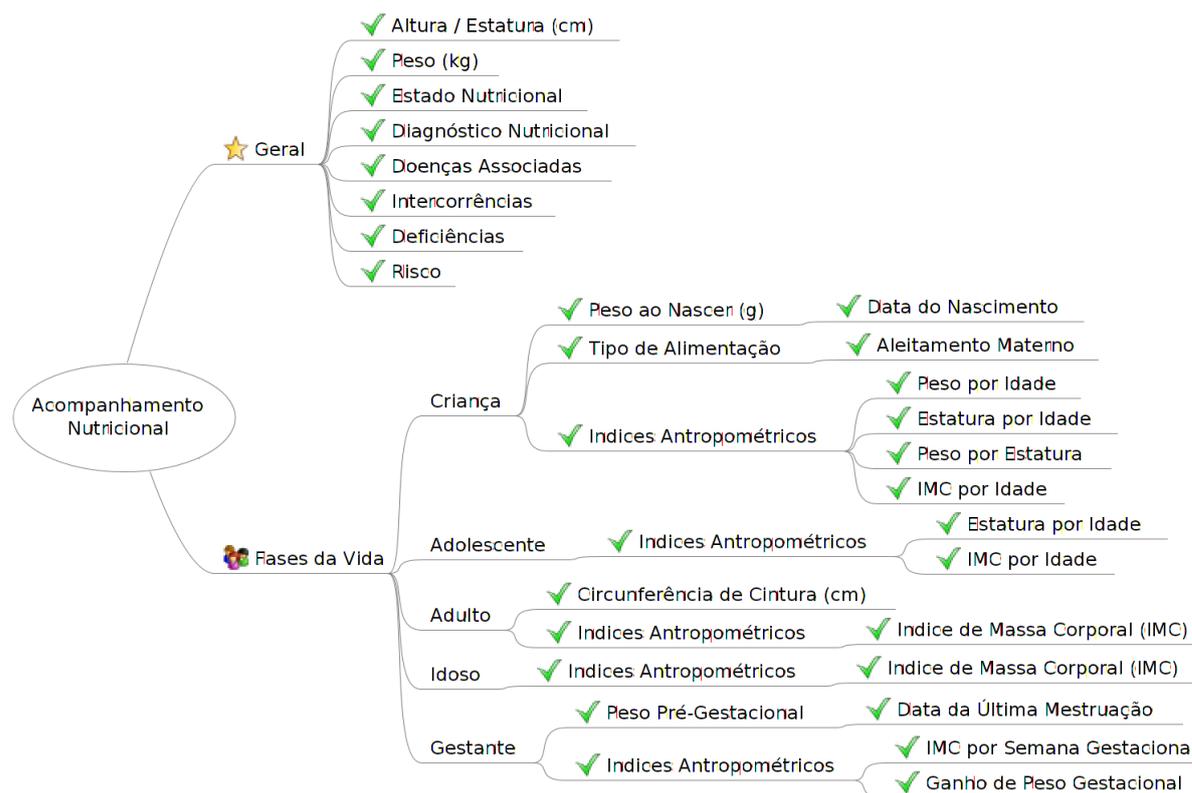


Figura 5.4: Mapa Mental dos Conceitos de VAN por Fase da Vida

5.3.1 O Problema da definição múltipla da DUM

Ao ser realizada a pesquisa foi identificado que a variável *Data da Última Menstruação* (DUM) ou *Last Menstrual Period* (LMP) estava definida em dois lugares diferentes e com descrições distintas.

A estrutura de arquétipos permite que elementos que estejam definidos dentro de algum conceito possa ser destacado (em níveis) do mesmo para poder ser reutilizado. Tendo isso em mente buscamos a variável tanto como um elemento (*ELEMENT*) definido de forma isolada quanto como um elemento organizado em subitens de um conceito mais geral. Feito isso, o conceito de DUM foi encontrado em três elementos:

- *ELEMENT.last_normal_menstrual_period.v1* (repositório openEHR);
- *ELEMENT (Last menstrual period)* dentro de *EVALUATION.pregnancy.v1* (repositório openEHR);
- *ELEMENT (Last normal menstrual period)* dentro de *EVALUATION.pregnancy_summary.v1* (repositório NEHTA).

Avaliando o histórico dos arquétipos se percebe que o último está em constantes revisões pela equipe do NEHTA e revisado por um grande grupo de pesquisadores do mundo, portanto parece ser o conceito mais maduro entre os três e será o escolhido. É importante ressaltar que apesar dos conceitos do primeiro e do terceiro item terem o mesmo nome não há nada que faça o link entre eles, logo não deveriam ser considerados o mesmo conceito, no entanto acreditamos que a intensão do autor foi essa.

5.3.2 Generalização dos Conceito de Estado de Peso e IMC

O conceito de “Estado de Peso” é um conceito que representa uma classificação do estado de peso de acordo com uma avaliação do IMC do paciente. O SISVAN usa esse conceito de forma mais geral, o conceito “Diagnóstico Nutricional” que representa a variação em uma classificação de acordo com uma avaliação usando algum índice antropométrico, de acordo com a fase de vida do paciente. Ou seja, ao tratar o conceito “Estado de Peso” (*Weight Status*), representado pelo arquétipo `ELEMENT.weight_status`, deve ser usado o conceito mais geral de “Diagnóstico Nutricional” associado a avaliação do estado de peso do paciente através do IMC.

O mesmo ocorre com o conceito IMC, é necessário fazer uma generalização do termo para um termo mais amplo que corresponde ao conceito de “Índice Antropométrico”.

Essa diferenciação ocorre, pois segundo as definições da OMS, a avaliação do estado de peso usando apenas o IMC não é precisa ao se avaliar crianças e adolescentes.

5.4 Análise e identificação de terminologias

A identificação das terminologias no domínio é fundamental na vinculação dos conceitos definidos pelos modelos de informação (arquétipos) com os conceitos definidos por essas terminologias e por outras formas similares de representação de conhecimento.

A associação com os termos podem ocorrer de duas formas, por definição (*term_definitions*) quando a lista de termos é definida dentro do próprio arquétipo, e por vinculação (*term_binding*) quando o termo é vinculado a uma lista formal.

Exemplo de lista de termo já representadas e consolidadas na área de saúde é a CID, atualmente utilizada na versão 10 (CID10). Proposta pela OMS, como visto na Seção 3.2.1, e construída com o apoio de vários países, já é bastante utilizada e consolidada nos SIS e na prática clínica dos profissionais de saúde. Da mesma forma a CIAP, atualmente utilizada na versão 2 (CIAP2), é amplamente utilizada em países como Portugal, Austrália e Holanda, no entanto não está consolidada nos SIS e na prática clínica dos profissionais de saúde no Brasil[46].

Observando a estrutura de registro de acompanhamento do SISVAN, podemos perceber, além dos conceitos a serem representados pelos arquétipos, as seguintes informações podem ser vinculadas às classificações CID e/ou CIAP:

- Doenças Associadas (condições crônicas);
- Deficiências (nutricionais);
- Intercorrências (associadas no atendimento);

O recorte de Doenças Associadas, Deficiências e Intercorrências, consideradas pelo SISVAN, podem ser vinculadas através do CID ou CIAP da seguinte forma:

Doenças Associadas • Anemia falciforme (CID D57.0 e D57.1)

- Diabetes mellitus (CID E14)
- Doenças cardiovasculares (CID I51.6)
- Hipertensão Arterial Sistêmica (CID I10 e I11)

- Osteoporose (CID M819)
- Outras doenças (CIAP A99)
- Sem doenças (CIAP A97)

Deficiências (CIAP T91 - Deficiência vitamínica/nutricional)

- Anemia ferropriva (CID D50)
- DDI (Distúrbio por Deficiência de Iodo) (CID E00X e E01X)
- Hipovitaminose A (E50.9)

Intercorrências • Diarreia (CIAP D11)

- Infecções intestinais virais (CID A084)
- IRA (Infecção Respiratória Aguda) (CID J00 a J22)

Todas as listas de termos, classificações e terminologias utilizadas nos extratos de RES, segundo a norma ISO 13606, precisam ser identificadas por códigos *Object Identifier* (OID). Isso objetiva garantir que tanto o sistema que enviou o extrato de dados, quanto o que recebeu, identifiquem corretamente a terminologia em uso. O controle desses números é realizado por entidades de registro semelhantes a da Internet [71].

A identificação das terminologias, considerando o contexto do Ministério da Saúde, devem considerar dois blocos de referência de termos, o primeiro bloco é o Catálogo de Padrões de Interoperabilidade definido na portaria de interoperabilidade, e o outro bloco que define um conjunto de tabelas de domínio conhecido como Repositório de Tabelas Nacionais definida pelo DATASUS em 2004, como reflexo das discussões da PNIIS.

O Brasil, apesar das iniciativas de 2004 na definição de um repositório de tabelas de domínio e da portaria de interoperabilidade, ainda não definiu uma entidade/local que faça o controle desses vocabulários. Regionalmente, no estado de Minas Gerais, está sendo formalizado um repositório de terminologias por meio da Base de Registro Eletrônico em Saúde do Estado de Minas Gerais.

No cenário internacional a organização *HL7 International* mantém uma lista⁵ de OID para saúde com mais de 350 objetos identificados, sendo esta lista composta por terminologias, classificações internacionais, lista de termos, dicionários, catálogos, e ainda subitens que fazem referência a alguns desses objetos traduzidos a outros idiomas. Na lista contemplada pela organização podem ser encontrados todas as terminologias e classificações que constam no CPIIS, inclusive o LOINC e o SNOMED-CT, que ainda não estão traduzidos para o português.

5.5 Modelagem dos Arquétipos

O Acompanhamento Nutricional é um conceito de mais alto nível, logo, como visto na Seção 3.1.2, pode ser composto por vários níveis de conceitos. Este ainda é feito de forma diferente de acordo com a fase da vida do indivíduo. Portanto, como podemos ver na Figura 5.5, usaremos um arquétipo de composição (*COMPOSITION*) para representar o

⁵A lista de OID para saúde organizadas pela HL7 Internacional pode ser visitada através do endereço eletrônico: <http://www.hl7.org/oid/>

conceito mais geral e especializações (*specialize*) desse arquétipo para estender o conceito às outras fases da vida. Para melhor organizar os conceitos e também para ficar mais didático foi proposta uma divisão nas seções (*SECTION*) Observação e Avaliação, que correspondem respectivamente às classes de entrada *OBSERVATION* e *EVALUATION*.

