

UNIVERSIDADE DE SÃO PAULO

FFCLRP-DEPARTAMENTO DE FÍSICA E MATEMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA APLICADA À MEDICINA E BIOLOGIA

Uma solução interoperável, baseada na UMLS, para apoiar a decisão diagnóstica colaborativa na Web

Daniel Facciolo Pires

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Doutor em Ciências, Área: Física Aplicada à Medicina e Biologia.

**Ribeirão Preto - SP
2007**

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

UNIVERSIDADE DE SÃO PAULO

FFCLRP-DEPARTAMENTO DE FÍSICA E MATEMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA APLICADA À MEDICINA E BIOLOGIA

Uma solução interoperável, baseada na UMLS, para apoiar a decisão diagnóstica colaborativa na Web

Daniel Facciolo Pires

Orientador: *Prof. Dr. Evandro Eduardo Seron Ruiz*

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Doutor em Ciências, Área: Física Aplicada à Medicina e Biologia.

**Ribeirão Preto – SP
2007**

Dedicatória

Ao meu pai e a minha mãe que sempre me deram apoio e incentivo durante toda a realização deste trabalho.

Agradecimentos

À minha esposa Melissa, que sempre esteve comigo nos melhores e piores momentos;

Ao meu filho Lucas, que acompanhou de perto o final deste trabalho, mesmo dentro da barriga da mamãe;

Ao meu irmão Alex e a minha irmã Milene, companheiros inseparáveis por toda a vida;

Ao meu orientador Evandro, por acreditar em mim e me dar a chance de realizar este trabalho, e ainda me apoiar e contribuir qualitativamente com o projeto, mesmo estando, no final do trabalho, afastado para seu Pós-Doutorado na Universidade de Columbia;

Ao Dr. Ricardo Halah, que acreditou no trabalho e contribuiu fortemente com suas idéias, e conhecimento clínico para o desenvolvimento do estudo de caso;

Aos meus colegas do laboratório IMAGCOM, em especial ao Adriano, ao Ivan, a Julia e a Rebeca, pelo companheirismo;

À UniCOC pelo apoio financeiro indireto às pesquisas realizadas envolvendo o projeto desta tese;

E a todos que, de alguma forma, tornaram este trabalho possível.

Resumo

Sistemas de Suporte a Decisão Clínica (CDSS) são sistemas baseados em conhecimento que utilizam dados de pacientes para gerar um parecer médico com objetivo de apoiar o usuário na decisão diagnóstica. Geralmente os sistemas hospitalares e clínicos fornecem informações criadas apenas por usuários locais ao CDSS para que uma base de conhecimento possa ser criada. Conseqüentemente, o CDSS apoiará seus usuários com o aprendizado realizado apenas em experiências locais. Por outro lado, uma situação interessante é quando os sistemas normalmente legados podem alimentar um CDSS com informações clínicas provenientes também de outros sistemas remotos, e criadas por médicos de outras instituições. Neste cenário, um médico poderia ser apoiado de forma colaborativa na decisão diagnóstica por vários outros médicos localizados remotamente, a partir do compartilhamento de bases de conhecimento. Neste cenário de computação distribuída, é comum a aparição de problemas de interoperabilidade técnica causados principalmente pelo uso de diferentes protocolos de comunicação para a criação e chamada de serviços clínicos remotos. Ainda no cenário apresentado anteriormente, podem surgir problemas de interoperabilidade semântica justificados pelo emprego de sistemas de terminologia e de ontologia clínica heterogêneos por parte das aplicações médicas. Assim sendo, estes problemas de interoperabilidade precisam ser explorados e solucionados de modo a permitir que a atividade de apoio à decisão diagnóstica colaborativa seja possível. Este trabalho investiga e discute os problemas de interoperabilidade semântica e técnica durante a troca de informações entre CDSSs e sistemas em saúde. A partir desta discussão, é proposta uma solução interoperável, baseada na UMLS, para apoiar a decisão diagnóstica colaborativa na Web. Esta solução é composta da criação de uma ontologia denominada DDSOnt, que estende a *UMLS Semantic Network*, e define um novo conjunto de tipos semânticos para a criação de uma estrutura que suporte o compartilhamento de bases de conhecimento de modo que possam ser utilizadas por médicos no apoio à decisão clínica colaborativa. A construção da ontologia busca promover a compatibilidade ontológica. Ainda, são confeccionados softwares que auxiliam os desenvolvedores na utilização da ontologia DDSOnt na construção de aplicações clínicas que precisam ser apoiadas na tomada de decisão diagnóstica: uma API denominada JDDSOnt para a definição e construção de documentos da Web Semântica em conformidade com a DDSOnt, e um conjunto de serviços Web, denominado DDSOntWs, que permite que esta troca de experiência clínica ocorra em ambiente distribuído, orientado a serviços, e independente de plataforma, promovendo assim a compatibilidade técnica. Desta maneira, um CDSS é projetado na forma de um serviço Web. Objetivando promover a compatibilidade terminológica, um dos serviços Web da DDSOntWs desenvolvido no trabalho faz uso da *UMLS Metathesaurus*. Um estudo de caso para a construção de aplicativos médicos que façam uso da solução proposta também foi realizado para teste e validação da metodologia apresentada.

Abstract

Clinical Decision Support Systems (CDSS) are knowledge based systems that make use of patient data to generate a diagnosis response aiming to support a clinical decision. Usually, hospital and clinical systems supply CDSS with information generated by local users. Thus, CDSS's knowledge base may be created based on local experiences learning and will support users with this type of experience. A different and interesting situation would be if legacy systems could supply CDSS with clinical information deriving from physically remote systems, and created by physicians from others affiliations. At this scenario, a physician could be supported in a collaborative environment during a clinical decision with physicians working remotely, and with shared clinical knowledge bases. Concerning this distributed computer scenario technical interoperability, communication problems may appear mainly due to different communication protocol usage to create and invoke remote clinical services. Still considering the scenario presented before, semantic interoperability problems can occur because heterogeneous terminology and ontology clinical systems may be used by distinctive medical applications. Therefore, these interoperability problems must be explored and solved to allow a collaborative clinical decision support activities. This works aims to investigate and discuss these technical and semantic interoperability problems caused by clinical information exchanged between CDSSs and health institutions systems. Based on that discussion, an interoperability and UMLS based solution is proposed to support a collaborative clinical decision environment on the Web. The solution is composed by a novel DDSOnt (Diagnosis Decision Support Ontology) ontology, that extends UMLS Semantic Network to create a new set of semantic types needful to define a structure that supports a shared knowledge databases to be used by physicians during the collaborative clinical decision. Ontology construction aims to promote ontology compatibility. Still, softwares have been developed to make DDSOnt useful by clinical applications that needs to be supported during clinical decision: a JDDSOnt API to define and to create semantic documents DDSOnt conformed, and a set of web services named DDSOntWs to enable a collaborative, service oriented, distributed computer environment, and plataform independent. Therefore, a CDSS is created as a web service. Aiming to promote terminology compatibility, one DDSOnt web service use the UMLS MetaThesaurus. An acute abdominal pain use case is also presented to demonstrate DDSOnt usability.

Sumário

| | |
|---|-----------|
| Lista de Figuras | xiii |
| Lista de Tabelas | xv |
| Lista de Abreviaturas | xvii |
| Trabalhos Publicados pelo Autor | xix |
| 1 Introdução | 1 |
| 1.1 Considerações Iniciais | 1 |
| 1.2 Motivação | 4 |
| 1.3 Objetivos | 5 |
| 1.4 Estrutura da tese | 6 |
| 2 Interoperabilidade Técnica em Sistemas de Informação em Saúde Distribuídos | 9 |
| 2.1 Considerações Iniciais | 9 |
| 2.2 Cenário Médico com Problemas de Interoperabilidade Técnica | 9 |
| 2.3 Serviços Web | 11 |
| 2.3.1 Funcionamento de Arquitetura de Serviços Web | 11 |
| 2.3.2 XML | 13 |
| 2.3.3 SOAP | 15 |
| 2.3.4 WSDL | 17 |
| 2.3.5 UDDI | 19 |
| 2.4 Serviços Web Semânticos | 21 |
| 2.4.1 OWL-S: OWL Web Ontology Language for Services | 21 |
| 2.5 Considerações Finais | 28 |
| 3 Incompatibilidade Terminológica em Sistemas de Informação em Saúde Distribuídos | 29 |
| 3.1 Considerações Iniciais | 29 |
| 3.2 Cenário Médico com Problemas de Incompatibilidade Terminológica | 29 |
| 3.3 Sistemas de Terminologia em Saúde | 31 |
| 3.3.1 CID | 31 |
| 3.3.2 SNOMED | 34 |
| 3.3.3 MeSH | 36 |
| 3.4 UMLS Metathesaurus – Um sistema que promove a compatibilidade terminológica | 37 |
| 3.4.1 Introdução ao UMLS | 37 |
| 3.4.2 Introdução ao UMLS Metathesaurus | 38 |

| | | |
|----------|---|-----------|
| 3.4.3 | Conceitos na UMLS Metathesaurus | 38 |
| 3.4.4 | <i>Strings</i> na UMLS Metathesaurus | 39 |
| 3.4.5 | Átomos na UMLS Metathesaurus | 39 |
| 3.4.6 | Termo Léxico na UMLS Metathesaurus | 39 |
| 3.4.7 | Átomos diferentes para representar um mesmo conceito | 40 |
| 3.4.8 | Átomos iguais para representar conceitos diferentes | 41 |
| 3.4.9 | UMLS API | 43 |
| 3.5 | Considerações Finais | 45 |
| 4 | Incompatibilidade Ontológica em Sistemas de Informação em Saúde Distribuídos | 47 |
| 4.1 | Considerações Iniciais | 47 |
| 4.2 | Cenário Médico com Problemas de Incompatibilidade Ontológica | 47 |
| 4.3 | Ontologias Clínicas | 48 |
| 4.3.1 | UMLS SN | 49 |
| 4.3.2 | Galen Common Reference Model - CRM | 54 |
| 4.3.3 | Modelo HL7 RIM | 57 |
| 4.4 | Web Semântica | 63 |
| 4.4.1 | RDF | 63 |
| 4.4.2 | OWL | 67 |
| 4.5 | Considerações Finais | 69 |
| 5 | Ontologia DDSOnt e Softwares Relacionados | 71 |
| 5.1 | Considerações Iniciais | 71 |
| 5.2 | Definição da Ontologia DDSOnt | 71 |
| 5.2.1 | Escolha da Metodologia | 71 |
| 5.2.2 | Execução do Passo 1 | 72 |
| 5.2.3 | Execução do Passo 2 | 73 |
| 5.2.4 | Execução do Passo 3 | 74 |
| 5.2.5 | Execução do Passo 4 | 77 |
| 5.2.6 | Execução dos Passos 5 e 6 | 78 |
| 5.2.7 | Execução do Passo 7 | 79 |
| 5.3 | Softwares Relacionados a DDSOnt | 81 |
| 5.3.1 | Uma Introdução aos Documentos Semânticos da DDSOnt | 81 |
| 5.3.2 | JDDSOnt | 83 |
| 5.3.3 | DDSOntWs | 85 |
| 5.4 | Considerações Finais | 90 |
| 6 | Resultados da Implementação do Estudo de Caso | 91 |
| 6.1 | Considerações Iniciais | 91 |
| 6.2 | Considerações Iniciais do Estudo de Caso | 91 |
| 6.2.1 | A Síndrome, os Sinais Clínicos e Laboratoriais, e os Diagnósticos | 91 |
| 6.2.2 | Entrada Estruturada dos Dados | 94 |
| 6.2.3 | Requisitos de Software do Estudo de Caso | 95 |
| 6.3 | Módulo de Apoio a uma Entidade Reguladora | 96 |
| 6.4 | Módulo de Apoio para Juntas Médicas ou Médicos Provedores | 102 |
| 6.5 | Módulo de Apoio para Médicos Requisitantes | 109 |
| 6.6 | Considerações Finais | 113 |

| | |
|--|------------|
| 7 Discussão dos Resultados | 115 |
| 7.1 Considerações Finais | 115 |
| 7.2 Trabalhos Relacionados | 116 |
| 7.3 Contribui original para a área de Informática em Saúde | 119 |
| 7.4 Trabalhos Futuros | 120 |
| Referências | 124 |

Lista de Figuras

| | | |
|------|---|----|
| 2.1 | Cenário ilustrando problemas de interoperabilidade | 10 |
| 2.2 | Cenário com especificações de serviços Web. | 11 |
| 2.3 | Funcionamento da arquitetura serviços Web. | 12 |
| 2.4 | Funcionamento da arquitetura serviços Web com serviço de publicação e descoberta. | 13 |
| 2.5 | Relacionamento entre as entidades reguladora, provedora e requisitante | 21 |
| 2.6 | Classes e propriedades da linguagem OWL-S para definir perfil do serviço | 23 |
| 2.7 | Classes e propriedades da linguagem OWL-S para definir perfil do serviço | 27 |
| 3.1 | Mesmo conceito médico sendo representado por diferentes termos | 30 |
| 3.2 | Mesmo termo médico com conceitos diferentes | 31 |
| 3.3 | CID-10 disponível para consulta na Internet. | 33 |
| 3.4 | Relacionamento entre as entidades da UMLS Metathesaurus. . . | 40 |
| 3.5 | Formas de acesso a UMLSKS. | 43 |
| 4.1 | Aplicações utilizando sistemas ontológicos diferentes | 48 |
| 4.2 | Porção da rede semântica UMLS: Entidades do tipo objeto físico. | 50 |
| 4.3 | Porção da rede semântica UMLS: Entidades do tipo conceitual . . | 51 |
| 4.4 | Porção da rede semântica UMLS: Tipos semânticos relacionados a evento. | 52 |
| 4.5 | Porção da rede semântica UMLS: Relacionamento entre os tipos de relação | 52 |
| 4.6 | Porção da rede semântica UMLS: relacionamento hierárquico e associativo entre tipos semânticos. | 54 |
| 4.7 | Galen CRM de alto nível. | 55 |
| 4.8 | Porção da Galen CRM: Especialização dos fenômenos | 56 |
| 4.9 | Porção da Galen CRM: Hierarquia dos modificadores. | 57 |
| 4.10 | Classes principais do modelo RIM | 59 |
| 4.11 | Classes Especializadas de uma 'Entidade' do modelo RIM | 60 |
| 4.12 | Classes Especializadas de uma 'Regra' do modelo RIM | 60 |
| 4.13 | Classes Especializadas de uma 'Ação' do modelo RIM | 61 |
| 4.14 | Documento RDF na forma gráfica. | 65 |
| 4.15 | Parte da ontologia DDSOnt. | 66 |
| 5.1 | Definição dos tipos semânticos da DDSOnt | 77 |

| | | |
|------|---|-----|
| 5.2 | Definição da DDSOnt | 80 |
| 5.3 | DDSOnt Java API - Auxilia na construção de documentos semânticos em conformidade a DDSOnt | 84 |
| 5.4 | Implementação e uso dos serviços Web da DDSOntWs | 85 |
| 5.5 | Serviços Web relacionados a DDSOnt e a classe CDKDescription | 87 |
| 5.6 | DDSOntWs com suporte UDDI | 89 |
| 6.1 | Diagrama de Requisitos de Software do Estudo de Caso | 96 |
| 6.2 | Diagrama de Caso de Uso para Módulo de Apoio a Entidade Reguladora. | 97 |
| 6.3 | Tela para a definição da patologia, dos sinais e diagnósticos. | 98 |
| 6.4 | Tela para a definição da patologia, dos sinais e diagnósticos suportados pelo serviço getUMLSInfo | 99 |
| 6.5 | Tela para registrar a definição de serviços da DDSOntWs. | 100 |
| 6.6 | Tela para registrar a implementação do serviço getUMLSInfo. | 101 |
| 6.7 | Diagrama de Caso de Uso para Módulo de Apoio à Junta Médica. | 103 |
| 6.8 | Tela para apoiar a busca por TModels no servidor UDDI | 104 |
| 6.9 | Tela para apoiar a entrada de dados relativos a identificação do paciente | 105 |
| 6.10 | Tela para apoiar a entrada de dados relativos a primeira parte da queixa/duração do paciente | 106 |
| 6.11 | Tela para apoiar a entrada de dados relativos a segunda parte da queixa/duração do paciente | 106 |
| 6.12 | Tela para apoiar a entrada de dados relativos aos exames físicos do paciente | 107 |
| 6.13 | Tela para apoiar a entrada de dados relativos aos exames laboratoriais do paciente | 107 |
| 6.14 | Tela para apoiar a entrada de dados relativos ao diagnóstico do paciente | 108 |
| 6.15 | Tela para apoiar a consulta por uma resposta diagnóstica a partir de base de conhecimento local | 108 |
| 6.16 | Tela para apoiar a consulta das estatísticas e da árvore de decisão da base de conhecimento local | 109 |
| 6.17 | Diagrama de Caso de Uso para Módulo de Apoio ao Médico | 110 |
| 6.18 | Tela para apoiar a busca por implementações de um <i>tModel</i> | 111 |
| 6.19 | Tela para apoiar a busca por um base de conhecimento | 112 |
| 6.20 | Tela para apoiar a consulta por uma resposta diagnóstica | 112 |

Lista de Tabelas

| | | |
|-----|---|----|
| 3.1 | Átomos diferentes para representar um mesmo conceito | 41 |
| 3.2 | Átomos iguais para representar conceitos diferentes | 42 |
| 3.3 | Métodos da API UMLSKS | 44 |
| 3.4 | Classes do pacote <i>UMLSKS Object Model</i> | 45 |
| 5.1 | Primeiro conjunto de termos presentes na DDSOnt | 74 |
| 5.2 | Termos do segundo conjunto presentes na DDSOnt | 75 |
| 5.3 | Termos do terceiro conjunto presentes na DDSOnt | 76 |
| 6.1 | Sinais clínicos de Queixa/Duração Definidos Para o Estudo de Caso | 93 |
| 6.2 | Sinais clínicos de Exame Físico Definidos Para o Estudo de Caso | 94 |
| 6.3 | Sinais Laboratoriais Definidos Para o Estudo de Caso | 94 |

Lista de Abreviaturas

ANSI (*American National Standards Institute*)
CID (*International Classification of Diseases*)
COSTAR (*Computer-Stored Ambulatory Records*)
DCOM (*Distributed Component Object Model*)
DSOM (*Distributed System Object Model*)
DTD (*Document Type Definition*)
ISO (*International Organization for Standardization*)
Galen CRM (*Galen Common Reference Model*)
GRAIL (*Galen Representation and Integration Language*)
HL7 (*Health Level Seven*)
HL7 CDA (*HL7 Common Document Architecture*)
HL7 RIM (*HL7 Reference Information Model*)
HTTP (*Hypertext Transfer Protocol*)
KRSS (*Knowledge Representation System Specification*)
LOINC (*Logical Observation Identifiers Names and Codes*)
MeSH (*Medical Subject Headings*)
NAICS (*North American Industry Classification System*)
NLM (*National Library of Medicine*)
OSI (*Open Systems Interconnection*)
OWL (*Ontology Web Language*)
PSY (*Thesaurus of psychological index*)
RDF (*Resource Descriptor Framework*)
RDFS (*Resource Description Framework*)
RMI (*Remote Method Invocation over Internet Inter-Orb Protocol*)
SDO (*Standard Developing Organization*)
SOAP (*Simple Object Access Protocol*)
SODont (*Second Opinion Diagnosis Ontology*)
SNOMED (*Systematized Nomenclature of Medicine*)
SNOMED CT (*Systematized Nomenclature of Medicine–Clinical Terms*)
SNOMED RT (*Systematized Nomenclature of Medicine–Reference Terminology*)
SNOP (*Standard Nomenclature of Pathology*)
UDDI (*Universal Description, Discovery and Integration*)
UML (*Unified Modeling Language*)
UMLS (*Unified Medical Language System*)
URI (*Unique Resource Identifier*)
OWL (*Web Ontology Language*)
W3C (*World Wide Web Consortium*)
WHO (*World Health Organization*)
WSDL (*Web Service Description Language*)
XML (*Extensible Markup Language*)

Trabalhos Publicados pelo Autor Envolvendo este Trabalho

Pires, D. F. and Ruiz, E. E. S. Uma Arquitetura para Suporte à Atividade de Telemedicina através do Compartilhamento do Conhecimento Clínico Extraído de Sistemas de Prontuário Eletrônico. In Proceedings of the X Simpósio Brasileiro de Sistemas Multimídia e Web. 226–229(2). Ribeirão Preto-SP. 2004.

Pires, D. and Halah, R. and Tinos, R. and Ruiz, E. Uma Arquitetura para Suportar o Compartilhamento de Conhecimento Clínico em Sistemas PEP Integrados a Sistemas de Auxílio ao Diagnóstico. In Proceedings of the IX Congresso Brasileiro de Informática em Saúde, 593–598(1). Ribeirão Preto-SP. 2004.

Pires, D. and Halah, R. and Tinos, R. and Ruiz, E. An Architecture to Support Clinical Knowledge Sharing Among Medical Community via Internet. In Proceedings of the V Workshop de Informática Médica. Congresso Brasileiro de Qualidade de Software. 114-124(1). Porto Alegre-RS. 2005.

Pires, D. F., Ruiz, E. E. S., Teixeira, C. A. C. Compartilhamento do Conhecimento Clínico Suportado por Ontologias e pela Comunicação Multimídia. In Proceedings of XII Simpósio Brasileiro de Sistemas Multimídia e Web. 23-25 (2). Natal-RN. 2006.

Pires, D. F., Ruiz, E. E. S. Uso da SODont e Softwares Relacionados para Suportar a Segunda Opinião Diagnóstica. In Proceedings of XII Simpósio Brasileiro de Sistemas Multimídia e Web. 73-77 (2). Natal-RN. 2006.

Pires, D. F., Halah, R. A., Ruiz, E. E. S. Sharing Web Clinical Knowledge Between Medical Community to Support a Second Opinion Diagnosis. In Proceedings of XXVI Congresso da Sociedade Brasileira de Computação. XXXIII Seminário Integrado de Software e Hardware - SEMISH. 361-375(1). Campo Grande-MS. 2006.

Pires, D. F., Halah, R. A., Ruiz, E. E. S. Interoperabilidade Semântica na Troca de Informações de Segunda Opinião Diagnóstica. In Proceedings of X Congresso Brasileiro de Informática em Saúde. 885-890 (1). Florianópolis-SC. 2006

Bataglião, L. L. Uso da UMLS Na Solução de Problemas de Incompatibilidade Terminológica em Saúde: Um Estudo de Caso. Trabalho de Conclusão de Curso. UniCOC. 56 p. Ribeirão Preto. 2006

Araujo, F. M. Padrão HL7: Um Estudo de Caso. Trabalho de Conclusão de Curso. UniCOC. 65 p. Ribeirão Preto. 2006

Araujo, F. M., Pires, D. F. Padrão HL7: Um Estudo de Caso. 14^o Simpósio In-

ternacional de Iniciação Científica da Universidade de São Paulo. Poster. São Paulo. 2006

Martins, J. A. C. M. and Pires, D. F. and Teixeira, C. A. C. Construção de bases de conhecimento clínicas para segunda opinião diagnóstica em ambiente suportado por ontologias e pela comunicação multimídia. Projeto de Pesquisa TT2 adjacente ao Processo 05/60653-1. TIDIA-Ae Fase II. FAPESP. 2007.

Pereira, C. M. M. and Pires, D. F. and Teixeira, C. A. C. Criação de novas atividades de aprendizagem médica em ambiente e-learning suportado por ontologias e pela comunicação multimídia. Projeto de Pesquisa TT1 adjacente ao Processo 05/60653-1. TIDIA-Ae Fase II. FAPESP. 2007.

Introdução

1.1 Considerações Iniciais

Sistemas Baseados em Conhecimento em geral são caracterizados por conter dois módulos principais e que mais se destacam: uma base de conhecimento e um mecanismo de inferência (Rezende 2002). A base de conhecimento contém uma representação computacional das informações específicas de um determinado domínio de aplicação, enquanto que o mecanismo de inferência consiste de algoritmos e técnicas computacionais responsáveis pelo desenvolvimento de raciocínio e aprendizado a partir das informações representadas na base de conhecimento. Assim, um programa que inferiu e aprendeu com as informações representadas em uma base de conhecimento de um determinado domínio, pode responder perguntas ou validar respostas do usuário neste domínio de acordo com seu aprendizado.

Sistemas de Suporte a Decisão Clínica (ou Diagnóstica) ou *Clinical Decision Support Systems* – CDSS são Sistemas Baseados em Conhecimento que se utilizam de dados de pacientes para gerar uma recomendação, ou um parecer médico com objetivo de apoiar o usuário durante uma decisão diagnóstica (Wyatt and Spiegelhalter 1991). Este tipo de sistema tem se tornado um importante recurso para as instituições em saúde como hospitais e clínicas médicas, pois podem melhorar a qualidade na prescrição de medicamentos, como demonstra Sjoborg e colaboradores (Sjöborg et al. 2007), podem evitar uma decisão diagnóstica tendenciosa, como explica Bradley e co-autores (Bradley 2005), e também podem reduzir custos administrativos em serviços ambulatoriais e clínicos, como exemplificado por Krause (Krause et al. 2004) e Haung (Haung et al. 2003), e seus respectivos colaboradores.

As conquistas apresentadas anteriormente tem se tornado realidade apesar de algumas barreiras apresentadas por Short e co-autores (Short et al. 2004). Ainda assim, acreditamos que é justificável pesquisar e aplicar CDSS em sistemas de informação em saúde, assim como outros que criaram importantes ferramentas para o suporte à decisão diagnóstica. Alguns exemplos presentes na literatura são os trabalhos de Tan e co-autores (Tan et al. 2005), que criam *softwares* para auxiliar o

diagnóstico de câncer no ovário, de Hassan (Hassan 2007), que propõe CDSS para evitar aplicações de doses de insulina em excesso, e de Samore e colaboradores (Samore et al. 2005), que utilizam de um CDSS para auxiliar a prescrição de antibióticos para pacientes com infecções respiratórias.

Instituições em saúde possuem ainda outros tipos de sistemas computacionais, muitas vezes legados, como Prontuários Eletrônicos de Pacientes ou *Electronic Healthcare Systems – EHS* (Zheng et al. 2005), Sistemas de Informações Radiológicas ou *Radiology Information System* (Smith-Akin et al. 2007), Sistemas de Informações Ambulatoriais ou *Ambulatory Information System* (Murff et al. 2003), dentre outros. São estes sistemas que, geralmente, fornecem informações ao CDSS para que uma base de conhecimento possa ser criada e representada, e inferências realizadas. Com isso, um médico pode então ser apoiado durante uma decisão diagnóstica.

Usualmente, os sistemas hospitalares e clínicos citados anteriormente contém informações criadas apenas por usuários locais. Consequentemente, o CDSS apoiará seus usuários com o aprendizado realizado em experiências locais. Por outro lado, uma interessante situação seria se os sistemas das instituições em saúde pudessem inserir dados em um CDSS com informações clínicas provenientes de outros sistemas, e criadas por médicos de outras instituições. Neste cenário, um médico poderia ser apoiado durante a decisão diagnóstica por vários outros médicos localizados remotamente, como um especialista ou uma junta médica de referência. Sendo assim, cria-se um ambiente colaborativo em que médicos podem auxiliar e serem auxiliados durante uma tomada de decisão clínica a partir do compartilhamento de bases de conhecimento.

Para que o cenário apresentado torne-se possível, e os CDSSs possam utilizar-se de informações provenientes de sistemas de instituições em saúde diversos e legados, é necessário que estes sistemas troquem e compartilhem dados clínicos. Esta situação pode criar um ambiente heterogêneo, de computação distribuída, compartilhado e orientado a serviços clínicos. Pode-se citar como exemplos destes serviços a criação de uma base de conhecimento clínica, a obtenção de uma resposta para apoiar a decisão diagnóstica, ou ainda a publicação ou a busca por uma base de conhecimento em específico (Pires et al. 2006a).

O ambiente computacional apresentado para compartilhamento de informações clínicas pode se tornar heterogêneo pois os serviços são normalmente criados e utilizados por diferentes protocolos de comunicação como DCOM

(Horstmann and Kirtland 1997), DSOM (Coskun and Sessions 1992) e RMI/I-IOP (Sun 2007). Como consequência do uso destes diferentes protocolos, o ambiente pode se tornar dependente de linguagem de programação e/ou de arquitetura computacional. Ainda, os serviços do ambiente geralmente não são descritos, nem publicados e nem descobertos na Web de forma padronizada e universal de modo a aumentar a colaboração e compartilhamento destes. Assim, exige-se o estudo e a solução de problemas de interoperabilidade técnica para se propor que o apoio à decisão diagnóstica possa ocorrer em ambiente distribuído.