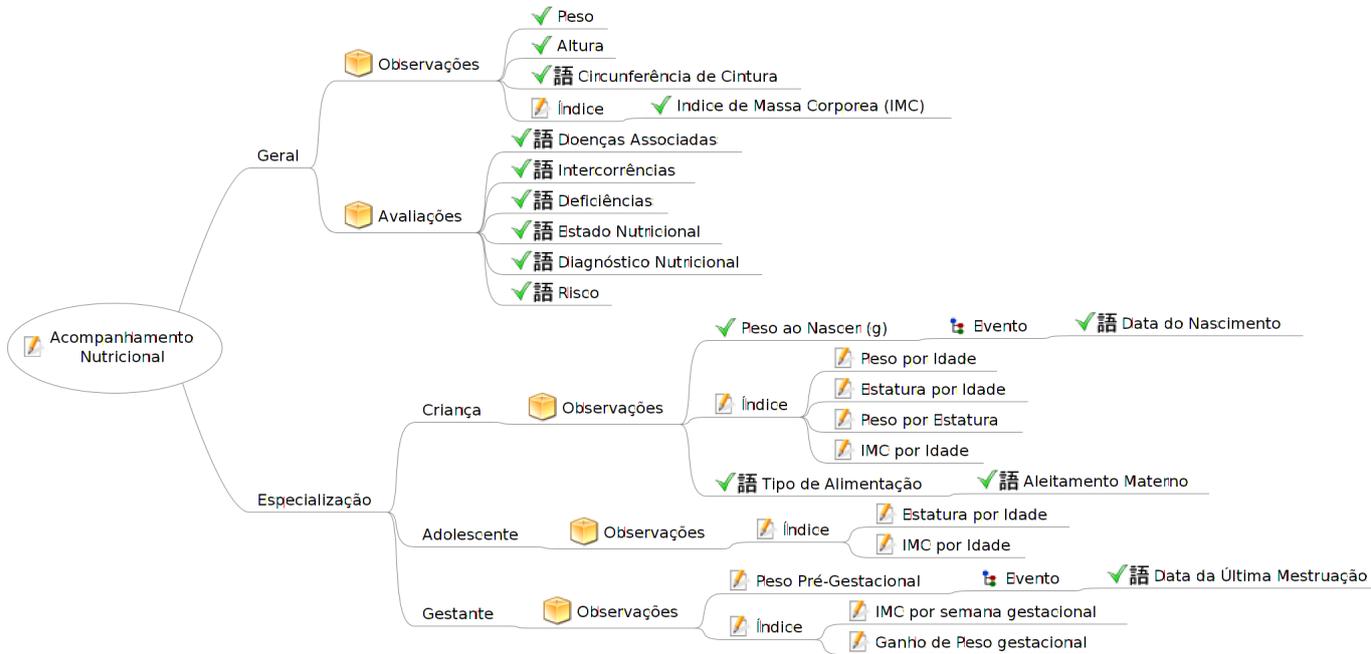


Figura 5.5: Mapa Mental de Acompanhamento Nutricional no padrão openEHR

Na Figura 5.5 podemos visualizar algumas marcas no mapa mental que indicam o estado dos arquétipos em relação às necessidades de codificação, note as diferentes representações:

- sinal de checado (em verde): conceitos que foram encontrados na pesquisa de repositório;
- sinal de linguagem (caractere chinês): conceitos encontrados que precisam ser traduzidos;
- sinal de lápis: conceitos que devem ser criados;

5.6 Codificação de arquétipos em ADL

A etapa de codificação, neste estudo de caso, se resume a duas ações: traduzir os arquétipos que não têm tradução para o português (pt_BR) e implementar os conceitos que não foram possíveis de encontrar nos repositórios disponíveis.

5.6.1 Ferramentas para edição de ADL

A linguagem ADL, atualmente em sua versão 1.4, é uma linguagem de texto puro e portanto se pode utilizar um editor simples de texto para codificar os conceitos em ADL. No entanto já existem editores que auxiliam no processo de edição dos arquétipos a depender, em especial, do modelo de referência escolhido para tal tarefa.

Em uma busca pela internet foram encontrados três ferramentas:

Ocean Archetype Editor⁶: um editor desenvolvido pela empresa australiana Ocean Informatics ligada a University College London (UK). Atualmente na versão 2.2 (Maio de 2011), é uma ferramenta mono-plataforma Windows. Oferece suporte nativo ao padrão openEHR e ao padrão CEN/ISO 13606.

LiU Archetype Editor⁷: um editor open-source desenvolvido pelo Department of Biomedical Engineering, Linköping University (LiU) da Suécia, em java, suporta o padrão openEHR. Atualmente se encontra na versão (0.5.2 Março de 2007).

LinkEHR Normalization Platform⁸: uma plataforma open-source desenvolvida pelo Biomedical Informatics Group (IBIME) dentro do Institute for Applications of Advanced Information and Communication Technologies (ITACA Institute) localizado na Universidade Politécnica de Valência na Espanha. Desenvolvido em java multi-plataforma, suporta os padrões CEN EN13606, openEHR, CCR, HL7 CDA e CDISC ODM. Atualmente está na versão RC4 (Novembro 2011).

Em uma rápida avaliação, considerando interface gráfica e funcionalidades, foi constatado que o editor LinkEHR apresenta melhores características que os outros. Funcionalidades que chamam a atenção e auxiliam no processo de construção dos arquétipos, são elas:

- pode-se escolher diferentes modelos de referência;
- é multi-plataforma;
- documentação clara e com diversos recurso de apoio, como tutoriais e vídeos;
- implementa um mecanismo de busca que localiza arquétipos definidos em repositórios externos ligados a internet, sem necessidade de prévia configuração;
- gera arquivos de mapa mental em formato aberto para FreeMind⁹ ou em XML para outros editores;
- tem ferramenta de transformação de linguagem, por exemplo, ADL para OWL e OWL para ADL;
- entre outras.

5.6.2 Arquétipos modelados

Usando a ferramenta LinkEHR, foram codificados os conceitos que não foram encontrados e traduzidos os conceitos que ainda não tinham tradução para o português do Brasil. Na Figura 5.6, podemos visualizar o conceito de Acompanhamento Nutricional (*Nutritional Monitoring*) sendo codificado fazendo uso da ferramenta LinkEHR, note que os itens dentro das seções Observação e Avaliação estão sendo usando referenciados pela função de *Archetype Slot (include)*, ou seja, inclui conceitos dentro do arquétipo definido em outro arquétipo de mais baixo nível.

⁶<https://wiki.oceaninformatics.com/confluence/display/TTL/Archetype+Editor>

⁷<http://www.imt.liu.se/mi/ehr/tools/>

⁸<http://www.linkehr.com/>

⁹FreeMind é um software livre capaz de editar diagramas usando os conceitos de mapa mental.

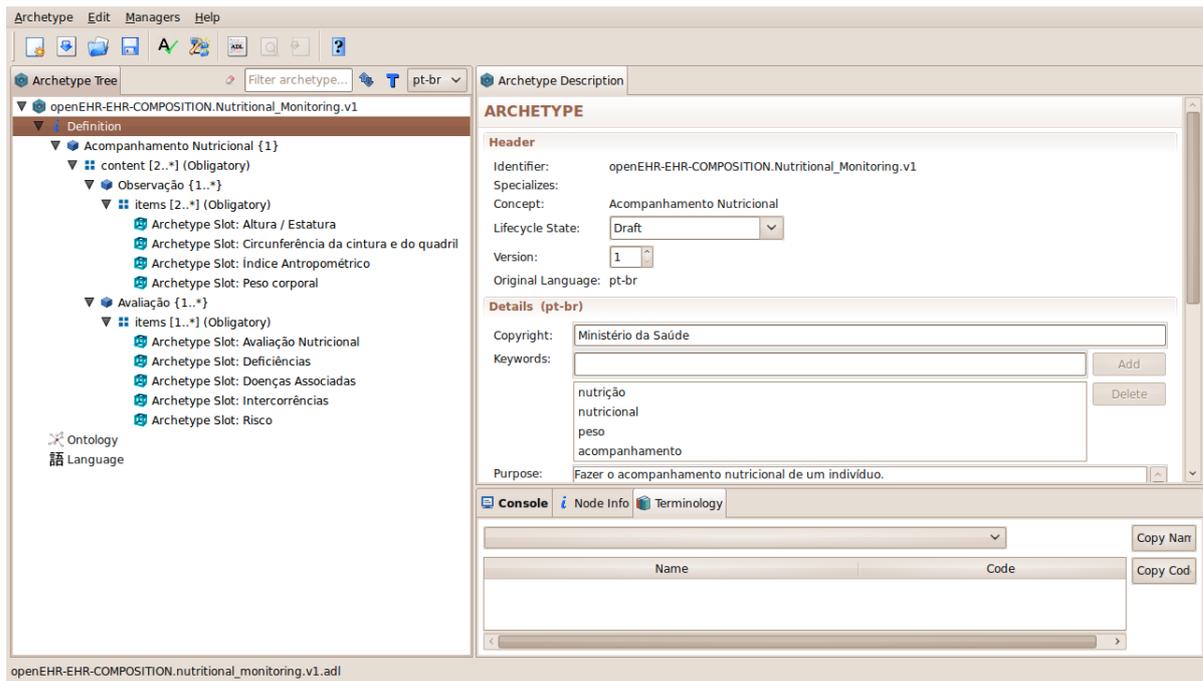


Figura 5.6: Acompanhamento Nutricional no LinkEHR

O uso da função de *include* da ADL, busca garantir maior reutilização dos conceitos, porém os arquétipos poderiam ter sido salvos ou criados dentro da própria estrutura do arquétipo.