As especificações SOAP (Mitra 2006), um protocolo de comunicação para computação distribuída, WSDL (Chinnici et al. 2006), uma linguagem de descrição de serviços, e UDDI (UDDI 2007), uma especificação e mecanismo para publicação e busca de serviços, mostram-se interessantes de serem aplicadas no ambiente distribuído descrito anteriormente, visto que são padrões independentes de plataforma computacional, serem soluções gratuitas e de código-fonte aberto, e ainda por utilizarem XML como formato de seus documentos associados. Estas especificações fazem parte da

arquitetura Serviços Web (Booth et al. 2006). Assim, um CDSS pode ser visto como um serviço Web.

O cenário de troca e compartilhamento de informações clínicas com objetivo de fazer com que um CDSS possa utilizar-se de informações provenientes de diversos sistemas de instituições em saúde ocasiona também problemas de interoperabilidade semântica. Estes problemas precisam ser explorados, especificamente, os de incompatibilidade terminológica e ontológica. Problemas terminológicos podem ocorrer quando aplicações médicas que precisam trocar dados clínicos utilizam-se de diferentes vocabulários médicos como por exemplo: *COSTAR* (Barnett et al. 1979), *MeSH* (Lipscomb 2000), *RxNorm* (Liu et al. 2005) e *SNOMED CT* (Cote et al. 1993), só para citar alguns. Problemas ontológicos podem acontecer quando sistemas em saúde precisam compartilhar informações médicas e fazem uso de ontologias clínicas incompatíveis, como pode ser o caso de utilização das seguintes ontologias: *UMLS Semantic Network* (McCray 2003), *Galen CRM* (Rector et al. 1996), *HL7 RIM* (Seven 2006) e *Gene Ontology* (Ashburner et al. 2000).

Na maioria das vezes, a compatibilidade terminológica em aplicações que precisam apoiar o médico na decisão diagnóstica pode ser conseguida com o uso do tesauro digital *UMLS Metathesaurus* (Schuyler et al. 1993), que relaciona os conceitos médicos entre termos de mais de 100 diferentes vocabulários, preservando os seus significados. Isso significa que o *UMLS Metathesaurus* pode ser utilizado por aplicações durante a troca de informações clínicas para apoiar a decisão diagnóstica, e portanto é utilizado neste trabalho para promover a compatibilidade terminológica.

Por outro lado, a compatibilidade ontológica mostra-se ainda sem solução na literatura, em função das dificuldades de se ter uma só ontologia padrão que consiga abranger todas as necessidades médicas. Entretanto, se o *UMLS MetaThesaurus* for utilizado para promover a compatibilidade terminológica durante a troca de dados clínicos, a compatibilidade ontológica pode ser conseguida frequentemente com a *UMLS Semantic Network*, já que esta cria tipos semânticos baseados nos conceitos do tesauro da UMLS.

Para que a *UMLS Semantic Network* possa ser utilizada especificamente para apoiar a decisão diagnóstica, uma extensão desta ontologia pode ser proposta de modo que satisfaça um novo conjunto de tipos semânticos necessários para a criação de uma estrutura que suporte o compartilhamento de bases de conhecimento de modo que possam ser utilizadas por médicos no apoio à decisão clínica em ambiente colaborativo.

O uso de especificações da Web Semântica (Miller et al. 2006c) como *RDF* (Miller et al. 2006b) e *OWL* (Miller and Hendler 2006) permite que ontologias possam ser criadas e representadas na forma de documentos na Web. E que estes documentos possam ainda expressar significado sobre seu conteúdo, podendo criar assim um ambiente na Internet no qual agentes de software possam obter informações que apenas o ser humano poderia concluir (Berners-Lee et al. 2001). Sendo assim, estas especificações podem ser utilizadas na implementação de uma ontologia clínica para o ambiente da Web que possa suportar a troca de dados para apoiar a decisão diagnóstica.

Por fim, a compatibilidade sintática é também tratada neste projeto, porém não é discutida porque considera-se que todos os documentos clínicos trocados entre as aplicações computacionais são representados em XML.

Este trabalho propõe uma solução interoperável baseada na UMLS que permita que aplicações de instituições em saúde legadas possam suprir um CDSS orientado a

serviços com informações provenientes de sistemas de outros profissionais médicos. Esta proposta cria portanto um cenário de apoio à decisão diagnóstica em ambiente colaborativo. A proposta ainda faz uso da *UMLS MetaThesaurus* para promover a compatibilidade terminológica, reutiliza a *UMLS Semantic Network* para a construção de uma ontologia que promove a compatibilidade ontológica, e ainda utiliza as especificações de Serviços Web e UDDI para promover a compatibilidade técnica.

1.2 Motivação

Com o cenário de um ambiente computacional que suporta a troca e o compartilhamento de informações clínicas para apoiar a decisão diagnóstica compartilhada e distribuída, novas atividades médicas podem se tornar possíveis, como as apresentadas a seguir (Pires et al. 2004b; Pires et al. 2004a):

- Um médico pode construir uma base de conhecimento clínico a partir de sua base contendo casos clínicos locais;
- Um clínico pode publicar e compartilhar sua base de conhecimento clínico através da Internet;
- Um médico pode descobrir e recuperar uma base de conhecimento clínico através da Internet;
- Um clínico pode especificar um conjunto de valores de sinais e diagnósticos relacionados a uma patologia, podendo esta ser especificada a partir da experiência do médico ou então através da frequência dos sinais e diagnósticos relevantes na literatura médica. Particularmente no primeiro caso, a habilidade de um médico em diagnosticar pode ser substituída por uma junta médica especializada.
- O conjunto de sinais e diagnósticos pode ser utilizado e compartilhado por médicos da mesma especialidade. Uma interface gráfica de um EHS pode também ser automaticamente configurada com campos de entrada estruturada de dados que se refiram aos mesmos sinais e diagnósticos sugeridos.
- Médicos participantes da comunidade Web podem ser responsáveis em gerenciar e moderar algumas situações interessantes como:
 - Um médico propõe um novo sinal ou um novo diagnóstico para uma comunidade de uma patologia em particular;
 - Um médico insere um novo caso clínico que ele considera interessante em uma base de conhecimento compartilhada;

Em ambos os casos, a comunidade médica é responsável em decidir se estes dados devem ser relevantes para a inclusão na base de conhecimento. A decisão por incluir um novo caso pode ser moderado por uma comunidade através de um agente de software. Esta situação é particularmente importante porque ela pode ser usada para auxiliar na descoberta de novas regras e inferências de diagnósticos, oferecendo a possibilidade de descoberta de novas patologias e doenças, podendo melhorar as estatísticas em saúde.

No contexto das atividades médicas descritas anteriormente, e aprimoradas com aspectos colaborativos, algumas situações interessantes podem ser possíveis (Pires et al. 2004b; Pires et al. 2004a):

- Um médico menos experimentado pode, antes de se decidir por um diagnóstico, consultar uma base de conhecimento de um colega;
- Um médico pode ainda consultar por duas ou mais diferentes bases de conhecimento antes de se decidir por um diagnóstico. Por exemplo, primeiro consultar uma base formada por uma junta médica, segundo consultar sua própria base de conhecimento, e terceiro consultar a base construída por um outro colega médico.

Entretanto, para se criar um ambiente computacional que promova as atividades e situações clínicas apresentadas nesta seção, permitindo assim a troca de conhecimento clínico entre a comunidade médica através da Internet de modo a apoiar a decisão diagnóstica, é necessário explorar problemas de interoperabilidade semântica e técnica descritos anteriormente na Seção 1.1.

1.3 Objetivos

Este trabalho investiga e discute os problemas de interoperabilidade semântica e técnica durante a troca de informações entre CDSSs e sistemas de instituições em saúde. E a partir desta discussão, é proposta Uma Solução Interoperável e Baseada na UMLS para Apoiar a Decisão Diagnóstica Colaborativa na Web. Esta solução é composta da criação de uma ontologia denominada DDSOnt (*Diagnosis Decision Support Ontology*), que estende a *UMLS Semantic Network*, e define um novo conjunto de tipos semânticos necessários para a criação de uma estrutura que suporte o compartilhamento de bases de conhecimento de modo que possam ser utilizadas por médicos no apoio à decisão clínica. A criação desta ontologia objetiva promover a compatibilidade ontológica.

A solução proposta neste trabalho ainda é composta por softwares que auxiliam os desenvolvedores na utilização da ontologia DDSOnt durante a construção de aplicações clínicas que precisam ser apoiadas na tomada de decisão diagnóstica. Para isso, foi construída uma API denominada JDDSOnt para a definição e construção de documentos da Web Semântica em conformidade com a DDSOnt, de modo que documentos contendo informações de apoio diagnóstico possam ser trocados entre as aplicações. Também foram projetados e implementados serviços Web denominados DDSOntWs que permite que esta troca de experiência clínica ocorra em ambiente compartilhado e distribuído, orientado a serviços, e independente de plataforma, promovendo assim a interoperabilidade técnica. Assim, um CDSS é projetado na forma de serviços Web. E objetivando promover a compatibilidade terminológica, um dos serviços Web da DDSOntWs desenvolvido no trabalho faz uso da *UMLS Metathesaurus* para relacionar termos de diversos vocabulários médicos com os conceitos da UMLS.

Um estudo de caso para a construção de aplicativos médicos que façam uso da solução proposta também faz-se necessário para teste a validação da metodologia apresentada. Assim, foram criados módulos de software para que entidades reguladoras, provedoras e requisitantes de apoio diagnóstico possam trocar experiências relacionadas a patologia de dor abdominal aguda. Podem ser consideradas entidades reguladoras profissionais de um conselho de medicina ou mesmo de um centro de referência de uma especialidade. Entidades provedoras são exemplificadas como qualquer profissional com experiência em uma especialidade, podendo ser uma junta médica. São exemplos de entidades requisitantes qualquer profissional que deseja ser

apoiado por um entidade provedora durante uma decisão diagnóstica.

Neste contexto, os objetivos específicos deste trabalho, bem como os passos realizados para o desenvolvimento do mesmo são:

- Estudar os problemas de interoperabilidade técnica em ambientes computacionais distribuídos de Sistemas de Informação em Saúde que precisam trocar informações relacionadas ao suporte à decisão diagnóstica;
- Compreender as soluções existentes para promover a interoperabilidade técnica em computação distribuída, em especial as especificações relacionadas a serviços Web e UDDI, de modo a justificar seu uso neste trabalho;
- Examinar os possíveis problemas de interoperabilidade semântica especificamente relacionadas a incompatibilidade terminológica em Sistemas de Informação em Saúde que compartilham dados para a decisão diagnóstica;
- Investigar alguns sistemas de terminologia na área da saúde e mostrar como a *UMLS Metathesaurus* consegue promover a compatibilidade terminológica, fundamentando assim seu uso neste trabalho;
- Estudar os possíveis problemas de interoperabilidade semântica especificamente relacionadas a incompatibilidade ontológica em Sistemas de Informação em Saúde que trocam informações para apoiar a decisão diagnóstica;
- Examinar as ontologias na área médica existentes na literatura, e justificar o uso da *UMLS Semantic Network* como uma extensão desta na criação da ontologia SSDOnt;
- Analisar de que forma as linguagens ontológicas RDF, RDFS e OWL da Web Semântica propõem promover a interoperabilidade semântica na Web;
- Compreender metodologias para a definição, modelagem e construção da DDSOnt;
- Implementar a JDDSOnt de modo a permitir que softwares clínicos possam criar documentos XML conforme a DDSOnt;
- Implementar o DDSOntWs para suportar a troca de informações para apoiar a decisão diagnóstica entre as aplicações médicas em ambiente distribuído e orientado a serviços;
- Criar um caso de uso em uma patologia clínica onde uma entidade reguladora possa definir sinais e diagnósticos, uma entidade provedora possa criar casos clínicos, gerar e compartilhar uma base de conhecimento conforme as definições da entidade reguladora, e que ainda uma entidade requisitante possa ser apoiada na decisão diagnóstica por bases de conhecimento das entidades provedoras. O caso de uso é construído utilizando-se dos softwares JDDSOnt e DDSOntWs.

1.4 Estrutura da tese

Este trabalho está organizado da seguinte forma:

O Capítulo 1 apresentou o estado da arte relacionado ao assunto do trabalho, as motivações e os objetivos gerais e específicos.

O Capítulo 2 apresenta os problemas relacionados a interoperabilidade técnica em ambientes computacionais distribuídos de Sistemas de Informação em Saúde através de um cenário de troca de informações entre aplicações médicas. Na sequência, é discutido como as tecnologias e especificações de serviços Web procuram promover a interoperabilidade técnica no contexto da computação distribuída.

O Capítulo 3 discute os problemas de incompatibilidade terminológica relacionados a interoperabilidade semântica em Sistemas de Informação em Saúde através de um cenário de compartilhamento de informações médicas. Em seguida, são apresentados alguns sistemas de terminologia na área da saúde, como o ICD, o SNOMED CT, e o MeSH. Na sequência, mostra-se como a *UMLS Metathesaurus* consegue promover a compatibilidade terminológica.

O Capítulo 4 expõe os problemas de incompatibilidade ontológica associados a interoperabilidade semântica em Sistemas de Informação em Saúde através de um cenário de troca de informações médicas. São apresentados conceitos relacionados a ontologia, e discutidos exemplos de ontologias na área médica, como a *UMLS Semantic Network*, a *Galen Common Reference Model*, e *HL7 RIM*. Em seguida, é analisada de que forma a iniciativa Web Semântica da W3C propõe promover a interoperabilidade semântica na Web com as especificações RDF, e OWL.

O Capítulo 5 mostra como a ontologia DDSOnt foi definida e projetada com a metodologia 101, e como esta foi estendida da *UMLS Semantic Network*. Ainda, é apresentada a construção dos softwares JDDSOnt e DDSOntWs. Ainda, é detalhado como um dos serviços Web utiliza-se das funcionalidades da *UMLS Metathesaurus*.

O Capítulo 6 discorre a respeito do estudo de caso em dor abdominal aguda desenvolvido para demonstrar como os softwares JDDSOnt e DDSOnt podem ser utilizados em aplicações da área de saúde que precisam ser apoiados por uma decisão diagnóstica. São apresentados os módulos de apoio para as entidades reguladora, provedora e requisitante.

O Capítulo 7 tece as considerações finais do trabalho, apresenta trabalhos relacionados, discute trabalhos futuros e mostra as contribuições desta tese para a área de Informática em Saúde.