Considerando as classes de entrada do padrão openEHR, percebe-se que foram utilizadas apenas as classes *OBSERVATION* e *EVALUATION*, as outras duas classes *INSTRUCTION* e *ACTION* não foram utilizadas neste modelo para dar mais simplicidade ao próprio, no entanto existem instruções e ações referidas a partir da avaliação de risco (*risk*) do indivíduo que poderiam ter sido modeladas, como por exemplo, instruções de suplementação de alguma vitamina em caso de uma deficiências nutricionais ou um exame (ação) de lipidograma completo para melhor diagnosticar o estado nutricional de um paciente caso o mesmo esteja obeso [15].

Os arquétipos reutilizados ou codificados em ADL, usando o modelo de referência do openEHR-EHR, podem ser vistos na lista que segue (conceito:arquétipo):

- Acompanhamento Nutricional (AN):
openEHR-EHR-COMPOSITION.nutritional_monitoring.v1
- AN para Gestante:
openEHR-EHR-COMPOSITION.nutritional_monitoring-pregnancy.v1
- AN para Adolescente:
openEHR-EHR-COMPOSITION.nutritional_monitoring-adolescent.v1
- AN para Criança:
openEHR-EHR-COMPOSITION.nutritional_monitoring-child.v1
- Índice Antropométrico (IA):
openEHR-EHR-OBSERVATION.anthropometric_index.v1

- IA Índice de Massa Corporal:
openEHR-EHR-OBSERVATION.anthropometric_index-body_mass_index.v1
- IA Peso por Altura:
openEHR-EHR-OBSERVATION.anthropometric_index-weight_per_height.v1
- IA Estatura por Idade:
openEHR-EHR-OBSERVATION.anthropometric_index-height_per_age.v1
- IA Peso por Idade:
openEHR-EHR-OBSERVATION.anthropometric_index-weight_per_age.v1
- IA IMC por Semana Gestacional:
openEHR-EHR-OBSERVATION.anthropometric_index-IMC_per_gestation_week.v1
- IA IMC por Idade:
openEHR-EHR-OBSERVATION.anthropometric_index-IMC_per_age.v1
- Alimentação Infantil:
openEHR-EHR-OBSERVATION.infant_feeding.v1
- Altura/Estatura:
openEHR-EHR-OBSERVATION.height.v1
- Peso Corporal:
openEHR-EHR-OBSERVATION.body_weight.v1
- Peso ao Nascer:
openEHR-EHR-OBSERVATION.body_weight-birth.v1
- Circunferência da Cintura e Quadril:
openEHR-EHR-OBSERVATION.waist_hip.v1
- Risco:
openEHR-EHR-EVALUATION.risk.v1
- Diagnóstico Nutricional:
openEHR-EHR-EVALUATION.nutrition_evaluation.v1
- Diagnóstico:
openEHR-EHR-EVALUATION.problem-diagnosis.v1
- Data da Última Menstruação:
openEHR-EHR-ELEMENT.last_normal_menstrual_period.v1

Perceba que sintaxe de nomeação dos arquétipos já indicam qual é o modelo de referência e qual é a classe pai do arquétipo definido. Também é importante destacar a função do hífen (-) na nomeação, que representa uma especialização de um conceito, por exemplo, o conceito de avaliação “Diagnóstico” é uma especialização de um “Problema de saúde” (EVALUATION.problem-diagnosis). A última parte do nome do arquétipo indica a versão desse arquétipo.

No Apêndice C pode ser visto o arquétipo definido em linguagem ADL para o conceito geral de Acompanhamento Nutricional (*Nutritional Monitoring*), conforme visto no item geral do mapa mental da Figura 5.5.

5.7 Análise dos Resultados

Como análise do arquétipo, foi feito um extrato de informação usando informações reais do sistema SISVAN. O resultado, extrato de dados produzido em XML, é apresentado no Apêndice E.

O extrato de dados deve respeitar a sintaxe do modelo de referência do padrão openEHR, logo existem esquemas¹⁰ (XSD) que restringem sintaticamente o modelo. Existem seis esquemas definidos pelas especificações do padrão openEHR para validar sintaticamente os extratos:

- *Basic types* - valida os tipos básicos (rm.data_types e rm.support.identification);
- *Structure* - valida os tipos para estruturas (rm.data_structures e rm.common.generic);
- *Resource* - valida os tipos para recursos (rm.common.resource);
- *Content* - valida *Item*, *Section*, *Entry* e e outros sub-tipos;
- *Composition* - valida a classe *Composition* (rm.composition);
- *Version* - valida a sintaxe das estruturas de controle de alterações (rm.common.change_control).

Na Figura 5.7 [7] pode ser visto um exemplo de relação das classes do padrão openEHR na criação de extratos de informação.

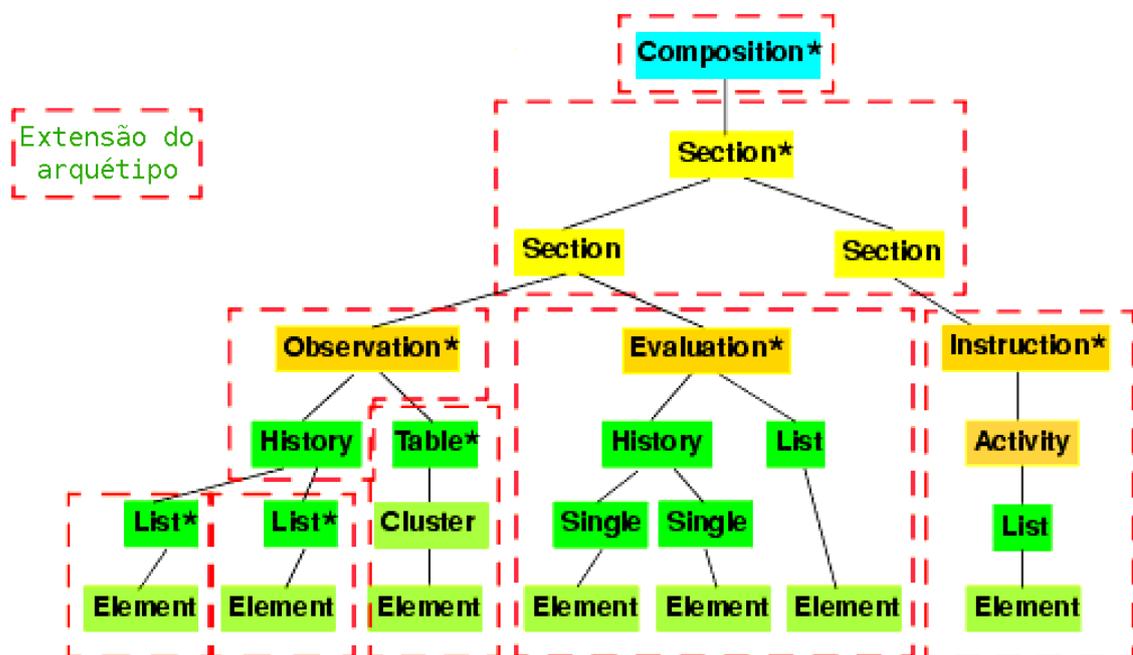


Figura 5.7: “Arquetipando” informação

A relação dos arquétipos de nível maior com os de nível menor pode ser feita através de funções de *slot* para fazer o detalhamento dos conceitos, como visto na Seção 5.6.2. Por

¹⁰Os esquemas mencionados podem ser encontrados em <http://www.openehr.org/releases/1.0.2/its/XML-schem>

exemplo, no Código 5.1 pode ser visto o conceito de estatura (*height*) incluído (*include*) dentro do arquétipo geral que representa o conceito de Acompanhamento Nutricional, apresentado por completo no Código C.1 do Apêndice C.

Código 5.1: Incluindo via slot o conceito de Estatura

```

33 ...
34   allow_archetype OBSERVATION[at0001]
35   occurrences matches {1..1}
36   matches {
37     include
38     archetype_id/value matches {
39       /openEHR-EHR-OBSERVATION.height.v1/
40     }
41     exclude
42     archetype_id/value matches {/.*}/
43   }
44 ...

```

O arquétipo que define o conceito de Estatura pode ser visto no Apêndice D, onde pode-se perceber que as funções de *include* ao mesmo tempo que "esconde" a complexidade do outro conceito, ainda permite que não apenas o conceito pai seja incluído, mas também os conceitos que estendem esse conceito. Para permitir esse tipo de construção em níveis, os arquétipos usam expressões regulares nos nomes dos arquivos dos arquétipos. Um exemplo desse tipo de construção pode ser visto no Código 5.2, linha 95, pela inclusão do conceito Dispositivo (*device*).

Código 5.2: Incluindo via slot o conceito de Dispositivo e todos seus filhos

```

90 ...
91   allow_archetype CLUSTER[at0011] occurrences matches {0..1}
92   matches {
93     include
94     archetype_id/value matches{
95       /openEHR-EHR-CLUSTER\.device(-[a-zA-Z0-9_]+)*\.v1/
96     }
97   }
98 ...

```

Ao criar os extratos de informação, mesmo usando a linguagem XML, os extratos, da mesma forma, usam os conceitos fazendo referência aos arquétipos de modo a poder rastrear os conceitos. No Código 5.3, podemos ver o conceito de Estatura sendo referenciado dentro do extrato de informação.

Código 5.3: Referenciamento do conceito de estatura no Extrato de Dados

```

10 ...
11   <items type="OBSERVATION">
12     <archetype_node_id>at0001</archetype_node_id>
13     <archetyped>
14       <archetype_id>openEHR-EHR-OBSERVATION.height.v1.adl</
15         archetype_id>
16       <rm_version>1.0.2</rm_version>
17     </archetyped>
18     <data type="ITEM_TREE">
19       <items type="ELEMENT">

```

```

19         <archetype_node_id>at0004</archetype_node_id>
20         <value type="DV_QUANTITY">
21             <magnitude>1.42</magnitude>
22             <units>m</units>
23         </value>
24         <name type="DV_TEXT">
25             <value>Estatura</value>
26         </name>
27     </items>
28 </data>
29 <name type="DV_TEXT">
30     <value>Estatura</value>
31 </name>
32 </items>
33 ...