Interoperabilidade Técnica em Sistemas de Informação em Saúde Distribuídos

2.1 Considerações Iniciais

Este capítulo tem como objetivo apresentar problemas relacionados a interoperabilidade técnica em ambiente distribuído durante a troca de informações clínicas entre sistemas de informação em saúde, especificamente para apoiar a decisão diagnóstica. Ainda, são discutidas como as especificações SOAP, WSDL e UDDI relacionadas a serviços Web procuram promover este tipo de interoperabilidade. SOAP é um protocolo de comunicação escrito em XML, de padrão aberto e não proprietário. WSDL é uma aplicação XML para descrever serviços Web, enquanto que UDDI é uma especificação utilizada na publicação e descoberta de serviços Web. Este Capítulo é finalizado apresentando como os estudos relacionados a estas especificações contribuíram no cumprimento dos objetivos deste trabalho.

2.2 Cenário Médico com Problemas de Interoperabilidade Técnica

Para que médicos possam remotamente trocar informações de apoio à decisão diagnóstica através da Internet, faz-se necessário um ambiente computacional que suporte a computação distribuída, ou seja, um ambiente computacional formado de componentes de software, ou objetos encapsulados, que se comunicam entre si e são executados em diferentes sistemas computacionais. Esses componentes ou objetos são na verdade procedimentos remotos que podem realizar atividades tais como: a criação de uma base de conhecimento clínica, o apoio durante a decisão diagnós-

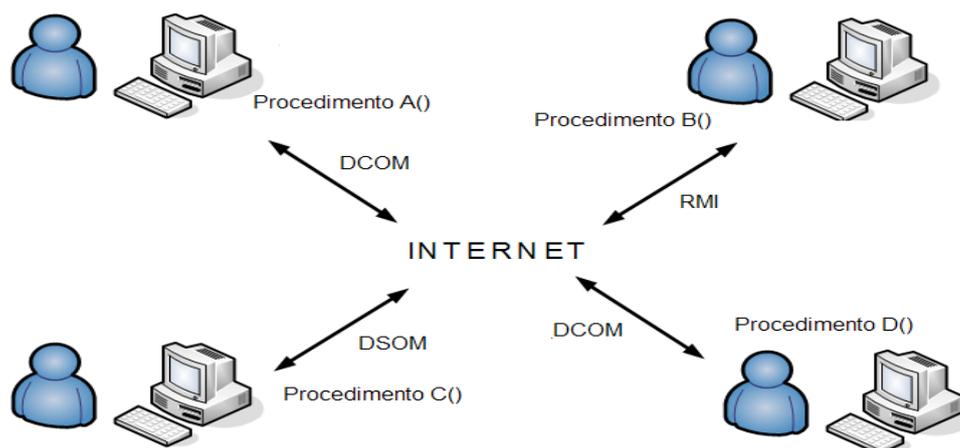


Figura 2.1: Cenário ilustrando problemas de interoperabilidade

tica, ou ainda a disponibilização ou descoberta de bases de conhecimento de uma determinada síndrome. Entretanto, este tipo de ambiente computacional distribuído caracteriza-se como sendo fortemente acoplado quando os procedimentos remotos são criados ou executados por protocolos de comunicação proprietários de computação distribuída. O forte acoplamento significa a criação de programas chamados de 'pontes', programas necessários para comunicação entre procedimentos remotos implementados por protocolos de comunicação diferentes. Se por acaso algum protocolo proprietário sofrer alteração em sua especificação durante o funcionamento do sistema, essa ponte deve sempre ser refeita. A Figura 2.1 ilustra um cenário típico de uso de diferentes protocolos de comunicação proprietários em computação distribuída sendo utilizados por médicos em suas aplicações clínicas. Nesta Figura são mostrados os seguintes protocolos: DCOM da Microsoft, DSOM IBM, e RMI/IIOP da Sun. Estes protocolos podem ainda tornar o sistema clínico distribuído dependente de sistema operacional, como o caso do protocolo DCOM, ou então dependente de linguagem de programação, como o protocolo RMI.

O cenário apresentado anteriormente ainda apresenta a ausência de dois recursos quando se trata de uso de procedimentos remotos por aplicações clínicas distribuídas que precisam trocar informações de modo a apoiar a decisão diagnóstica. O primeiro deles é a ausência de um mecanismo que possa facilitar a chamada de um procedimento clínico através da descrição do mesmo. Essa descrição pode envolver o nome do procedimento clínico, a quantidade e o tipo de parâmetros de entrada, e se existe algum parâmetro e o tipo de parâmetro de saída do procedimento. O segundo recurso é a falta de um mecanismo que possa disponibilizar ou publicar os procedimentos de um sistema clínico em algum local onde possam posteriormente ser encontrados por possíveis entidades interessadas. Além disso, esse mesmo mecanismo deveria oferecer às entidades interessadas uma forma de procurar e recuperar informações destes procedimentos disponibilizados.

No entanto, o cenário apresentado anteriormente pode ser alterado de modo a solucionar as limitações discutidas. Para isso, a solução proposta neste trabalho é substituir os protocolos de comunicação por um único protocolo XML de comunicação de padrão aberto e não proprietário, chamado SOAP, e ainda adicionar as especificações de padrão aberto WSDL e UDDI, que permitem descrever procedimentos, e publicá-los e recuperá-los, respectivamente. O resultado da substituição e das adições é ilustrado

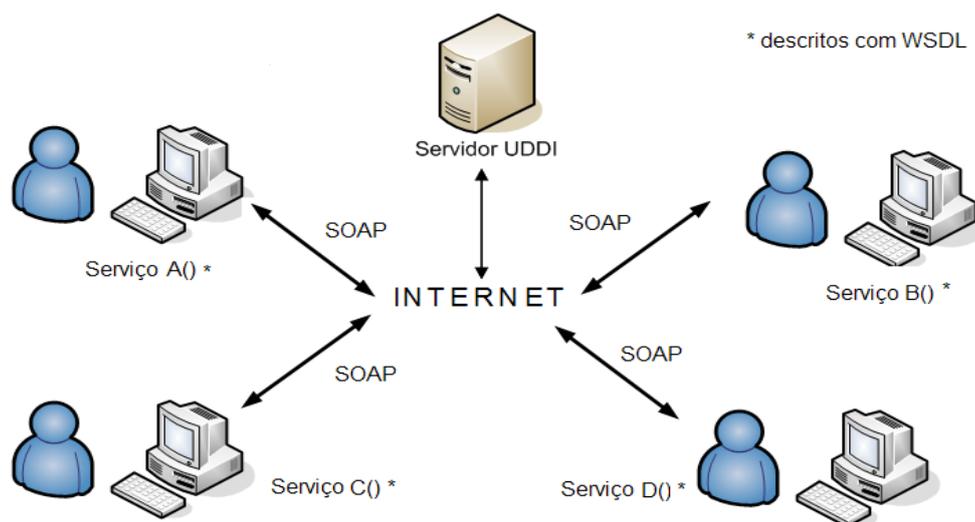


Figura 2.2: Cenário com especificações de serviços Web.

na Figura 2.2. Assim, as aplicações clínicas que estiverem em conformidade com as especificações SOAP, WSDL e UDDI podem se comunicar independentemente de arquitetura computacional ou linguagem de programação, deixando também o sistema fracamente acoplado. As especificações SOAP, WSDL e UDDI fazem parte da iniciativa de Serviços Web e são apresentadas na seção a seguir.

2.3 Serviços Web

Serviços Web (Booth et al. 2006) é uma iniciativa proposta pela W3C para padronizar a forma com que aplicações distribuídas na Web possam se comunicar, independentemente de plataforma computacional ou linguagem de programação, desde que estas aplicações obedeçam as especificações estabelecidas pelo consórcio. Os serviços Web são caracterizados por sua interoperabilidade e extensibilidade e, principalmente por fazerem uso da XML, linguagem esta introduzida neste documento na sub-seção 2.3.2, como linguagem para estruturação das mensagens trocadas pelas aplicações (Haas 2006). De modo a fazer com que essa comunicação ocorra, é necessário que alguns padrões comuns sejam adotados de forma que as aplicações possam entender o que uma comunica a outra. Atualmente a organização responsável para promover a interoperabilidade entre as aplicações na Web é conhecida como a WS-I.org (*Web Services Interoperability Organization*)¹.

2.3.1 Funcionamento de Arquitetura de Serviços Web

A Figura 2.3 a seguir apresenta um resumo dos passos necessários para que uma entidade requisitante, possivelmente uma aplicação de um médico que precisa ser apoiado por uma resposta diagnóstica, possa convidar uma entidade provedora, provavelmente outra aplicação de outro médico que possui sua base de conhecimento compartilhada, para utilizar seus serviços. O passo (1) apresenta as entidades requisitante e provedora conhecendo-se uma a outra, ou pelo menos uma conhecendo a

¹<http://www.ws-i.org>

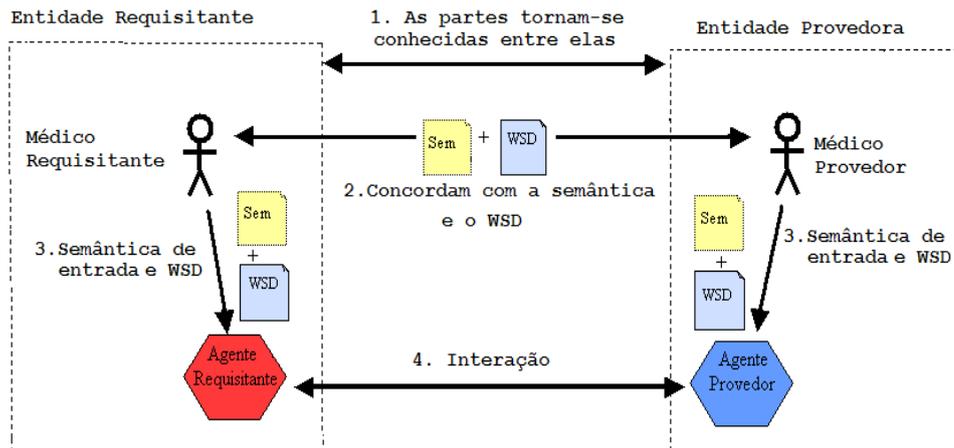


Figura 2.3: Funcionamento da arquitetura serviços Web.

outra. O passo (2) ilustra a entidade requisitante e a entidade provedora firmando um acordo, de alguma maneira, com as descrições e semânticas do(s) serviço(s) em questão que irão dar as regras da interação entre os agentes das duas entidades. O passo (3) exemplifica a realização dos serviços de descrição e de semântica. Essa realização ocorre por parte dos agentes da entidade requisitante e da entidade provedora juntas. Finalmente, o passo (4) apresenta a troca de mensagens entre os agentes da entidade requisitante e da entidade provedora, e portanto, desempenhando alguma tarefa em nome de suas entidades. Esse último passo significa que a troca de mensagens com o agente provedor representa a manifestação concreta de interação com a entidade requisitante do serviço Web.

A Figura 2.4 apresenta uma ilustração das tarefas necessárias para que uma entidade requisitante, possivelmente uma aplicação de um médico precisando criar uma base de conhecimento a partir de seus casos clínicos, descubra os serviços que uma entidade provedora forneça. Provavelmente, a entidade provedora é utilizada por um médico que possui um serviço que cria e retorna uma base de conhecimento, e neste contexto a entidade requisitante não conhece as descrições e semânticas do serviço do agente provedor. Neste caso, a entidade requisitante precisa descobrir este serviço. ‘Descobrir’ é o ato de localizar um serviço Web apropriado. Isso é feito através da tentativa de se encontrar uma máquina processadora de descrição de um serviço Web que possa ser inicialmente desconhecido e que coincida com alguns critérios funcionais. Um serviço de descoberta de serviços Web é um serviço que facilita o processo de pesquisa de serviços Web. Esse tipo de serviço pode ser desempenhado por um indústria, ou um agente controlador, por exemplo, em casos aplicados a Medicina, a um conselho ou associação regional, nacional ou internacional, ou ainda por uma junta médica especializada na área.

O passo (1a) da Figura 2.4 apresenta a entidade provedora registrando o descritor de serviço Web (WSD) e o descritor funcional associado (FD) na máquina servidora de serviço de descoberta. O descritor funcional é uma estrutura contendo a descrição das funcionalidades do serviço que a entidade provedora está oferecendo, provavelmente informando que o serviço, por exemplo, cria uma base de conhecimento relacionada a síndrome de dor abdominal aguda. O passo (1b) ilustra a situação em que a entidade requisitante fornece os critérios ao serviço de descoberta para selecionar uma descrição de serviço Web baseado em sua descrição funcional, capacidades e outras

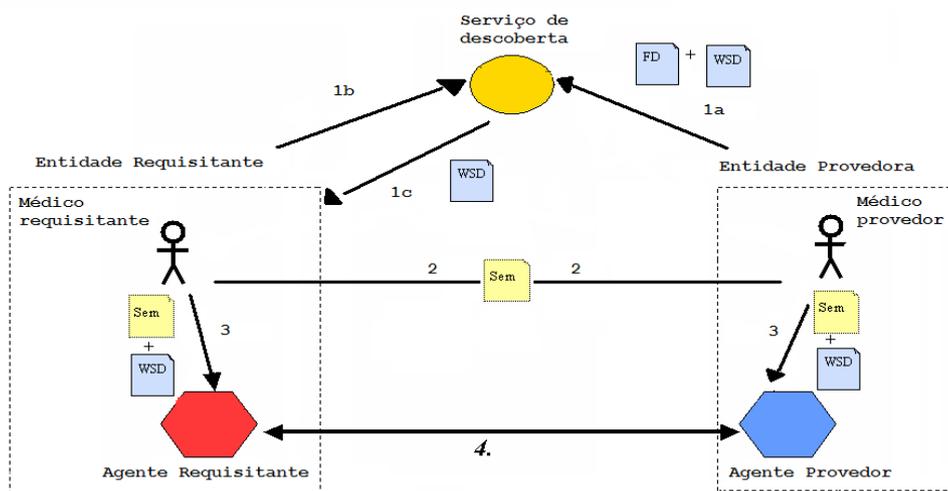


Figura 2.4: Funcionamento da arquitetura serviços Web com serviço de publicação e descoberta.

características potenciais relacionadas. No contexto deste projeto, possivelmente o critério seria solicitar serviços que criem bases de conhecimento relacionados a dor abdominal aguda. Um aspecto interessante nesta pesquisa pode estar relacionado a serviços tendo certas funcionalidades ou semântica desejada.