```

Extrair informações de um sistema legado usando arquétipos se apresenta como uma potencial forma de comunicar sistemas legados através de padrões de informação. O tipo de construção apresentada nesta seção associada a mecanismos de comunicação, e.g. HL7, possibilita o desenvolvimento de mecanismos de troca de informação. Ou seja, amplia a capacidade de SIS de interoperar e ainda ao se utilizar padrões de informação e metodologias apropriadas pode alcançar a interoperabilidade semântica entre os SIS.

Capítulo 6

Conclusões e Trabalhos Futuros

Concluimos neste trabalho que o modelo em dois níveis, proposto pela Fundação openEHR, se apresenta como um modelo de agregação gradual de semântica aos SIS de relativa simplicidade frente a modelos como o OWL. A proposta do modelo, de separar os níveis de conhecimento invariantes dos variantes e a linguagem ADL, aceleram o desenvolvimento de ontologias em contextos de aplicação no domínio da saúde, além de favorecerem o reuso de conhecimento através de mecanismos de *slots* e *binding*.

Beale and Heard [6] ao definir a modelagem de dois níveis, colocam a questão de deixar a cargo dos especialistas do domínio da saúde a construção dos arquétipos, no entanto, apesar de não ter sido realizado um experimento mais apropriado em relação a essa avaliação, se percebe uma certa dificuldade dos profissionais de saúde em modelar conhecimento a partir de estruturas primárias e de formalização de conhecimento. Quando se insere o profissional na modelagem utilizando diretamente os editores de arquétipos em ADL a resistência é ainda maior. No entanto, isso pode ser contornado por meio do uso de outras técnicas como Mapa Mental, que é suportado pela própria ferramenta de edição dos arquétipos.

Também se percebe, como apontado em trabalhos correlatos, que a discussão dos modelos conceituais provocam algumas reações interessantes no grupo de trabalho, principalmente no sentido de buscar conceitos mais gerais do domínio e fortalecer esses conceitos no grupo. Por meio das discussões de conceitos, se percebe que também os conceitos de sistema de informação e sistema de software ficam mais evidentes, podendo os detalhes da aplicação serem discutidos em separado, considerando contextos tecnológicos distintos. Desta forma a representação do conhecimento pode ser feita de forma incremental, utilizando camadas ontológicas, viabilizadas pela metodologia de modelagem em dois níveis.

Neste trabalho foi possível realizar uma prova de conceito da metodologia de desenvolvimento de arquétipos, a partir de um experimento no SISVAN, definindo e reutilizando arquétipos relacionados ao sub-domínio da VAN. A partir dos arquétipos foi gerado um extrato de dados abrindo a possibilidade de automatizar o processo de extração de informação semanticamente enriquecidos pela implementação de algum mecanismo de comunicação entre SIS. As especificações do openEHR trazem vários elementos, a partir da classe *EXTRACT_EHR*, para viabilizar essa comunicação.

Em trabalhos futuros pretendemos usar os arquétipos definidos para implementar um *middleware* semântico apoiado nas especificações do openEHR, de forma a poder avaliar

os possíveis problemas de comunicação entre sistema fundamentado no padrão openEHR e sistema “não-openEHR”. Em iniciativas recentes discute-se no MS, a composição de um RES que permita troca de informação entre SIS do SUS, em um primeiro momento com foco no nível da Atenção Básica e nas informações essenciais para continuidade do cuidado ao paciente.

Também pretendemos fazer uso do *framework* open-EHR-Gen [21], a fim de avaliar uma forma de validação dos modelos de conhecimentos, em protótipos funcionais, junto aos especialistas do domínio. E com isso possibilitar um maior refinamento dos arquétipos definidos.

Referências

- [1] Vladimir Alexiev and et al. *Information Integration with Ontologies: Ontology based Information Integration in an Industrial Setting*. John Wiley & Sons, West Sussex, UK, April 2005. ISBN 0-470-01048-7. 1
- [2] Mauricio B. Almeida and Marcello P. Bax. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação*, 32, 2003. URL <http://www.scielo.br/pdf/ci/v32n3/19019>. 10
- [3] Javier Carnicero Giménez Azcárate. *De la historia clínica a la historia de salud electrónica*. Informes SEIS. Sociedad Española de Informática de la Salud, Pamplona, España, 2 edition, 2003. 1
- [4] André René Barboni. *Sistema de Informações Estratégicas Municipais - SIEM: o desafio da construção participativa de um sistema de informações epidemiologicamente sustentado*. PhD thesis, Departamento de Saúde, Universidade Estadual de Feira de Santana, Feira de Santana, Bahia, 2006. 2
- [5] T Beale and H Frankel. *Extract Information Model*. The openEHR Foundation, February 2007. URL http://www.openehr.org/releases/1.0.2/architecture/rm/ehr_extract_im.pdf. Release 1, In development. 45
- [6] T. Beale and S. Heard. An ontology-based model of clinical information. *Studies in health technology and informatics*, 129(1):760, 2007. 21, 70
- [7] T Beale and S Heard. *Architecture Overview*. The openEHR Foundation, April 2007. URL <http://www.openehr.org/releases/1.0.2/architecture/overview.pdf>. Release 1.0.2. 22, 67
- [8] T Beale and S Heard. *Archetype Definition Language*. The openEHR Foundation, December 2008. URL <http://www.openehr.org/releases/1.0.2/architecture/am/adl.pdf>. Release 1.0.2, ADL 1.4. 25
- [9] T Beale, S Heard, D Kalra, and D Lloyd. *EHR Information Model*. The openEHR Foundation, April 2007. URL http://www.openehr.org/releases/1.0.2/architecture/rm/ehr_im.pdf. Release 1.0.2. 21, 39

- [10] T. Beale, H. Frankel, D. Kalra, and D. Lloyd. *Data Types Information Model*. The openEHR Foundation, November 2008. URL http://www.openehr.org/releases/1.0.2/architecture/rm/data_types_im.pdf. Release 1.0.2, ADL 1.4. 54
- [11] Thomas Beale. Archetypes: Constraint-based domain models for future-proof information systems. *OOPSLA*, 2002. 1, 18, 19, 20, 24, 26
- [12] T. Berners-Lee and J. Hendler. Scientific publishing on the semantic web. *Nature*, 410:1023–1024, 2001. 16
- [13] J. Bisbal, G. Engelbrecht, and A. Frangi. Archetype-based semantic mediation: Incremental provisioning of data services. In *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*, pages 132–137. IEEE, 2010. 37, 48
- [14] R. Bonita, R. Beaglehole, and T. Kjellström. *Epidemiologia básica*. Santos, São Paulo, 2a edição edition, 2010. 2
- [15] Brasil. *Protocolos do Sistema de Vigilância Alimentar e Nutricional - SISVAN na assistência à saúde*. Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Atenção Básica, Brasil, 2008. URL http://189.28.128.100/nutricao/docs/geral/protocolo_sisvan.pdf. 50, 51, 65
- [16] Brasil. *Orientações para a coleta e análise de dados antropométricos em serviços de saúde : Norma Técnica do Sistema de Vigilância Alimentar e Nutricional - SISVAN*. Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Atenção Básica, 2011. URL <http://nutricao.saude.gov.br/publicacoes.php>. 46, 51, 54, 57
- [17] L.C. Burton, G.F. Anderson, and I.W. Kues. Using electronic health records to help coordinate care. *Milbank Quarterly*, 82(3):457–481, 2004. 34, 35
- [18] T. Buzan and B. Buzan. *How to mind map*. Thorsons London, 2002. 42
- [19] Maria Luiza Campos. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, 33(1), 2004. ISSN 1518-8353. URL <http://revista.ibict.br/index.php/ciinf/article/view/56/52>. 13, 14
- [20] E. Cardillo, A. Taminin, and L. Serafini. A methodology for knowledge acquisition in consumer-oriented healthcare. *Knowledge Discovery, Knowledge Engineering and Knowledge Management*, pages 249–261, 2011. 15
- [21] Leandro Carrasco and Pablo Pazos Gutiérrez. Traumagen: historia clínica electrónica con acceso telemático a imágenes médicas de pacientes de trauma. Technical report, Universidad de la República, Uruguay, Julio 2010. 47, 71
- [22] Morris F. Collen. Origins of medical informatics. In *Medical Informatics*, 1986. 34

- [23] Ministério da Saúde and CONASEMS. *O SUS de A a Z: garantindo saúde nos municípios*. Comunicação e Educação em Saúde. Ministério da Saúde, Brasília - DF, 3 edition, 2009. 2, 3, 31
- [24] Ministério da Saúde. Secretaria Executiva. Departamento de Informática do SUS. *DATASUS TRAJETÓRIA 1991-2002*. Estatística e Informação em Saúde. Ministério da Saúde, Brasília - DF, 1 edition, 2002. 2
- [25] I. Dahlberg. Teoria do conceito. *Ciência da informação*, 7(2), 1978. 14
- [26] Maria Luiza de Almeida Campos, Rosali de Souza, and Maria Luiza Machado Campos. Organização de unidades de conhecimento em hiperdocumentos: o modelo conceitual como espaço comunicacional para a realização da autoria. *Ciência da Informação*, 32(2), 2003. ISSN 1518-8353. URL <http://revista.ibict.br/index.php/ciinf/article/view/111/92>. 14
- [27] Departamento de Atenção Básica (SAS/MS). *Política Nacional de Atenção Básica*, volume 4 of *Série Pactos pela Saúde*. Ministério da Saúde, Brasília - DF, 2006. 31, 32
- [28] João Alberto de Oliveira Lima. *Modelo Genérico de Relacionamentos na Organização da Informação Legislativa e Jurídica*. PhD thesis, Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2008. 14
- [29] Organización Mundial de Salud. Acerca de la oms. Internet, Dezembro 2011. URL <http://www.who.int/about/es/>. 30
- [30] Rigoleta Dutra Mediano Dias and Sergio Miranda Freire. Arquétipos para representar as informações demográficas em saúde. In *XI Congresso Brasileiro de Informática em Saúde*, 2008. 47, 53
- [31] A. Dogac, G.B. Laleci, S. Kirbas, Y. Kabak, S.S. Sinir, A. Yildiz, and Y. Gurcan. Artemis: deploying semantically enriched web services in the healthcare domain. *Information Systems*, 31(4-5):321–339, 2006. 34
- [32] R.H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F.M. Behlen, P.V. Biron, and A.S. Shvo. HL7 clinical document architecture, release 2. *Journal of the American Medical Informatics Association*, 13(1):30–39, 2006. 23
- [33] Marco Eichelberg, Thomas Aden, Jörg Riesmeier, Asuman Dogac, and Gokce B. Laleci. A survey and analysis of electronic healthcare record standards. *ACM Comput. Surv.*, 37:277–315, December 2005. ISSN 0360-0300. 18, 34, 44
- [34] Luiz Renato Evangelisti, José Luiz Barbosa Bevilacqua, Adriana Cláudia Martins Mendoza Cuellas, and Luiz Arnoldo Haertel. Uma proposta de integração de sistema ged a um repositório comunitário ihe brasileiro. In *XII Congresso Brasileiro de Informática em Saúde*, 2010. 48

- [35] Luiz Augusto Facchini, Ernani Bento Bandarra, Beatriz de Faria Leão, Cláudio Giuliano Alves da Costa, Beatriz Figueiredo Dobashi, Ilara Hämmerli Sozzi de Moraes, and Marcus Vinicius da Silva. *Política Nacional de Informação e Informática em Saúde*. Ministério da Saúde - Secretaria Executiva - Departamento de Informação e Informática do SUS, Brasília - DF, 2nd edition, 2004. 3
- [36] I. P. Fellegi and A. B. Sunter. A theory for record linkage. *Journal of the American Statistical Association*, 40:1183–1210, 1969. 8
- [37] Dieter Fensel. *Ontologies: a silver bullet for knowledge management and electronic commerce*. Springer Verlag, 2004. 16
- [38] Aldo Gangemi, Domenico M. Pisanelli, and Gerardo Steve. Some requirements and experiences in integrating terminological ontologies in medicine. In *Eleventh Workshop on Knowledge Acquisition, Modeling and Management*, 1998. 27, 28
- [39] Aldo Gangemi, Domenico M. Pisanelli, and Geri Steve. Ontology integration: Experiences with medical terminologies. In *Formal Ontology in Information Systems*, pages 163–178. Press, 1998. 28
- [40] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993. ISSN 1042-8143. doi: DOI:10.1006/knac.1993.1008. 9
- [41] Thomas R. Gruber. *Encyclopedia of Database Systems*, chapter Ontology. Springer-Verlag, 2008. 1, 9
- [42] Nicola Guarino. Formal ontology and information systems. In *Formal Ontology and Information Systems*, pages 3–15. IOS Press, 1998. 1, 9, 10, 11, 14
- [43] Giancarlo Guizzardi. Desenvolvimento para e com reuso: Um estudo de caso no domínio de vídeo sob demanda. Master’s thesis, Universidade Federal do Espírito Santos, Vitória, E.S., Brasil, Junho 2000. 13
- [44] Giancarlo Guizzardi. *Ontological foundations for structural conceptual models*. PhD thesis, Centre for Telematics and Information Technology, University of Twente, Enschede, 2005. URL <http://doc.utwente.nl/50826/>. 9
- [45] Giancarlo Guizzardi and Terry Halpin. Ontological foundations for conceptual modelling. *Appl. Ontol.*, 3:1–12, January 2008. ISSN 1570-5838. URL <http://portal.acm.org/citation.cfm?id=1412417.1412423>. 11
- [46] Gustavo Diniz Ferreira Gusso. *Diagnóstico de demanda em Florianópolis utilizando a Classificação Internacional de Atenção Primária: 2a edição (CIAP-2)*. PhD thesis, Faculdade de Medicina da Universidade de São Paulo, São Paulo, 2009. 39, 61
- [47] Asunción Gómez-Pérez, Mariano Fernández-López, and Oscar Corcho-Garcia. *Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003. ISBN: 1852335513. 1, 9, 10, 14, 28

- [48] Farshad Hakimpour and Andreas Geppert. Resolution of semantic heterogeneity in database schema integration using formal ontologies. *Inf. Technol. and Management*, 6(1):97–122, 2005. ISSN 1385-951X. doi: <http://dx.doi.org/10.1007/s10799-004-7777-0>. 1
- [49] K. Häyrinen, K. Saranto, and P. Nykänen. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. *International journal of medical informatics*, 77(5):291–304, 2008. 35
- [50] Newcombe HB and JM. Kennedy. Record linkage making maximum use of the discrimination power of identifying information. *Communications of the ACM*, 5:563–566, 1962. 8
- [51] *SNOMED Clinical Terms Technical Reference Guide*. International Health Terminology Standards Development Organisation, July 2008. URL <http://www.ihtsdo.org/snomed-ct/release-of-snomed-ct/>. 29
- [52] A. Jovic, M. Prcela, and D. Gamberger. Ontologies in medical knowledge representation. In *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on*, pages 535–540, june 2007. doi: 10.1109/ITI.2007.4283828. 27
- [53] Dipak Kalra. *Clinical Foundations and Information Architecture for the Implementation of a Federated Health Record Service*. PhD thesis, University College London, São Paulo, 2002. 1
- [54] Dipak Kalra. Electronic health record standards. *IMIA yearbook of medical informatics*, 2006:136–144, 2006. 19, 22, 24
- [55] A. Kumar and B. Smith. The ontology of processes and functions: a study of the international classification of functioning, disability and health. In *AIME 2005 Workshop on Biomedical Ontology Engineering*, 2005. 15
- [56] Maurizio Lenzerini. Data integration: a theoretical perspective. In *PODS 02: Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 233–246, New York, NY, USA, 2002. ACM. ISBN 1-58113-507-6. doi: <http://doi.acm.org/10.1145/543613.543644>. 2
- [57] Leonardo Lezcano, Miguel-Angel Sicilia, and Carlos Rodríguez-Solano. Integrating reasoning and clinical archetypes using owl ontologies and swrl rules. *Journal of Biomedical Informatics*, 44(2):343 – 353, 2011. ISSN 1532-0464. doi: 10.1016/j.jbi.2010.11.005. URL <http://dl.acm.org/citation.cfm?id=1961723>. 25
- [58] G. C. Liu, J. G. Cooper, K. M. Schoeffler, and W. E. Hammond. Standards for the electronic health record, emerging from health care’s tower of babel. In *AMIA Symposium*, pages 388–392. AMIA, 2001. 34
- [59] *Logical Observation Identifiers Names and Codes, User Guide*. LOINC, December 2011. URL <http://loinc.org/downloads/files/LOINCManual.pdf>. 29

- [60] Daniel Luna, Adrian Gomez, Fernando Plazzotta, Fernando Campos, Marcela Martinez, Hernan Navas, Daniela Canosa, Enrique Soriano, and Fernan González Bernaldo de Quirós. Implementación del clinical document architecture standard en un sistema de información clínico. In *X Congresso Brasileiro de Informática em Saúde*, 2006. 46
- [61] A. C. M. Martins, Machado J. P., and D.P. Silveira. Review about the usage of probabilistic linkage in the brazilian health area. *The 5th Institution of Engineering and Technology International Seminar on Appropriate Healthcare Technologies for developing Countries*, 15(2):34–45, May 2008. 8
- [62] Isabel Román Martínez. *Aportaciones Metodológicas a la Integración de Sistemas de Información Sanitarios Basada en Gestión Semántica*. PhD thesis, Ingenieria de Sistemas y Automática, Universidad de Sevilla, Sevilla, España, 2006. 1
- [63] Catalina Martínez-Costa, Marcos Menárguez-Tortosa, and Jesualdo Tomás Fernández-Breis. An approach for the semantic interoperability of iso en 13606 and openehr archetypes. *Journal of Biomedical Informatics*, 43(5): 736 – 746, 2010. ISSN 1532-0464. doi: 10.1016/j.jbi.2010.05.013. URL <http://www.sciencedirect.com/science/article/pii/S1532046410000821>. 25, 38
- [64] D.L. McGuinness, R. Fikes, J. Hendler, and L.A. Stein. Daml+ oil: an ontology language for the semantic web. *Intelligent Systems, IEEE*, 17(5):72–80, 2002. 16
- [65] Fernando da Silva Medeiros. *Um mar de possibilidades: A medicina no passado presente e futuro*. Biblioteca24horas, Seven System Internacional, 2010. 24
- [66] C.F. Naiman and A.M. Ouksel. A classification of semantic conflicts in heterogeneous database systems. *Journal of Organizational Computing and Electronic Commerce*, 5(2):167–193, 1995. 8
- [67] Fabiane Bizinella Nardon, Tony França, and Humberto Naves. Construção de aplicações em saúde baseadas em arquétipos. In *XI Congresso Brasileiro de Informática em Saúde*, 2008. 47
- [68] Ricardo Alfredo Quintano Neira, Fabiane Bizinella Nardon, Lincoln de Assis Moura Jr, and Beatriz de Faria Leão. Como incorporar conhecimento aos sistemas de registro eletrônico em saúde? In *XI Congresso Brasileiro de Informática em Saúde*, 2008. 47
- [69] *Digital Imaging and Communications in Medicine (DICOM)*. NEMA, National Electrical Manufacturers Association, 2011. URL http://medical.nema.org/Dicom/2011/11_01pu.pdf. 30
- [70] Eluzai Souza dos Santos. Uma proposta de integração de sistemas computacionais utilizando ontologias. Dissertação de mestrado, Universidade de Brasília, 2006. 47
- [71] Marcelo Rodrigues dos Santos. *Sistema de registro eletrônico de saúde baseado na norma ISO 13606: aplicações na Secretaria de Estado de Saúde de Minas Gerais*. PhD thesis, Universidade Federal de Minas Gerais, Belo Horizonte, 2011. 6, 20, 21, 25, 40, 41, 49, 62