Contudo, pode ser possível especificar um critério não-funcional relacionado com o agente provedor, como o nome da entidade provedora, um critério de confiança ou desempenho, ou ainda um critério relacionado com a entidade provedora, como o nome da faculdade, do departamento, do hospital, ou da clínica médica que oferece o serviço. Finalmente, o passo (1c) apresenta o serviço de descoberta retornando uma ou mais descrições de serviço Web (uma referências para eles) relacionado ao critério especificado. A partir do passo (2), a entidade requisitante e provedora comunicam-se diretamente, conforme já explicado anteriormente.

O funcionamento da arquitetura serviços Web está relacionado com suas especificações. Toda comunicação que ocorre entre entidade requisitante, entidade provedora e serviço de descoberta ocorre através do protocolo SOAP, que será apresentado na sub-seção 2.3.3. Os descritores de serviços Web (WSD) são realizados através da linguagem WSDL, a ser descrita na sub-seção 2.3.4. O serviço de descoberta é especificado através do UDDI, que será apresentado na sub-seção 2.3.5. Finalmente, os descritores funcionais são registrados no UDDI nos chamados *UDDI Registry*, e na forma da estrutura de dados *tModel*, a serem descritos nas sub-seções 2.3.5.1 e 2.3.5.2, respectivamente. A linguagem XML, presente nas especificações SOAP, WSDL e UDDI, é introduzida a seguir.

2.3.2 XML

A linguagem de marcação XML² é definida com sendo uma meta-linguagem, ou seja, uma linguagem que possibilita a criação de outras linguagens XML. Assim, os desenvolvedores podem criar seus próprios elementos de marcação, de acordo com

²<http://www.w3.org/xml>

sua aplicação, de modo a refletir a importância do conteúdo e da estrutura da informação.

Para um documento XML ser definido como bem-formatado, este deve obedecer algumas regras, como: todos elementos que aparecem no documento precisam de um elemento de marcação de abertura e de um elemento de marcação de fechamento. Os elementos de marcação devem ser criados considerando-se as características de XML, como por exemplo, é uma ligação sensível ao tamanho da caixa da fonte (*case sensitive*); os elementos de marcação devem ser aninhados, ou seja, caso o elemento B seja aberto no escopo do elemento A, o elemento B deve ser fechado primeiro que o A; todo documento XML deve ter um elemento chamado raiz; e finalmente o conteúdo dos atributos devem estar sempre entre aspas duplas. O documento XML apresentado a seguir é considerado bem-formatado. Esse exemplo ilustra um documento contendo informações de sinais clínicos relacionados a síndrome de dor abdominal aguda, juntamente com seu diagnóstico.

```
1 <?xml version="1.0"?>
2 <ClinicalCase>
3   <pathology> Acute Abdominal Pain </pathology>
4   <vomitus> Biliious Finding </vomitus>
5   <severities> Moderate </severities>
6   <diseaseOrSindrome> Cholecystitis </diseaseOrSindrome>
7 </ClinicalCase>
```

Num contexto de troca de documentos XML entre aplicações clínicas diferentes, para que uma aplicação possa reconhecer o documento eletrônico recebido como válido, esse documento XML precisa estar de acordo com um documento chamado de DTD, *Document Type Definition*, como veremos a seguir.

2.3.2.1 Linguagem DTD

Um documento DTD ³ define os elementos e atributos que podem aparecer em um documento XML, e ainda define quais são os elementos que são nós filhos de um elemento, e a ordem destes elementos filhos. Um exemplo de documento DTD que valida o documento XML mostrado na seção 2.3.2 é apresentado a seguir. No exemplo, é mostrado na linha 1 que o elemento *ClinicalCase* deve conter exatamente um elemento *pathology*, um elemento *vomitus*, um elemento *severities*, e um elemento *diseaseOrSindrome*.

```
1 <!ELEMENT ClinicalCase (pathology, vomitus, severities, diseaseOrSindrome)>
2 <!ELEMENT pathology (#PCDATA)>
3 <!ELEMENT vomitus (#PCDATA)>
4 <!ELEMENT severities (#PCDATA)>
5 <!ELEMENT diseaseOrSindrome (#PCDATA)>
```

Como pode ser notado no exemplo apresentado, nas linhas de 2 a 5, os tipos de dados dos elementos não podem ser melhor especificados. PCDATA indica apenas que valores alfa-numéricos são permitidos. Isso demonstra a simplicidade e a limitação de um DTD, que restringe o desenvolvedor na criação de um mecanismo para validar um documento XML. Entretanto, caso o desenvolvedor não precise de recursos mais completos durante a validação, DTD torna-se muito útil em função de sua simplicidade.

³<http://www.w3.org/2002/xmlspec/dtd/2.9/xmlspec.dtd>

2.3.2.2 XML Schema

Procurando resolver as limitações de um documento DTD, a linguagem XML SCHEMA⁴ possui todas as funcionalidades de um DTD e ainda possui recursos adicionais como: possibilidade de informar tipos de dados aos elementos e atributos, como real, inteiro, ponto-flutuante; permite a criação de TADs (Tipos Abstratos de Dados) definida pelo desenvolvedor; e ainda suporta *namespaces*, que é um recurso que possibilita que desenvolvedores possam distinguir sintaticamente elementos com o mesmo nome, mas com propósitos diferentes. O documento XML da seção 2.3.2 é válido em relação ao documento SCHEMA a seguir. Com o recurso dos elementos `<xsd:complexType>` (linha 4) e `<xsd:simpleType>` (linha 12), pode-se criar TADs, como por exemplo *ClinicalCaseType* e *diseaseOrSindromeType*. Com o uso do elemento `<xsd:sequence>`, pode-se definir que um elemento é composto por uma sequência exata de elementos (veja as linhas 5 a 10). Além disso, através dos elementos `<xsd:restriction>` e `<xsd:enumeration>`, pode-se definir os possíveis valores que um elemento pode conter (linhas 13 a 18).

```
1 <?xml version="1.0"?>
2 <xsd:schema>
3 <xsd:element name="ClinicalCase" type="ClinicalCaseType"/>
4 <xsd:complexType name="ClinicalCaseType">
5 <xsd:sequence>
6 <xsd:element name="pathology" type="xsd:string"/>
7 <xsd:element name="vomitus" type="xsd:string"/>
8 <xsd:element name="severities" type="xsd:string"/>
9 <xsd:element name="diseaseOrSindrome" type="diseaseOrSindromeType"/>
10 </xsd:sequence>
11 </xsd:complexType>
12 <xsd:simpleType name="diseaseOrSindromeType">
13 <xs:restriction base="xsd:string">
14 <xs:enumeration value="cholecystitis"/>
15 <xs:enumeration value="appendicitis"/>
16 <xs:enumeration value="pepticulceration"/>
17 <xs:enumeration value="gastroenteritis"/>
18 </xs:restriction>
19 </xsd:simpleType>
20 </xsd:schema>
```

Como pode ser visto, em um documento XML Schema pode-se especificar quais os sinais clínicos relevantes de uma síndrome, quais os valores possíveis que um sinal clínico pode apresentar, e quais os valores de diagnósticos válidos como documentos Web.

2.3.3 SOAP

SOAP, *Simple Object Access Protocol* (Mitra 2006), é um protocolo orientado o objetos que define um formato XML para as mensagens que são trocadas e utilizadas para chamadas a serviços Web ou objetos remotos em ambiente distribuído e descentralizado, utilizando para isso o protocolo HTTP. Por trabalhar sob HTTP, o protocolo SOAP utiliza a porta 80 dos servidores Web, ficando sua execução normalmente livre de regras de *firewalls*. Este é considerado uma característica vantajosa sobre propostas semelhantes pois SOAP pode ser utilizado de modo transparente em servidores que liberam a porta 80 por definição. Por outro lado, o uso da porta 80 pode ser um fator negativo do ponto de vista de segurança, aspecto não coberto neste trabalho por fugir do escopo do mesmo.

Uma mensagem SOAP é um documento XML contendo dois elementos obrigató-

⁴<http://www.w3.org/XML/Schema>

rios, que são: o *Envelope* e o *Body*, e dois elementos opcionais, *Header* e *Fault*. O documento XML mostrado a seguir ilustra um exemplo de uma mensagem SOAP. Cada um destes elementos é brevemente explicado na sequência.

```
1 <?xml version="1.0"?>
2   <soap:Envelope
3     xmlns:soap="http://www.w3.org/2001/12/soap-envelope"
4     soap:encodingStyle="http://www.w3.org/2001/12/soap-encoding">
5     <soap:Header>
6       ...
7     </soap:Header>
9     <soap:Body>
10      ...
11      ...
12      <soap:Fault>
13        ...
14      </soap:Fault>
16    </soap:Body>
17  </soap:Envelope>
```

2.3.3.1 Elemento Envelope

O elemento obrigatório *Envelope* é chamado de elemento raiz de uma mensagem SOAP, já que este deve envolver os demais elementos. É importante notar o uso dos *namespaces soap* (linha 3) e *encoding*, responsáveis em definir, respectivamente, que o elemento *Envelope* é um elemento SOAP, e quais são os tipos de dados utilizados no documento.

2.3.3.2 Elemento Header

O elemento opcional *Header*, caso apareça, deve suceder exatamente o elemento *Envelope*. Esse elemento opcional pode conter informações específicas da aplicação, como uma autenticação ou uma assinatura digital para ser apoiado durante a decisão diagnóstica, por exemplo.

2.3.3.3 Elemento Body

O elemento obrigatório *Body* contém o conteúdo da mensagem que será entregue a uma entidade receptora, seja ela requisitante ou provedor de um serviço de resposta diagnóstica. Caso a entidade receptora da mensagem seja provedora, ou seja, está-se requisitando por um serviço, o elemento *Body* pode ser construído como no documento a seguir. No documento, percebe-se um elemento que indica o serviço *Web DiagnosisDecisionSupport* sendo chamado (linha 2 – *getDiagnosisDecisionSupport*). Os sub-elementos de *Body* representam os valores passados durante a chamada do serviço (linhas de 4 a 6).

```
1 <soap:Body>
2   <ddson: GetDiagnosisDecisionSupport
3     xmlns:ddson="http://143.107.220.180/DDSONt">
4     <ddson: pathology> Acute Abdominal Pain </ddson: pathology>
5     <ddson: vomitus> Biliious Finding </ddson: vomitus>
6     <ddson: severities> Biliious Finding </ddson: severities>
7   </ddson: GetDiagnosisDecisionSupport>
8 </soap:Body>
```

Já o documento a seguir ilustra um exemplo de um documento parcial SOAP sendo entregue à entidade que requisitou o serviço. No documento, o elemento da

linha 2 identifica que a mensagem é de retorno ou de resposta (*GetDiagnosisDecisionSupportResponse*). Note que o conteúdo da resposta de apoio à decisão diagnóstica é *cholecystitis* (linhas de 4 a 6).

```
1 <soap:Body>
2   <ddsonst: GetDiagnosisDecisionSupportResponse
3     xmlns: ddsont="http://143.107.220.180/DDSOnt">
4     <ddsonst: diseaseOrSindrome>
5       cholecystitis
6     </ddsonst: diseaseOrSindrome>
7   </ddsonst: GetDiagnosisDecisionSupportResponse>
8 </soap:Body>
```

2.3.3.4 Elemento Fault

Finalmente, o elemento opcional *Fault*, localizado dentro do elemento *Body*, contém informações das possíveis situações de erro de uma mensagem SOAP. São informações que podem estar contidas no elemento *Fault* relativas a uma falha, como: código, descrição, quem a causou e problemas no elemento *Body*.

2.3.4 WSDL

WSDL, *Web Services Description Language* (Chinnici et al. 2006), é uma linguagem de marcação XML para descrever interfaces de serviços Web, a localização destas, bem como os métodos e operações expostas. O documento a seguir ilustra a estrutura de um documento WSDL. Cada um destes elementos é brevemente explicado na sequência.

```
1 <definitions
2   <!-- Definição dos tipos de dados utilizados -->
3   <message> </message>
4   <portType> </portType>
5   <binding> </binding>
6   <services> </services>
7 </definitions>
```

2.3.4.1 Elemento definition

O elemento raiz *definitions* define e concentra os outros elementos WSDL. É nesta área onde são definidos os tipos de dados ou *namespaces* utilizados em todo o documento. Exemplos de *namespaces* são: *wSDL*, definido pela URL `xmlns:wSDL="http://schemas.xmlsoap.org/wSDL/"`, e utilizado para qualificar os elementos da linguagem WSDL, *wSDLsoap*, definido pela URL `xmlns:wSDLsoap="http://schemas.xmlsoap.org/wSDL/soap/"`, e usada na qualificação dos elementos WSDL relacionados à ligação dos serviços com o endereço da localização destes, e *xSD*, definido pela URL `xmlns:xSD="http://www.w3.org/2001/XMLSchema"`, e utilizada para qualificar os tipos de dados de XML Schema, como *string*, *float*, e *int*.