- [72] Peter Schloeffel, Thomas Beale, George Hayworth, Sam Heard, and Heather Leslie. The relationship between CEN 13606, HL7, and openEHR. In *Health Informatics Conference*, 2006. 21, 22, 38
- [73] E. Sciore, M. Siegel, and A. Rosenthal. Using semantic values to facilitate interoperability among heterogeneous information systems. *ACM Transactions on Database Systems (TODS)*, 19(2):254–290, 1994. 7
- [74] Pablo Serrano, David Moner, Tomás Sebastián, José A. Maldonado, Rafael Navalón, Monserrat Robles, and Ángel Gómez. Utilidad de los arquetipos iso 13606 para representar modelos clínicos detallados. *Revista e Salud*, 5(18), 2009. 18
- [75] A.P. Sheth. *Interoperating geographic information systems*, chapter Changing Focus on Interoperability Information Systems: From System, Syntax, Structure to Semantics. Kluwer Academic, 1999. 7, 8
- [76] A.P. Sheth and J.A. Larson. Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Surveys (CSUR)*, 22(3):183–236, 1990. 7
- [77] Daniel Sigulem. *Um Novo Paradigma de Aprendizado na Prática Médica da UNIFESP/EPM*. PhD thesis, Centro de Informática em Saúde, EPM/UNIFESP, São Paulo, 1997. 34
- [78] João Paulo Lyra da Silva, Claudia Travassos, Miguel Murat de Vasconcellos, and Linair Maria Campos. Revisão sistemática sobre encadeamento ou linkage de bases de dados secundários para uso em pesquisa em saúde no brasil. *Cadernos de Saúde Coletiva*, 14(2):197–204, April 2006. 8
- [79] *O Papel do Nutricionista na Atenção Primária à Saúde*. Sistema Conselhos Federal e Regionais de Nutricionistas, 2008. URL <http://www.cfn.org.br/eficiente/repositorio/Cartilhas/61.pdf>. 46
- [80] R.R. Souza and L. Alvarenga. A web semântica e suas contribuições para a ciência da informação. *Ciência da informação*, 33(1):132–141, 2004. 16
- [81] Bárbara Starfield. *Atenção primária: equilíbrio entre necessidades de saúde, serviços e tecnologia*, volume 4. UNESCO, Ministério da Saúde, Brasília, Brasil, 2002. 32, 34, 39
- [82] György Surján. *Barriers and challenges of using medical coding systems*. PhD thesis, Academic Medical Center, University of Amsterdam, Amsterdam, 2011. 15
- [83] P.C. Tang, J.S. Ash, D.W. Bates, J.M. Overhage, and D.Z. Sands. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association*, 13(2):121, 2006. 35
- [84] Frank van Harmelen and Deborah L. McGuinness. OWL web ontology language overview. W3C recommendation, W3C, February 2004. URL <http://www.w3.org/TR/2004/REC-owl-f-20040210/>. 17

- [85] Gertjan van Heijst, Sabina Falasconi, Ameen Abu-Hanna, Guus Schreiber, and Mario Stefanelli. A case study in ontology library construction. *Artificial Intelligence in Medicine*, 7, 1995. 10
- [86] P. Yu. Archetypes, gehr, openehr and electronic health records. *The Journal on Information Technology in Healthcare*, 200:369, 2003. 18

Apêndice A

CPIISS - Ministério da Saúde

Os padrões de interoperabilidade, publicado pelo Departamento de Informática do SUS (DATASUS/SGEP/MS), constarão do **Catálogo de Padrões de Interoperabilidade de Informações de Sistemas de Saúde** (CPIISS), disponível em uma primeira versão como Anexo a Portaria GM nº 2.073 que "Regulamenta o uso de Padrões de Interoperabilidade e Informação em Saúde para sistemas de informação em saúde no âmbito do Sistema Único de Saúde, nos níveis municipal, distrital, estadual e federal, e para os sistemas privados e do setor de saúde suplementar", pelo Ministério da Saúde do Brasil.

A.1 Catálogo de Padrões de Informação

Os padrões são definidos em nível lógico (negócios) e não físico de arquivamento de banco de dados. Estes padrões não documentam propriedades de exibição. Os sistemas legados podem ter suas respostas, para integração e interoperação, encapsuladas em padrões XML aderentes aos padrões do Catálogo, de forma que, mesmo sem obedecer internamente ao padrão catalogado, possam comunicar-se fazendo uso dele, por meio de XML Schemas

1. Para a definição do Registro Eletrônico em Saúde (RES) será utilizado o modelo de referência OpenEHR, disponível em <http://www.openehr.org/home.html>.
2. Para estabelecer a interoperabilidade entre sistemas, com vistas à integração dos resultados e solicitações de exames, será utilizado o padrão HL7 - Health Level 7.
3. Para codificação de termos clínicos e mapeamento das terminologias nacionais e internacionais em uso no país, visando suportar a interoperabilidade semântica entre os sistemas, será utilizada a terminologia SNOMED-CT, disponível em <http://www.ihtsdo.org/snomed-ct/>.
4. Para a interoperabilidade com sistemas de saúde suplementar serão utilizados os padrões TISS (Troca de Informações em Saúde Suplementar).
5. Para a definição da arquitetura do documento clínico será utilizado o padrão HL7 CDA.
6. Para a representação da informação relativa a exames de imagem será utilizado o padrão DICOM.

7. Para a codificação de exames laboratoriais será utilizado o padrão LOINC (Logical Observation Identifiers Names and Codes).
8. Para a codificação de dados de identificação das etiquetas de produtos relativos ao sangue humano, de células, tecidos e produtos de órgãos, será utilizada a norma ISBT 128.
9. Para a interoperabilidade de modelos de conhecimento, incluindo arquétipos, templates e metodologia de gestão, será utilizado o padrão ISO 13606-2.
10. Para o cruzamento de identificadores de pacientes de diferentes sistemas de informação, será utilizada a especificação de integração IHE-PIX (Patient Identifier Cross-Referencing).
11. Outras classificações que serão utilizadas para suporte à interoperabilidade dos sistemas de saúde: CID, CIAP-2 (Atenção primária de saúde), TUSS e CBHPM (Classificação brasileira hierarquizada de procedimentos médicos) e tabela de procedimentos do SUS.

Apêndice B

Formulário do SISVAN

Usado pelo Sistema de Vigilância Alimentar e Nutricional e definido pelo Ministério da Saúde (SAS/DAB/CGAN), este formulário, é utilizado para coletar os dados do Cadastro Individual e do primeiro Acompanhamento Nutricional.

	Ministério da Saúde/ SAS/ DAB/ CGPAN		
	SISTEMA DE VIGILÂNCIA ALIMENTAR E NUTRICIONAL		
	Estabelecimento de Saúde	Nº CNES*	
	Nome ou Matrícula do Profissional de Saúde		
Ficha: __ Inclusão __ Alteração			
DADOS CADASTRAIS			
Cadastro de domicílio			
Endereço completo (tipo de logradouro, nome do logradouro, número, complemento)*			
Bairro*	CEP	DDD	Telefone
Nº CNES do domicílio*	Estabelecimento de Saúde		
Cadastro do indivíduo			
Nome completo (sem abreviaturas)*		Data de Nascimento*	Data do Cadastro
		/ /	/ /
Nome completo da mãe (sem abreviaturas)*		Nome completo do pai	
Sexo*	Raça / Cor*	Escolaridade* ⁽¹⁾	Nacionalidade
<input type="checkbox"/> 1. Masculino <input type="checkbox"/> 2. Feminino	<input type="checkbox"/> 1. Branca <input type="checkbox"/> 2. Negra <input type="checkbox"/> 3. Amarela <input type="checkbox"/> 4. Parda <input type="checkbox"/> 5. Indígena		<input type="checkbox"/> Brasileira <input type="checkbox"/> Estrangeira
Data de naturalização	UF Nascimento	Município Nascimento	Situação familiar ⁽²⁾
/ /			
Documentação do indivíduo			
NIS (Nº Identificação Social)	NCNS (Nº Cartão Nac. Saúde)	NPCNS (Nº Provisório Cartão Nac. Saúde)	Outro código identificador:
O registro de pelo menos um documento oficial é obrigatório* (consulte lista dos documentos oficiais no verso):			
Tipo ⁽³⁾	Dados do documento*		
Tipo ⁽³⁾	Dados do documento		
Programas Vinculados:	<input type="checkbox"/> Programa Bolsa Família <input type="checkbox"/> _____ <input type="checkbox"/> _____		

Figura B.1: Formulário, frente, Dados Cadastrais

ACOMPANHAMENTO NUTRICIONAL**						
Data do acompanhamento*: / /						
Criança (<10 anos)	Peso (em kg)*:	Altura (em cm)*:	Estado nutricional:		Tipo de alimentação ^{(4)***} :	Peso ao nascer (em gramas):
			Peso por idade:	Altura por idade:		
Adolescente (≥10 e <20 anos)	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:		IMC por idade:	Altura por idade:
Adulto (≥20 e <60 anos)	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:	Circunferência da cintura (em cm):	Risco aumentado:	
					<input type="checkbox"/> Sim	<input type="checkbox"/> Não
Idoso (≥60 anos)	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:			
Gestante	Peso (em kg)*:	Altura (em m)*:	Estado nutricional:	Peso pré-gestacional (em kg):	Data da última menstruação*: / /	
Doenças*:		Deficiências e/ou intercorrências*:			Tipo de Acompanhamento*:	
<input type="checkbox"/> Anemia falciforme <input type="checkbox"/> Diabetes mellitus <input type="checkbox"/> Doenças cardiovasculares <input type="checkbox"/> Hipertensão Arterial Sistêmica <input type="checkbox"/> Osteoporose <input type="checkbox"/> Outras doenças <input type="checkbox"/> Sem doenças		<input type="checkbox"/> Anemia ferropriva <input type="checkbox"/> DDI (Distúrbio por Deficiência de Iodo) <input type="checkbox"/> Diarréia <input type="checkbox"/> Infecções intestinais virais <input type="checkbox"/> IRA (Infecção Respiratória Aguda) <input type="checkbox"/> Hipovitaminose A <input type="checkbox"/> Outras deficiências e/ou intercorrências <input type="checkbox"/> Sem deficiências e/ou intercorrências			<input type="checkbox"/> Atendimento na Atenção Básica <input type="checkbox"/> Chamada Nutricional <input type="checkbox"/> Saúde na Escola <input type="checkbox"/> _____ <input type="checkbox"/> _____	

* Campos de preenchimento obrigatório (fundo cinza).