2.3.4.2 Elemento portType

O elemento *portType* é considerado o mais importante dos elementos WSDL pois é ele quem descreve as interfaces ou funções ou ainda métodos expostos por um serviço Web, bem como quais as mensagens estão envolvidas. Normalmente, para cada

método exposto, existem duas mensagens, a de requisição (*request*) e a de resposta (*response*). O documento WSDL parcial a seguir ilustra um exemplo de construção do elemento *portType*. Na linha 1 são definidos o nome do método e os parâmetros de entrada do mesmo. Das linhas de 2 a 4 são definidas as operações de requisição e resposta relacionadas ao método *DiagnosisDecisionSupport*.

```

1 <portType name="DiagnosisDecisionSupport" parameterOrder="pathology_vomitus_severities">
2   <operation name="getDiagnosisDecisionSupport" >
3     <input message="getDiagnosisDecisionSupportRequest"/>
4     <output message="getDiagnosisDecisionSupportResponse"/>
5   </operation>
6 </portType>

```

2.3.4.3 Elemento message

O elemento *message* fica responsável em definir todas as operações declaradas na definição dos métodos. O documento a seguir ilustra a construção deste elemento. A operação de requisição possui 3 parâmetros do tipo *string* (linhas de 1 a 5), enquanto a operação de resposta retorna um valor também do tipo *string* (linhas de 6 a 8).

```

1 <message name="getDiagnosisDecisionSupportRequest">
2   <part name="pathology" type="xs:string"/>
3   <part name="vomitus" type="xs:string"/>
4   <part name="severities" type="xs:string"/>
5 </message>
6 <message name="getDiagnosisDecisionSupportResponse">
7   <part name="diseaseOrSindrome" type="xs:string"/>
8 </message>

```

2.3.4.4 Elemento binding

O objetivo do elemento *binding* é descrever a forma de acesso aos serviços Web definidos através do protocolo SOAP. Assim, esse elemento conecta os serviços Web ao protocolo SOAP por meio de um nome. O documento a seguir ilustra um exemplo que associa o serviço Web *DiagnosisDecisionSupport*, assim como as mensagens de entrada e saída, ao *bind* *DDSOntImpl*.

```

1 <binding type="DDSOnt" name="DDSOntImpl">
2 <soap:binding style="RPC" transport="http://schemas.xmlsoap.org/soap/http" />
3 <operation>
4   <soap:operation
5     soapAction="http://143.107.220.180/DDSOnt/getDiagnosisDecisionSupport"/>
6   <input
7     name="DiagnosisDecisionSupportRequest">
8     <soap:body use="literal"/>
9   </input>
10  <output
11    name="DiagnosisDecisionSupportResponse">
12    <soap:body use="literal"/>
13  </output>
14 </operation>
15 </binding>

```

2.3.4.5 Elemento service

Finalmente, o elemento *service* é responsável em definir o ponto de acesso (*endpoint*) dos serviços Web, ou ainda, por associar o nome do *bind* ao local onde estão disponíveis os serviços. O documento a seguir apresenta um exemplo. O serviço Web *DiagnosisDecisionSupport* pode ser obtido no arquivo *DDSOntImpl* localizado em <http://143.107.220.180:8080/DDSOnt>, por meio do *bind* *DDSOntImpl*.

```
1 <wsdl:service name="DDSOntImplService">
2   <wsdl:port binding="impl:DDSOntImplSoapBinding" name="DDSOntImpl">
3     <wsdlsoap:address location="http://143.107.220.180:8080/DDSOnt/DDSOntImpl.jws" />
4   </wsdl:port>
5 </wsdl:service>
```

2.3.5 UDDI

A especificação ou *framework* UDDI ([UDDI 2007](#)) define um serviço de registro para serviços Web. O serviço de registro UDDI é um serviço Web que gerencia informações sobre provedores de serviço, implementações de serviços e metadados de um serviço. Provedores de serviço podem utilizar UDDI para publicar os serviços que estes oferecem. Os consumidores de serviço podem utilizar um UDDI para descobrir serviços que estejam de acordo com seus requisitos e para obter os meta-dados necessários para consumir esses serviços.

Um cenário típico de uso do UDDI, e aplicado ao contexto deste trabalho, é uma junta médica especializada em dor abdominal aguda que implementa um serviço Web que retorna uma resposta diagnóstica para um caso clínico duvidoso a partir de uma base de conhecimento. Assim, essa junta médica deve registrar a implementação deste serviço em um UDDI. Entidades interessadas nesse tipo de serviço podem descobrir essa implementação fazendo uma consulta ao UDDI, e verificando quais as informações necessárias para o uso do serviço, como o nome, os sinais clínicos necessários ou ainda as possíveis respostas diagnósticas. Em seguida, o interessado faz uso do serviço descoberto diretamente com a máquina da junta médica que o hospeda.

Outro cenário também relacionado ao contexto deste trabalho é a possibilidade de um médico ser apoiado na decisão diagnóstica por várias juntas médicas, utilizando a mesma especificação de serviço. Isso quer dizer que a interface do serviço de qualquer junta médica registrada no UDDI será a mesma, com o mesmo nome de serviço, a mesma síndrome, e os mesmos sinais e diagnósticos. Um situação contrária seria que cada junta médica implementasse o mesmo serviço ao seu modo, contrariando o padrão de utilização. Voltando à situação correta, o *software* do médico interessado utiliza as várias implementações do serviço encontrado no UDDI sempre do mesmo jeito, bastando alterar o endereço de localização do serviço quando da chamada deste.

Para a implementação do cenário anterior, é necessário que uma organização ou uma entidade reguladora, como representantes de um conselho regional, ou associações médicas, ou especialidades médicas, defina os descritores funcionais e não-funcionais dos serviços Web, na forma de *tModels*, explicado na sub-seção [2.3.5.2](#) e os disponibilize no UDDI. Ainda, é preciso que uma junta médica, antes de implementar os serviços Web de interesse, consulte o UDDI e verifique se existe um padrão de implementação do mesmo.

2.3.5.1 UDDI Registry

Um servidor UDDI organiza e relaciona o registro dos serviços utilizando estrutura de diretórios, chamados de *Registry*. Essa entidade visa facilitar a descoberta dos serviços e suas implementações. Cada unidade inserida em um *Registry* é chamada de *tModel*.

2.3.5.2 *tModels*

O *tModel* é uma estrutura de dados que pode conter até 8 elementos, descritos como:

TModelKey : chave única no servidor UDDI para um *tModel*;

AuthorizedName : nome do usuário que está publicando o *tModel* no servidor;

Operator : nome do administrador do servidor responsável pela manutenção técnica do *tModel*;

Name : nome do *tModel*. Deve ser inteligível, um nome que tenha sentido para pessoas que estão procurando por um tipo de serviço;

Description : descrição *tModel*. É possível uma descrição por idioma;

OverviewDoc : referência externa à descrição do serviço. Normalmente a referência é para um documento WSDL localizado em um servidor remoto;

IdentifierBag e CategoryBag : oferecem a possibilidade de associar categorias ou classificações às entidades do *Registry* através de pares (nome, valor), onde nome é a identificação da classificação (*identifierBag*) e valor é a categoria ou significado da classificação (*categoryBag*). Exemplos na área médica utilizando o NAICS como referência são os pares (6215, *Medical and Diagnostic Laboratories*), (621512, *Diagnostic Imaging Centers*), e (621111, *Offices of Physicians (except Mental Health Specialists)*).

Com esses elementos do *tModel*, é possível então que um órgão regulador possa definir os descritores de serviços Web de modo que os mesmos possam ser encontrados e implementados por juntas médicas, ou ainda encontrados e utilizados por médicos interessados.

A Figura 2.5 ilustra o relacionamento entre as entidades reguladora, provedora e requisitante de decisão diagnóstica. De acordo com a figura:

- As entidades reguladora, provedora, requisitante, e o servidor UDDI comunicam-se através do protocolo SOAP;
- A entidade reguladora é responsável por definir e registrar no servidor UDDI um *tModel* para cada cada serviço Web. É através deste *tModel* que as entidades provedoras e requisitantes tomam conhecimento sobre qual síndrome, e quais os sinais e os diagnósticos clínicos são utilizados nos serviços Web;
- A entidade provedora deve consultar por *tModels* de interesse e registrados no servidor UDDI, e a partir da definição destes, implementar os serviços oferecidos pela entidade reguladora. Após as implementações, a entidade provedora deve registrá-las no UDDI;
- A entidade requisitante precisa consultar no servidor UDDI os *tModels* de seu interesse, e a partir do resultado consultar as implementações do mesmo. Assim, a entidade requisitante pode utilizar os serviços implementados pela entidade provedora, e ser apoiado por esta nas atividades de consulta por uma base de conhecimento, de obtenção por uma resposta diagnóstica, e eventualmente de criação de uma própria base de conhecimento.

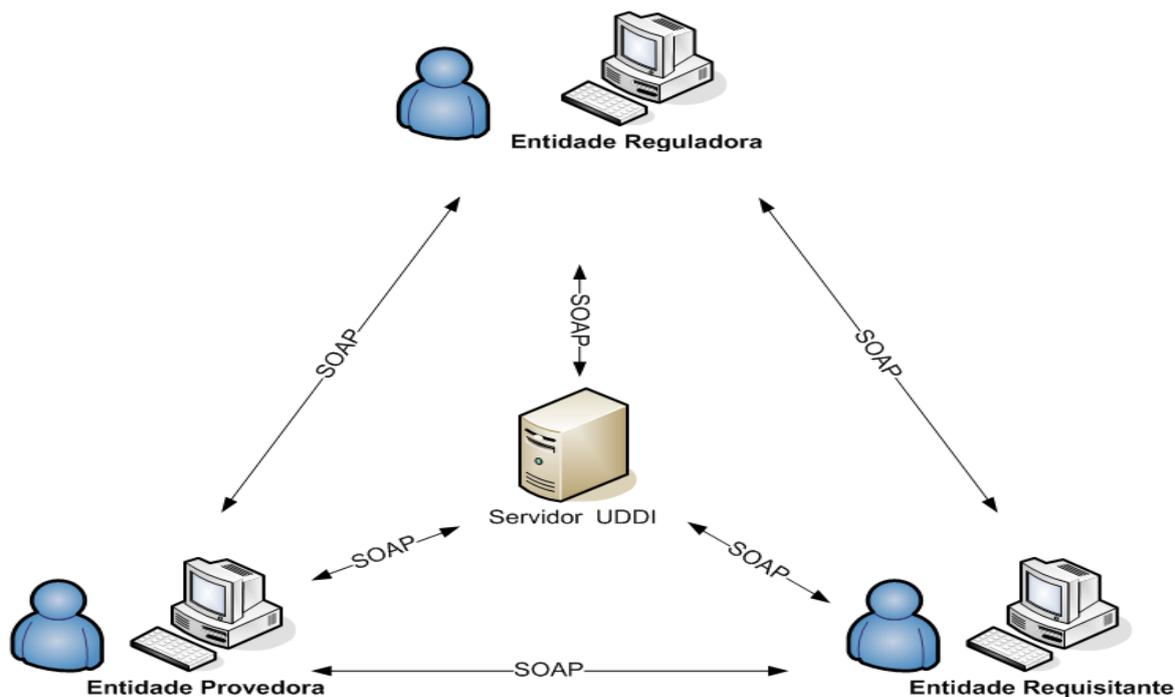


Figura 2.5: Relacionamento entre as entidades reguladora, provedora e requisitante

2.4 Serviços Web Semânticos

As especificações SOAP, WSDL, e UDDI da arquitetura serviços Web apresentadas neste Capítulo não tem como propósito fornecer a semântica dos dados, a lógica do negócio e nem a definição de uma sequência de processos. Assim, a iniciativa da criação de serviços Web Semânticos⁵ é a de trazer benefícios no tratamento da informação, facilitando operações de anúncio, de descoberta, de representação, de interpretação, e a de manutenção dos dados relativos aos serviços na Web. Dessa forma, com o uso de serviços Web Semânticos neste projeto, um médico pode melhor representar o significado semântico dos seus serviços Web sendo publicados, e um outro médico pode obter melhores resultados na busca por serviços de interesse.

2.4.1 OWL-S: OWL Web Ontology Language for Services

OWL-S é uma especificação submetida à W3C em 2005 para a criação de uma linguagem de descrição semântica de serviços Web utilizando a tecnologia OWL da Web Semântica, esta apresenta na seção 4.4.2. A linguagem OWL-S possui um conjunto de classes e propriedades para descrever as características e capacidades dos serviços Web, de maneira não ambígua e de forma a serem processados automaticamente por computador. Essa linguagem de marcação pode facilitar e aumentar a automação das tarefas relacionadas a serviços Web, como a descoberta, a execução, a composição e a interoperabilidade (Barstow et al. 2004). Para isso, a ontologia OWL-S representa o conhecimento dos serviços Web em 3 facetas (Martin et al. 2004), que são: o perfil do

⁵<http://www.w3.org/2002/ws/swsig/>

serviço; o serviço na forma de um processo; e, o fundamento do serviço, apresentados a seguir.

O documento a seguir ilustra um serviço (*DiagnosisDecisionSupport*) sendo associado a 3 arquivos diferentes: o primeiro para se definir o perfil, o segundo para modelá-lo em forma de serviço, e ainda um terceiro para se definir o fundamento do serviço. Para isso, a classe *Service* disponibiliza três propriedades, que são: *presents*, *describedBy* e *supports*.

```
1 <service:Service
2   rdf:ID="DDSOnt_DiagnosisDecisionSupport">
3   <service:presents
4     rdf:resource="http://143.107.220.180/DDSOnt/DDSPerfil.owl#Profile_DDSONt_DDS"/>
5   <service:describedBy
6     rdf:resource="http://143.107.220.180/DDSOnt/DDSProcesso.owl#Processo_DDSONt_DDS"/>
7   <service:supports
8     rdf:resource="http://143.107.220.180/DDSOnt/DDSFundamento.owl#Fundamento_DDSONt_DDS"/>
9 </service:Service>
```

2.4.1.1 Perfil do Serviço

O perfil do serviço é utilizado para melhorar a publicação e a descoberta dos serviços. Assim, descrever o perfil do serviço significa exibir os serviços oferecidos pelo provedor e expor os serviços necessários para o requisitante. Na linguagem OWL-S, para informar que um serviço possui um perfil, é necessário que a propriedade *presents* da classe *Service* tenha como conteúdo uma instância da classe *Profile*.

As descrições do perfil de um serviço Web são realizadas na forma de três tipos de informação (Martin et al. 2004):

1. a organização que provê o serviço;
2. as funcionalidades que o serviço dispõe, e
3. as características do serviço.

Essa forma de se descrever o perfil do serviço se parece muito com o modelo adotado pela especificação UDDI, descrita na seção 2.3.5. Entretanto, outras 28 formas de registro de serviços Web também existem na literatura (Sycara et al. 1999). Assim, descrever o perfil de um serviço Web utilizando OWL-S torna-o independente da forma de registro, e portanto potencialmente pode-se registrá-lo em qualquer um dos 28 tipos de registros.