** Para maiores informações sobre o registro do acompanhamento nutricional, consulte os materiais técnicos do SISVAN.

*** Campo obrigatório apenas para crianças menores de 2 anos.

Figura B.2: Formulário, frente, Dados de Acompanhamento Nutricional

Legendas:

(1) Escolaridade:

1. Não sabe ler/escrever
2. Alfabetizado (indivíduo lê e escreve pelo menos um bilhete)
3. Nível fundamental incompleto (1º grau incompleto)
4. Nível fundamental completo (1º grau completo)
5. Nível médio incompleto (2º grau incompleto)
6. Nível médio completo (2º grau completo)
7. Superior incompleto
8. Superior completo
9. Especialização/ Residência
10. Mestrado
11. Doutorado
12. Sem informação

(3) Tipo de documentos oficiais:

01. Registro geral / Identidade (RG)
02. Cadastro de pessoa física (CPF)
03. Carteira de Trabalho e Previdência Social (CTPS)
04. Carteira Nacional de Habilitação (CNH)
05. Título Eleitoral
06. Número de Identificação Social (NIS/PIS)
08. Documento estrangeiro
09. Passaporte
10. Certificado de Reservista Militar (CRM)
12. Carteira Funcional
13. Registro de Órgão de Classe
14. Certificado de naturalização
91. Certidão de Nascimento
92. Certidão de Casamento
93. Certidão de Separação ou Divórcio
95. Certidão Administrativa - Índio

(2) Situação familiar:

1. Convive com companheira(o) e filho(s)
2. Convive com companheira(o), com laços conjugais e sem filho(s)
3. Convive com companheira(o), com filho(s) e/ou outros familiares
4. Convive com familiar(es), sem companheira(o)
5. Convive com outras pessoas sem laços consanguíneos e/ou laços conjugais
6. Vive só

(4) Tipo de Alimentação:

1. Aleitamento materno exclusivo
2. Aleitamento materno predominante
3. Alimentação complementar (leite materno e alimentos)
4. Não recebe leite materno
5. Sem informação

Figura B.3: Formulário, verso, Legendas

Apêndice C

Arquétipo do SISVAN em linguagem ADL

Código C.1: Arquétipo de define o conceito de Acompanhamento Nutricional

```
1 archetype (adl_version=1.4)
2   openEHR-EHR-COMPOSITION.Nutritional-Monitoring.v1
3
4 concept
5   [at0000]
6
7 language
8   original_language = <[ISO_639-1::pt-br]>
9
10 description
11   lifecycle_state = <"Draft">
12   details = <
13     ["pt-br"] = <
14       language = <[ISO_639-1::pt-br]>
15       purpose = <"Fazer o acompanhamento nutricional de um individuo.">
16       keywords = <"nutricao","nutricional","peso","acompanhamento">
17       copyright = <"Ministerio da Saude">
18       use = <"Usado para acompanhar individuos adultos e idosos">
19       misuse = <"Nao usar para acompanhar criancas e gestantes, nesses
20         casos usar as especializacoes COMPOSITION.nutritional_monitoring
21         -child e COMPOSITION.nutritional_monitoring-pregnancy">
22     >
23   >
24
25 definition
26   — Acompanhamento Nutricional
27   COMPOSITION[at0000] occurrences matches {1..1} matches {
28     content existence matches {1..1}
29     cardinality matches {2..*; ordered}
30     matches {
31       — Observacao
32       SECTION[at0011] occurrences matches {1..*} matches {
33         items existence matches {1..1}
34         cardinality matches {2..*; unordered; unique}
35         matches {
36           allow_archetype OBSERVATION[at0001]
```

```

35     occurrences matches {1..1}
36 matches {
37     include
38         archetype_id/value matches {
39             /openEHR-EHR-OBSERVATION.height.v1/
40         }
41     exclude
42         archetype_id/value matches {/.*}/
43 }
44 allow_archetype OBSERVATION[at0013]
45 occurrences matches {1..1}
46 matches {
47     include
48         archetype_id/value matches {
49             /openEHR-EHR-OBSERVATION.body_weight.v1/
50         }
51     exclude
52         archetype_id/value matches {/.*}/
53 }
54 allow_archetype OBSERVATION[at0006]
55 occurrences matches {0..1}
56 matches {
57     include
58         archetype_id/value matches {
59             /openEHR-EHR-OBSERVATION.anthropometric_index-
60             body_mass_index.v1/
61         }
62     exclude
63         archetype_id/value matches {/.*}/
64 }
65 allow_archetype OBSERVATION[at0014]
66 occurrences matches {0..1}
67 matches {
68     include
69         archetype_id/value matches {
70             /openEHR-EHR-OBSERVATION.waist_hip.v1/
71         }
72     exclude
73         archetype_id/value matches {/.*}/
74 }
75 }
76 — Avaliacao
77 SECTION[at0010] occurrences matches {1..*} matches {
78     items existence matches {1..1}
79     cardinality matches {1..*; unordered; unique}
80 matches {
81     allow_archetype EVALUATION[at0002]
82     occurrences matches {1..1}
83 matches {
84     include
85         archetype_id/value matches {
86             /openEHR-EHR-EVALUATION.nutrition_summary.v1/
87         }
88     exclude
89         archetype_id/value matches {/.*}/

```

```

90     }
91     allow_archetype EVALUATION[at0007]
92     occurrences matches {0..*}
93     matches {
94         include
95             archetype_id/value matches {
96                 /openEHR-EHR-EVALUATION.problem-diagnosis.v1/
97             }
98         exclude
99             archetype_id/value matches {/.*}/
100     }
101     allow_archetype EVALUATION[at0008]
102     occurrences matches {0..*}
103     matches {
104         include
105             archetype_id/value matches {
106                 /openEHR-EHR-EVALUATION.risk.v1/
107             }
108         exclude
109             archetype_id/value matches {/.*}/
110     }
111     allow_archetype EVALUATION[at0009]
112     occurrences matches {0..*}
113     matches {
114         include
115             archetype_id/value matches {
116                 /openEHR-EHR-EVALUATION.problem-diagnosis.v1/
117             }
118         exclude
119             archetype_id/value matches {/.*}/
120     }
121     allow_archetype EVALUATION[at0012]
122     occurrences matches {0..*}
123     matches {
124         include
125             archetype_id/value matches {
126                 /openEHR-EHR-EVALUATION.problem-diagnosis.v1/
127             }
128         exclude
129             archetype_id/value matches {/.*}/
130     }
131     }
132     }
133     }
134     }
135
136 ontology
137     terminologies_available = <...>
138     term_definitions = <
139         ["pt-br"] = <
140             items = <
141                 ["at0000"] = <
142                     text = <"Acompanhamento Nutricional">
143                     description = <"Sumario de acompanhamento nutricional do
144                         individuo">

```

```

145 ["at0001"] = <
146   text = <"Altura / Estatura">
147   description = <"Altura ou comprimento do corpo, e medida a partir
      da coroa da cabeça a sola do pé. A altura e medida com o
      indivíduo na posição de pé e comprimento do corpo na posição
      decubito dorsal.">
148 >
149 ["at0002"] = <
150   text = <"Avaliação Nutricional">
151   description = <"Sumário da avaliação nutricional do indivíduo
      acompanhado.">
152 >
153 ["at0003"] = <
154   text = <" ">
155   description = <" ">
156 >
157 ["at0006"] = <
158   text = <"Índice Antropométrico">
159   description = <" ">
160 >
161 ["at0007"] = <
162   text = <"Doenças Associadas">
163   description = <"Doenças (crônicas) associadas ao indivíduo no
      momento do acompanhamento. ">
164 >
165 ["at0008"] = <
166   text = <"Risco">
167   description = <"Risco avaliado com base na informação alimentar/
      nutricional">
168 >s
169 ["at0009"] = <
170   text = <"Intercorrelações">
171   description = <"Doenças ou queixas, durante ou relacionada, na
      avaliação">
172 >
173 ["at0010"] = <
174   text = <"Avaliação">
175   description = <"Este é um objeto SECTION"> — Organiza blocos de
      informação
176 >
177 ["at0011"] = <
178   text = <"Observação">
179   description = <"Este é um objeto SECTION"> — Organiza blocos de
      informação
180 >
181 ["at0013"] = <
182   text = <"Peso corporal">
183   description = <"A medição do peso corporal de um indivíduo.">
184 >
185 ["at0014"] = <
186   text = <"Circunferência da cintura e do quadril">
187   description = <"Circunferência (ou perímetro) da cintura e do
      quadril">
188 >
189 ["at0015"] = <
190   text = <" ">

```

```
191         description = <">
192     >
193     ["at0012"] = <
194         text = <"Deficiencias">
195         description = <"Deficiencias diagnosticadas na avaliacao">
196     >
197 >
198 >
199 >
200 constraint_definitions = <
201 >
202 term_binding = <
203 >
204 constraint_binding = <
205 >
```

Apêndice D

Arquétipo de Estatura em linguagem ADL

Este arquétipo foi carregado através da ferramenta CKM da Fundação openEHR, no entanto foram feitas simplificações no conteúdo do código para ficar mais didático e enxuto¹.