A informação relativa ao provedor do serviço é relacionada ao nome do serviço, a uma descrição textual do serviço e ao contato de um administrador. Esse administrador pode ser o operador que mantém o serviço em execução ou um representante que pode fornecer maiores informações do serviço. Na linguagem OWL-S, essas informações são representadas pelas respectivas propriedades da classe *Profile*: *serviceName*, *textDescription*, e *contactInformation*, como mostra a Figura 2.6. O conteúdo da propriedade *contactInformation* não é especificado pela ontologia OWL, tendo o usuário que utilizar uma outra ontologia, como VCard ou FOAF. O documento a seguir ilustra um exemplo:

```
1 <profile:serviceName>
2   Acute Abdominal Pain Diagnosis DecisionSupport
3 </profile:serviceName>
4 <profile:textDescription>
5   Este serviço recebe um conjunto de valores de sinais clínicos relacionados
```

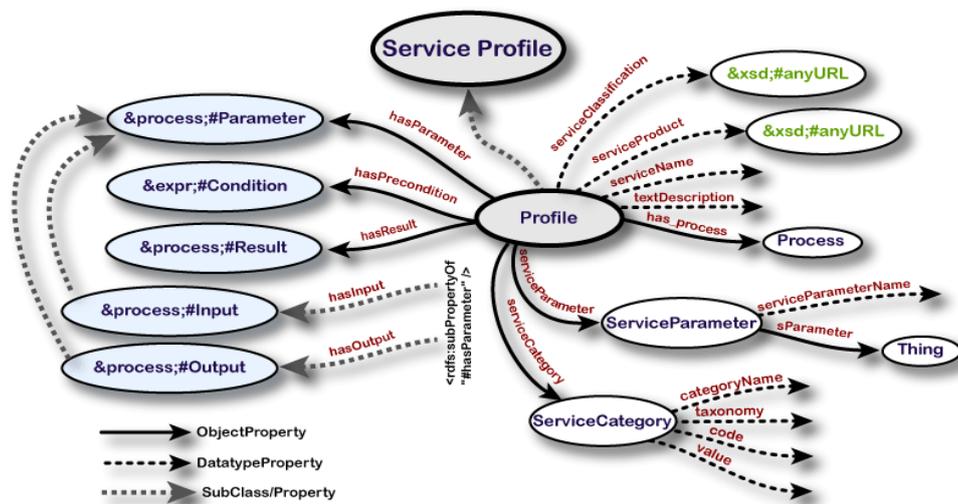


Figura 2.6: Classes e propriedades da linguagem OWL-S para definir perfil do serviço (Martin et al. 2004)

```

6 com a dor abdominal aguda e responde um possível diagnóstico
7 </profile:textDescription>
8 <profile:contactInformation>
9 <actor:phone>551636039773 </actor:phone>
10 <actor:fax>551636039773</actor:fax>
11 <actor:email>dfpires@imagcom.org</actor:email>
12 </profile:contactInformation>

```

As descrições funcionais do serviço informam quais devem ser os dados de entrada, e quais são os dados de saída, além de informar as pré-condições e os efeitos esperados. Como exemplo, as descrições funcionais de um serviço de apoio à decisão diagnóstica na síndrome de diabetes informaria quais os sinais clínicos devem ser informados, e quais são os possíveis diagnósticos como resposta, e como efeito o percentual de confiabilidade da resposta. Na linguagem OWL-S, os dados de entrada e de saída são representados, respectivamente, pelas sub-propriedades *hasInput* e *hasOutput* da propriedade *hasParameter* da classe *Profile*. Essas propriedades possuem como conteúdo objetos das classes *Input* e *Output*, respectivamente. As informações de pré-condição e efeitos são representados na linguagem OWL-S pelas propriedades *hasPrecondition* e *hasResult*, respectivamente, tendo essas propriedades objetos das classes *Condition* e *Result*, também respectivamente.

A Figura 2.6 (Martin et al. 2004) apresenta essas classes e propriedades. O documento a seguir apresenta um exemplo da síndrome de dor abdominal aguda.

```

1 <profile:hasInput
2   rdf:resource="http://143.107.220.180/DDSOnt/DDSPprocesso.owl#pathology"/>
3 <profile:hasInput
4   rdf:resource="http://143.107.220.180/DDSOnt/DDSPprocesso.owl#vomitus"/>
5 <profile:hasInput
6   rdf:resource="http://143.107.220.180/DDSOnt/DDSPprocesso.owl#severities"/>
7 <profile:hasOutput
8   rdf:resource="http://143.107.220.180/DDSOnt/DDSPprocesso.owl#diseaseOrSindrome"/>
9 <profile:hasEffect
10  rdf:resource="http://143.107.220.180/DDSOnt/DDSPprocesso.owl#confiability"/>

```

Finalmente, as características do serviço dizem respeito aos seguintes tipos de informação são:

1. a categoria do serviço, proveniente de um sistema de classificação como o

NAICS;

2. uma taxa indicando a qualidade do serviço, como por exemplo se é confiável ou rápido, o que facilitaria o provedor publicar esses números em um servidor que indica taxa de qualidade de serviços na Web, e o requisitante podendo consultar esse sistema; e
3. uma lista com tamanho indefinido que pode conter qualquer outro tipo de informação relevante, como o tempo máximo de resposta, e o público-alvo, como apenas para médicos iniciantes, ou apenas para médicos da região sudeste.

A categoria do serviço é representada na linguagem OWL-S pela propriedade *serviceCategory* da classe *Profile*. Essa propriedade tem como conteúdo uma instância da classe *ServiceCategory*. Essa classe possui as propriedades *categoryName*, que informa o nome da categoria, por exemplo NAICS, *taxonomy*, que representa a taxonomia da categoria do serviço, por exemplo *medical*, *value*, que informa o valor em uma taxonomia em específico, por exemplo *Medical and Diagnostic Laboratories* e *code*, que informa o código na taxonomia em específico, como 6215. A qualidade do serviço é especificada pela classe *QualityRating*. Como é possível de se ver na Figura 2.6, o conteúdo dessas propriedades refere-se a uma taxonomia não presente na ontologia OWL. O documento a seguir ilustra um exemplo:

```
1 <profile : serviceCategory>
2   <addParam:NAICS rdf:ID="NAICS-category">
3     <profile : value>
4       Medical and Diagnostic Laboratories
5     </profile : value>
6     <profile : code> 6215 </profile : code>
7   </addParam:NAICS>
8 </profile : serviceCategory>
9 <profile : qualityRating>
10  <profile : QualityRating
11    rdf:ID="DDSOnt-goodRating">
12    <profile : ratingName> 90% </profile : ratingName>
13    <profile : rating
14      rdf:resource="http://143.107.220.180/DDSOnt/DDSConceitos.owl#GoodRating"/>
15  </profile : QualityRating>
16 </profile : qualityRating>
```

A lista ilimitada de parâmetros que pode conter outras informações relevantes de um serviço é definida pela propriedade *serviceParameter* da classe *Profile*. O conteúdo dessa propriedade é uma instância da classe *ServiceParameter*, que contém a propriedade *serviceParameterName* indicando o nome do parâmetro não definido na ontologia OWL, e a propriedade *sParameter*, contendo o conteúdo do parâmetro. O documento a seguir ilustra um exemplo:

```
1 <profile : serviceParameter>
2   <addParam:publicoAlvo
3     rdf:ID="DDSOnt-publicoAlvo">
4     <profile : serviceParameterName>
5       Público-alvo da DDSOnt
6     </profile : serviceParameterName>
7     <profile : sParameter
8       rdf:resource="http://143.107.220.180/DDSOnt/DDSPublico.owl#Iniciantes"/>
9   </addParam:publicoAlvo>
10 </profile : serviceParameter>
```

Para se classificar o tipo de serviço fornecido, bem como os produtos tratados pelo serviço, duas propriedades da classe *Profile* podem ser utilizadas para tal. Essas propriedades são, respectivamente, *serviceClassification* e *serviceProduct*. Essas propriedades possuem conteúdo que pertence à taxonomia da ontologia OWL. O documento *DDSOntProfile.owl* contendo o perfil dos cinco serviços Web construídos no

estudo de caso deste projeto pode ser encontrado em na Web ⁶.

2.4.1.2 Modelando Serviços como Processos

A segunda forma de se representar o conhecimento de um serviço Web através da OWL-S é mostrando-o na forma de processos. De acordo com Martin e co-autores em (Martin et al. 2004), pode-se mostrar todas as maneiras que um cliente pode interagir com o mesmo, tentando facilitar assim o uso.

Existem três tipos de processo: atômico, simples e composto. O processo atômico corresponde às ações que um serviço pode desempenhar em uma única interação, não possuem sub-processos e são invocados diretamente. Esse tipo de processo obtém uma mensagem de entrada, processa-a, e retorna uma mensagem de saída. Para cada processo atômico, é necessário especificar a forma de acesso ao serviço associado, chamado de fundamento do serviço, que será apresentado na seção 2.4.1.3. Finalmente, sempre existem dois participantes em um processo como esse, o cliente e o servidor. O documento a seguir apresenta a definição do único processo atômico UMLService, que recebe 2 parâmetros de entrada (*Atom_In* e *SourceSystemName_In*), que são o átomo e o nome de um sistema de terminologia, e retorna os valores *UMLSAui_Out*, *UMLSSui_Out*, *UMLSLui_Out*, *UMLSCui_Out*, *UMLSConcept_Out*, *UMLSSemanticType_Out*, *UMLSDefinition_Out* e *UMLSSourceSystemName_Out*, que são alguns valores associados ao átomo no servidor da UMLS. Outros quatro processos atômicos foram criados neste trabalho e no contexto do estudo de caso apresentado no Capítulo 5: *ClinicalKnowledgeCreatorWebService*, *DiagnosisDecisionSupportWebService*, *ClinicalKnowledgePublishWebService* e *ClinicalKnowledgeDiscoverWebService*.

```
1 <process:AtomicProcess rdf:ID="UMLService">
2   <profile:hasInput rdf:resource="#Atom_In"/>
3   <profile:hasInput rdf:resource="#SourceSystemName_In"/>
4   <process:hasOutput rdf:resource="#UMLSAui_Out"/>
5   <process:hasOutput rdf:resource="#UMLSSui_Out"/>
6   <process:hasOutput rdf:resource="#UMLSLui_Out"/>
7   <process:hasOutput rdf:resource="#UMLSCui_Out"/>
8   <process:hasOutput rdf:resource="#UMLSConcept_Out"/>
9   <process:hasOutput rdf:resource="#UMLSSemanticType_Out"/>
10  <process:hasOutput rdf:resource="#UMLSDefinition_Out"/>
11  <process:hasOutput rdf:resource="#UMLSSourceSystemName_Out"/>
12 </process:AtomicProcess>
```

O processo simples é parecido com o atômico pois também não possui sub-processos. Entretanto, não são invocados diretamente, ou seja, não se especifica a forma de acesso ao processo simples. Sendo assim, é utilizado como elemento de abstração, seja para mostrar uma visão especializada de um processo atômico, seja para mostrar uma representação simplificada de um processo composto. Pode-se dizer que um processo simples é realizado por um processo atômico, que por sua vez realiza um processo simples. Nenhum processo simples foi criado neste projeto.

O processo composto é aquele que pode ser decomposto em outros processos. A decomposição deve ser especificada utilizando-se de controladores que indicam de que forma o cliente deve se comportar durante o envio e o recebimento das mensagens. Os controladores são, segundo Martin e colaboradores em (Martin et al. 2004):

Sequence : os sub-processos devem ser executados em uma ordem específica;

Split : os sub-processos são executados concorrentemente. A operação se completa quando todos os sub-processos tiverem sido iniciados;

⁶<http://143.107.220.180/DDSOnt/Semantic/DDSOntProfile.owl>

Split+Join : os sub-processos são executados concorrentemente. A operação se completa apenas quando todos os sub-processos tiverem sido finalizados;

Choice : permite a escolha de um sub-processo, dentre vários outros, para ser executado;

Any-Order : os sub-processos são executados em qualquer ordem. A operação se completa apenas quando todos os sub-processos tiverem sido finalizados;

It-Then-Else : permite o teste de uma condição lógica, que caso verdadeira executa um sub-processo A, e caso seja falso executa-se o sub-processo B;

Repeat-While : enquanto uma condição lógica for verdadeira um sub-processo é executado;

Repeat-Until : um sub-processo é executado até que uma condição lógica seja falsa. Neste caso, o sub-processo é executado ao menos uma vez.

O documento a seguir apresenta um exemplo de um processo composto chamado DDSOnt_Process criado neste projeto no contexto do estudo de caso apresentado no Capítulo 5. Esse processo é composto dos cinco processos do tipo átomo também criados no estudo de caso, que devem ser executados com o controlador de sequência.

```
1 <process:CompositeProcess rdf:ID="DDSOnt_Process">
2   <rdfs:label>
3     This is the top level process for DDSOnt
4   </rdfs:label>
5   <process:composedOf>
6     <process:All>
7       <process:components rdf:parseType="Collection">
8         <process:AtomicProcess
9           rdf:about="#UMLSWebService"/>
10        <process:AtomicProcess
11          rdf:about="#ClinicalKnowledgeCreatorWebService"/>
12        <process:AtomicProcess
13          rdf:about="#DiagnosiDecisionSupportsWebService"/>
14        <process:AtomicProcess
15          rdf:about="#ClinicalKnowledgePublishWebService"/>
16        <process:AtomicProcess
17          rdf:about="#ClinicalKnowledgeDiscoverWebService"/>
18      </process:components>
19    </process:Sequence>
20  </process:composedOf>
21 </process:CompositeProcess>
```

A linguagem OWL-S também permite especificar o fluxo de dados entre os sub-processos, ou seja, pode-se definir que os dados de entrada do sub-processo B são exatamente os dados de saída do sub-processo A. Ainda, caso um processo tenha dados de saída de vários sub-processos, estas saídas podem ser nomeadas e identificadas de modo a facilitar sua utilização pelo processo geral. O documento DDSOnt-Process.owl contendo a modelagem de todos os cinco serviços Web construídos no estudo de caso deste projeto pode ser encontrado na web ⁷.