Código D.1: Arquétipo reutilizado de Estatura do repositório da Fundação openEHR

```
1 archetype (adl_version=1.4)
2   openEHR-EHR-OBSERVATION.height.v1
3
4 concept
5   [at0000] — Height/Length
6
7 language
8   original_language = <[ISO_639-1::en]>
9
10 description
11   original_author = <
12     ["name"] = <"Sam Heard">
13     ["organisation"] = <"Ocean Informatics">
14     ["email"] = <"sam.heard@oceaninformatics.com">
15     ["date"] = <"9/03/2006">
16   >
17   details = <
18     ["pt-br"] = <
19       language = <[ISO_639-1::pt-br]>
20       purpose = <"Para registrar o comprimento do corpo de um individuo,
21         medindo da coroa da cabeça a sola do pe. A medida pode ser tanto
22         real como aproximada, quer seja com a posicao do individuo de pe
23         ou em decubito dorsal.">
24       use = <"Texto removido.">
25       keywords = <"encolhimento", "crescimento", "diminuicao", "diminuicao
26         da altura", "altura", "comprimento", "crescimento">
27       misuse = <"Texto removido.">
28     >
29   >
30 definition
```

¹Para acessar o código sem alterações é possível carregar o código original através do endereço: <http://www.openehr.org/knowledge/retrieveFile?archetype-cid=1013.1.138&format=ADL>

```

28 | OBSERVATION[at0000] matches {
29 |   data matches {
30 |     HISTORY[at0001] matches {
31 |       events cardinality matches {1..*; unordered} matches {
32 |         EVENT[at0002] occurrences matches {1..*} matches {
33 |           data matches {
34 |             ITEM_TREE[at0003] matches {
35 |               items cardinality matches {1..*; unordered} matches {
36 |                 ELEMENT[at0004] matches {
37 |                   value matches {
38 |                     C_DV_QUANTITY <
39 |                       property = <[openehr::122]>
40 |                       list = <
41 |                         ["1"] = <
42 |                           units = <"cm">
43 |                           magnitude = <|0.0..1000.0|>
44 |                         >
45 |                         ["2"] = <
46 |                           units = <"in">
47 |                           magnitude = <|0.0..250.0|>
48 |                         >
49 |                       >
50 |                     >
51 |                   }
52 |                 }
53 |               ELEMENT[at0018] occurrences matches {0..1} matches {
54 |                 value matches {
55 |                   DV_TEXT matches {*}
56 |                 }
57 |               }
58 |             }
59 |           }
60 |         }
61 |       state matches {
62 |         ITEM_TREE[at0013] matches {
63 |           items cardinality matches {0..*; unordered} matches {
64 |             ELEMENT[at0014] occurrences matches {0..1} matches {
65 |               value matches {
66 |                 DV_CODED_TEXT matches {
67 |                   defining_code matches {
68 |                     [local::
69 |                       at0016,
70 |                       at0020;
71 |                       at0016]
72 |                   }
73 |                 }
74 |               }
75 |             }
76 |           ELEMENT[at0019] occurrences matches {0..1} matches {
77 |             value matches {
78 |               DV_TEXT matches {*}
79 |             }
80 |           }
81 |         }
82 |       }
83 |     }

```

```

84     }
85   }
86 }
87 }
88 protocol matches {
89   ITEM_TREE[at0007] matches {
90     items cardinality matches {0..*; ordered} matches {
91       allow_archetype CLUSTER[at0011] occurrences matches {0..1}
92       matches {
93         include
94           archetype_id/value matches{
95             /openEHR-EHR-CLUSTER\.device(-[a-zA-Z0-9_]+)*\.v1/
96           }
97       }
98     }
99   }
100 }
101 }
102
103
104 ontology
105 term_definitions = <
106   ["pt-br"] = <
107     items = <
108       ["at0000"] = <
109         text = <"Altura / comprimento">
110         description = <"Altura ou comprimento do corpo, e medida a
111           partir da coroa da cabeça a sola do pé. A altura e medida
112           com o indivíduo na posição de pé e comprimento do corpo na
113           posição decubito dorsal.">
114       >
115       ["at0001"] = <
116         text = <"history">
117         description = <"@ internal @">
118       >
119       ["at0002"] = <
120         text = <"Qualquer evento">
121         description = <"Medida a qualquer momento da altura ou
122           comprimento do corpo.">
123       >
124       ["at0003"] = <
125         text = <"Simple">
126         description = <"@ internal @">
127       >
128       ["at0004"] = <
129         text = <"Altura / comprimento">
130         description = <"0 comprimento do corpo da coroa da cabeça a
131           sola do pé.">
132       >
133       ["at0007"] = <
134         text = <"List">
135         description = <"@ internal @">
136       >
137       ["at0011"] = <
138         text = <"Dispositivo">

```

```

134         description = <"Descricao do dispositivo utilizado para medir
135             altura ou comprimento do corpo.">
136     >
137     ["at0013"] = <
138         text = <"Tree">
139         description = <"@ internal @">
140     >
141     ["at0014"] = <
142         text = <"Posicao">
143         description = <"Posicao individual quando medido.">
144     >
145     ["at0016"] = <
146         text = <"De pe">
147         description = <"A altura e medida de pe sobre os dois pes com o
148             peso distribuido uniformemente, calcanhaes juntos e as
149             nadeegas e os calcanhaes em contato com uma placa traseira
150             vertical.">
151     >
152     ["at0018"] = <
153         text = <"Comentario">
154         description = <"Comentario sobre a medicao da altura/
155             comprimento do corpo.">
156     >
157     ["at0019"] = <
158         text = <"Fatore de erro">
159         description = <"Registrar quaisquer problemas ou fatores que
160             possam ter impacto sobre a medicao da altura/comprimento do
161             corpo, por exemplo, observacao de amputacao.">
162     >
163     ["at0020"] = <
164         text = <"Decubito dorsal">
165         description = <"O comprimento e medido em uma posicao
166             totalmente estendida, deitada com a pelve plana, pernas
167             estendidas e os pes flexionados.">
168     >
169 >
170 >
171 >
172 >

```

Apêndice E

Extrato de Informação do SISVAN em XML

Código E.1: XML com extrato de informação de Acompanhamento Nutricional

```
1 <?xml version="1.0" encoding="iso-8859-1"?>
2 <composition type="COMPOSITION">
3   <archetyped>
4     <archetype_id>openEHR-EHR-COMPOSITION.nutritional_monitoring.v1.adl</archetype_id>
5     <rm_version>1.0.2</rm_version>
6   </archetyped>
7   <archetype_node_id>at0000.1</archetype_node_id>
8   <content type="SECTION">
9     <archetype_node_id>at0011</archetype_node_id>
10    <items type="OBSERVATION">
11      <archetype_node_id>at0001</archetype_node_id>
12      <archetyped>
13        <archetype_id>openEHR-EHR-OBSERVATION.height.v1.adl</archetype_id>
14        <rm_version>1.0.2</rm_version>
15      </archetyped>
16      <data type="ITEM_TREE">
17        <items type="ELEMENT">
18          <archetype_node_id>at0004</archetype_node_id>
19          <value type="DV_QUANTITY">
20            <magnitude>1.42</magnitude>
21            <units>m</units>
22          </value>
23          <name type="DV_TEXT">
24            <value>Estatura</value>
25          </name>
26        </items>
27      </data>
28      <name type="DV_TEXT">
29        <value>Estatura</value>
30      </name>
31    </items>
32  <items type="OBSERVATION">
33    <archetype_node_id>at0013</archetype_node_id>
34    <archetyped>
```

```

35         <archetype_id>openEHR-EHR-OBSERVATION.body_weight.v1.adl
36         </archetype_id>
37         <rm_version>1.0.2</rm_version>
38     </archetyped>
39     <data type="ITEM_TREE">
40         <items type="ELEMENT">
41             <archetype_node_id>at0004</archetype_node_id>
42             <value type="DV_QUANTITY">
43                 <magnitude>40</magnitude>
44                 <units>kg</units>
45             </value>
46             <name type="DV_TEXT">
47                 <value>Peso</value>
48             </name>
49         </items>
50     </data>
51     <name type="DV_TEXT">
52         <value>Peso</value>
53     </name>
54 </items>
55 <items type="OBSERVATION">
56     <archetype_node_id>at0014</archetype_node_id>
57     <archetyped>
58         <archetype_id>openEHR-EHR-OBSERVATION.waist_hip.v1</
59         archetype_id>
60         <rm_version>1.0.2</rm_version>
61     </archetyped>
62     <data type="ITEM_TREE">
63         <items type="ELEMENT">
64             <archetype_node_id>at0004</archetype_node_id>
65             <value type="DV_QUANTITY">
66                 <magnitude>68</magnitude>
67                 <units>cm</units>
68             </value>
69             <name type="DV_TEXT">
70                 <value>Circunferencia da Cintura</value>
71             </name>
72         </items>
73     </data>
74     <name type="DV_TEXT">
75         <value>Circunferencia da cintura e do quadril</value>
76     </name>
77 </items>
78 <items type="OBSERVATION">
79     <archetype_node_id>at0006</archetype_node_id>
80     <archetyped>
81         <archetype_id>openEHR-EHR-OBSERVATION.
82         anthropometric_index-body_mass_index.v1.adl</
83         archetype_id>
84         <rm_version>1.0.2</rm_version>
85     </archetyped>
86     <data type="ITEM_TREE">
87         <items type="ELEMENT">
88             <archetype_node_id>at0004</archetype_node_id>
89             <value type="DV_QUANTITY">
90                 <magnitude>19.84</magnitude>

```

```

87         <units>kg/m2</units>
88     </value>
89     <name type="DV_TEXT">
90         <value>Indice de massa corporea</value>
91     </name>
92 </items>
93 </data>
94 <name type="DV_TEXT">
95     <value>Indice de massa corporea</value>
96 </name>
97 </items>
98 <name type="DV_TEXT">
99     <value>Observacao</value>
100 </name>
101 </content>
102 <content type="SECTION">
103     <archetype_node_id>at0010</archetype_node_id>
104     <items type="EVALUATION">
105         <archetype_node_id>at0002</archetype_node_id>
106         <archetyped>
107             <archetype_id>openEHR-EHR-EVALUATION.nutrition.summary.v1
108                 </archetype_id>
109             <rm_version>1.0.2</rm_version>
110         </archetyped>
111         <data type="ITEM_TREE">
112             <items type="ELEMENT">
113                 <archetype_node_id>at0003</archetype_node_id>
114                 <value type="DV_CODED_TEXT">
115                     <value>at0007</value>
116                 </value>
117                 <name type="DV_TEXT">
118                     <value>Diagnostico Nutricional</value>
119                 </name>
120             </items>
121         </data>
122         <name type="DV_TEXT">
123             <value>Sumario Nutricional</value>
124         </name>
125     </items>
126     <name type="DV_TEXT">
127         <value>Avaliacao</value>
128     </name>
129 </content>
130 <name type="DV_TEXT">
131     <value>Acompanhamento Nutricional</value>
132 </name>
</composition>

```