2.4.1.3 Fundamento do Serviço

A terceira forma de se representar o conhecimento de um serviço Web através da linguagem OWL-S é modelando o fundamento do serviço. Modelar o fundamento do serviço significa especificar os detalhes de como acessá-lo, como por exemplo apresentar o protocolo de comunicação, o formato de mensagens, a forma de serialização dos objetos, e o meio de transporte. Ao contrário das representações abstratas dos serviços Web em termos de perfil e processo apresentados nas sub-seções 2.4.1.1 e 2.4.1.2, a apresentação do fundamento do processo ocorre em uma representação

⁷<http://143.107.220.180/DDSOnt/Semantic/DDSOntProcess.owl>

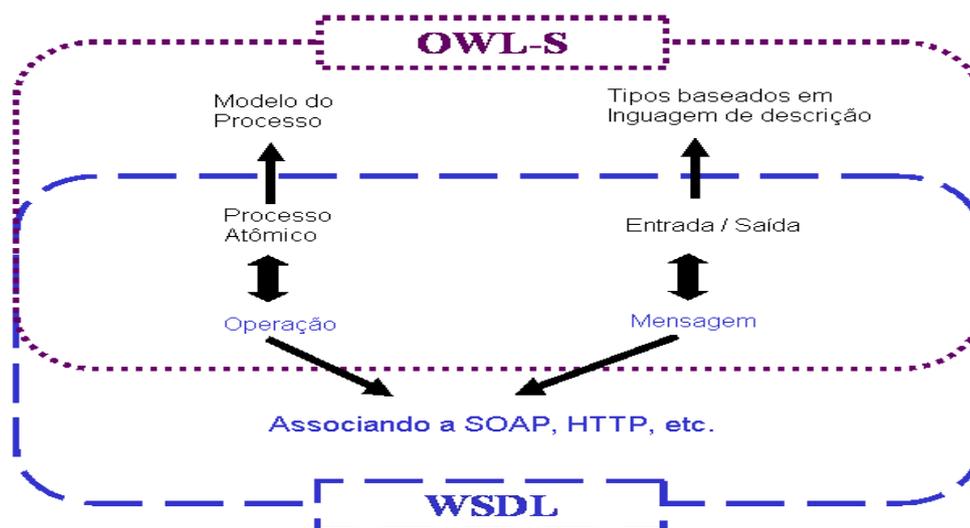


Figura 2.7: Classes e propriedades da linguagem OWL-S para definir perfil do serviço (Martin et al. 2004)

concreta. Assim, o objetivo do OWL-S é mostrar como as informações abstratas de entrada e saída de um processo atômico são representadas concretamente na forma de mensagens, ou seja, como ligar as mensagens aos protocolos de comunicação. Para isso, a linguagem WSDL, apresentada na seção 2.3.4, é utilizada para representar concretamente as mensagens. Segundo Martin e co-autores em (Martin et al. 2004), a vantagem de utilizar OWL-S e WSDL na descrição do fundamento de um serviço é poder se aproveitar da modelagem dos serviços em termos de processo, no caso implementado na OWL-S, e poder reutilizar recursos já presentes na WSDL, como a associação das mensagens ao protocolo SOAP. Ainda, utilizar a linguagem WSDL neste caso é estar de acordo com programas que já foram desenvolvidos para suportá-la.

A Figura 2.7 (Martin et al. 2004) apresenta a interação entre OWL-S e WSDL durante o processo de representação do fundamento de um serviço, mostrando que as duas linguagens neste momento são complementares. As classes da linguagem OWL-S são utilizadas para definição dos tipos abstratos de dados declarados nas mensagens WSDL, já que os tipos podem ser construídos com uma linguagem de descrição nativa da linguagem OWL, ao contrário do WSDL que define tipos de dados com XML Schema. WSDL é utilizada para construir a ligação do formato das mensagens aos protocolos de comunicação, visto que OWL-S não possui esse recurso em nível concreto.

Pelas correspondências da Figura 2.7, é possível perceber que:

- um processo atômico da OWL-S é correspondente a uma operação da WSDL, já apresentada na seção 2.3.5;
- o conjunto de entradas e o conjunto de saídas de um processo atômico da OWL-S corresponde cada um a uma mensagem WSDL, também já mencionada na seção 2.3.5;
- os tipos de dados da entrada e saída de um processo atômico da OWL-S corresponde a um tipo abstrato da WSDL; e

- existe um mecanismo pelo qual os elementos WSDL são referenciados junto às propriedades da classe *WsdAtomicProcessGrounding* da ontologia OWL-S.

O documento *DDSOntFundamento.owl* contendo o fundamento de todos os cinco serviços Web construídos no estudo de caso deste projeto pode ser encontrado na web ⁸.

2.5 Considerações Finais

As aplicações clínicas na Web que precisam trocar dados médicos em ambiente colaborativo e compartilhado possuem como opção computacional vários protocolos de comunicação, como DCOM, RMI e DSOM. Estes protocolos, caracterizados como sendo proprietários, criam um cenário heterogêneo, fortemente acoplado, e que pode exigir alguma linguagem de programação ou arquitetura computacional em particular para seu ideal funcionamento. Sendo assim, faz-se necessário que as aplicações clínicas que precisam trocar experiências clínicas para apoiar a decisão diagnóstica façam uso do protocolo de comunicação SOAP, caracterizado como sendo um padrão aberto, que funciona sobre o protocolo HTTP, e utiliza XML para formatar suas mensagens. É desta forma que este trabalho justifica a utilização de SOAP na implementação dos serviços Web relacionados à ontologia DDSOnt.

Os softwares clínicos que se utilizam de procedimentos espalhados na Web precisam, além do recurso de computação distribuída para executá-lo, de mecanismos de descrição, de publicação e de descoberta destes procedimentos em geral ou de serviços relacionados ao apoio à decisão diagnóstica em particular. Seguindo o mesmo raciocínio da escolha do protocolo de comunicação, as especificação WSDL e UDDI são utilizadas neste trabalho para descrever, e disponibilizar e recuperar, respectivamente, estes serviços clínicos relacionados ao apoio à decisão diagnóstica em ambiente colaborativo. A linguagem OWL-S foi utilizada neste trabalho para se criar documentos XML que possam adicionar informações semânticas de descrição dos serviços clínicos de apoio à decisão diagnóstica da DDSOnt, durante as etapas de publicação e recuperação dos mesmos. Até o presente momento, a OWL-S ainda não é uma recomendação W3C, e portanto encontra-se em fase de análise e discussão pela comunidade Web.

O próximo capítulo aborda problemas e possíveis soluções da incompatibilidade terminológica durante a troca de informações para apoiar a decisão diagnóstica em ambiente colaborativo na Web. Ainda, são discutidos alguns sistemas de terminologia em saúde, em especial a UMLS MetaThesaurus, e o porquê da utilização deste sistema na solução proposta por este projeto.

⁸<http://143.107.220.180/DDSOnt/Semantic/DDSOntGrounding.owl>

Incompatibilidade Terminológica em Sistemas de Informação em Saúde Distribuídos

3.1 *Considerações Iniciais*

Este Capítulo tem como objetivo discutir os problemas relacionados a incompatibilidade terminológica durante a troca de informações clínicas entre aplicações médicas, especificamente para apoiar a decisão diagnóstica colaborativa. São apresentados alguns sistemas de terminologia em saúde, como CID, MeSH, SNOMED e UMLS Metathesaurus. Na sequência, este Capítulo discute como o sistema UMLS Metathesaurus consegue na maioria das vezes, promover a compatibilidade terminológica e por isto é utilizado na proposta deste projeto. O capítulo é finalizado apresentando de que forma o estudo destes sistemas contribuiu para alcançar os objetivos deste trabalho.

3.2 *Cenário Médico com Problemas de Incompatibilidade Terminológica*

É comum aplicações clínicas fazerem uso de diferentes sistemas de terminologia médica em razão principalmente da não existência de um sistema único que seja reconhecido, aceito e adotado pela comunidade médica para todas as aplicações e finalidades. Consequentemente, associar dois termos de sistemas terminológicos diferentes é fundamental durante a troca de informações entre aplicações clínicas.

A Figura 3.1 ilustra um exemplo de troca de dados clínicos entre aplicações médicas localizadas remotamente. Nesta Figura, os softwares estão utilizando sistemas

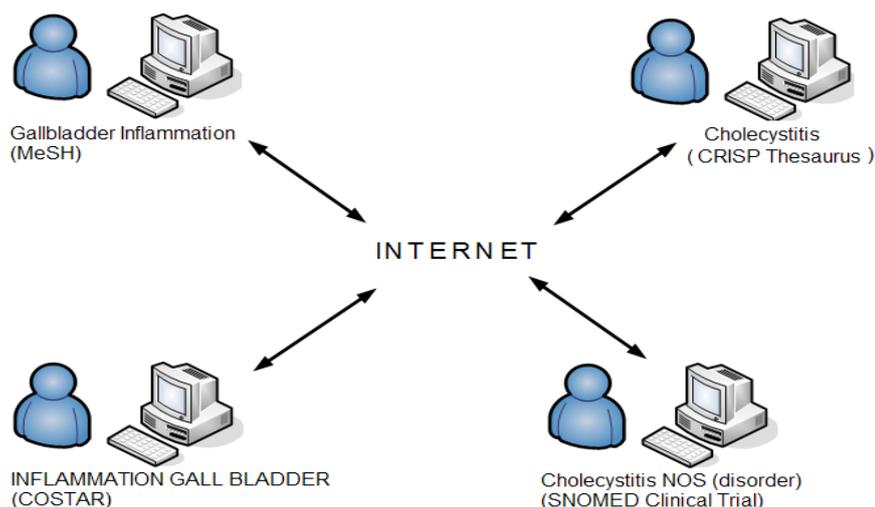


Figura 3.1: Mesmo conceito médico sendo representado por diferentes termos

de terminologia diferentes, que são neste caso: MeSH, CRISP, COSTAR e SNOMED CT. Todos eles representam o conceito *Cholecystitis* com termos e átomos diferentes: o software que utiliza o sistema MeSH utiliza o termo *Gallbladder Inflammation* para representar a *Cholecystitis*; já o aplicativo que faz uso do sistema COSTAR utiliza o termo INFLAMMATION GALL BLADDER para representar a *Cholecystitis*; enquanto que o programa que utiliza SNOMED CT utiliza o átomo *Cholecystitis NOS (disorder)* para exibir a *Cholecystitis*; e finalmente o software que faz uso do CRISP Thesaurus utiliza o mesmo termo *Cholecystitis* para representar o conceito *Cholecystitis*. Este cenário portanto ilustra um problema terminológico.

Outro cenário de incompatibilidade terminológica é quando um mesmo termo clínico de fontes terminologicamente heterogêneas estão relacionados a diferentes conceitos. A Figura 3.2 ilustra um exemplo de troca de dados clínicos entre aplicações médicas localizadas remotamente. Na Figura, os softwares estão utilizando sistemas de terminologia diferentes: MeSH, MTH, e COSTAR. No sistema MeSH, 'COLD' significa temperatura fria, no sistema MTH, 'COLD' significa resfriado, enquanto que no sistema COSTAR, 'COLD' representa o acrônimo da palavra *Chronic Obstructive Lung Disease*, que significa obstrução crônica nos pulmões.

Existe também a possibilidade de, dentro de um mesmo vocabulário, um mesmo termo expressar conceitos diferentes. Para ilustrar este caso, o termo *Interferon-Alpha* no sistema MESH o cataloga como um fármaco, ou seja, uma droga. Ainda, MESH o classifica como uma proteína natural, já que *Interferon-Alpha* é uma proteína natural do corpo humano.

Neste contexto, faz-se necessário estudar e projetar um mecanismo que permita que aplicações médicas troquem dados clínicos objetivando auxílio diagnóstico compartilhado sem que ocorram os problemas apresentados nos cenários anteriores. Com este propósito, este trabalho apresenta os resultados da investigação que busca características presentes em alguns dos sistemas de terminologia na literatura para solucionar o problema de incompatibilidade terminológica. Estes estudos são apresentados a seguir.

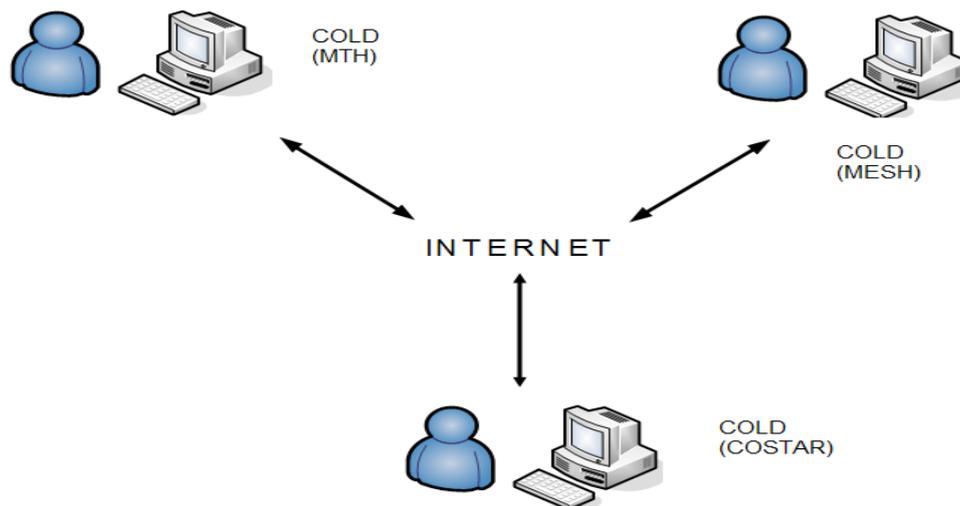


Figura 3.2: Mesmo termo médico com conceitos diferentes

3.3 *Sistemas de Terminologia em Saúde*

Sistemas de terminologia e de codificação em saúde são sistemas controlados que consistem em padronizar um conjunto de termos ou descritores arranjados em uma determinada estrutura. E a partir do conhecimento desta estrutura, pode-se criar, como exemplo, uma aplicação computacional que realize inserções e buscas em vários níveis de especificidade. Exemplos tradicionais de sistemas de codificação clínicos mais encontrados na literatura são CID, SNOMED, MeSH, e o metatesauro da UMLS. Eles serão descritos a seguir.

3.3.1 *CID*

O CID é um sistema internacional de classificação de doenças e de problemas relacionados a saúde mantido pela comunidade médica. O objetivo do CID é fazer estatísticas sobre mortalidade e também morbidade, não sendo portanto um sistema voltado para utilização da prática clínica diretamente, evidenciando histologia da doença, localização, etc. A sua primeira edição foi publicada em 1893, tendo novas revisões de 10 em 10 anos, aproximadamente. A nona revisão (CID-9) foi publicada em 1977 pela WHO ¹, e a décima e última revisão (CID-10) foi publicada em 1992 também pela mesma organização (Shortliffe et al. 2000). Os objetivos gerais destas revisões é tornar o sistema de classificação mais adequado aos estudos epidemiológicos para a avaliação de assistência à saúde, às necessidades dos profissionais da área, e também durante a prestação de contas de hospitais e outras instituições de saúde.

3.3.1.1 *CID-9*

O CID-9 consiste de um sistema de classificação de três dígitos, tendo ainda um quarto dígito que fornece detalhes adicionais à doença. Usualmente, os décimos de

¹<http://www.who.int/en/>

.0 a .7 são utilizados para representar as formas mais específicas de um termo, o décimo .8 é utilizado para representar *other*, e o .9 representa *unspecified*. Os termos são organizados numa hierarquia que é baseada nos dígitos que o termo é codificado. CID-9 também inclui termos que são familiares no diagnóstico da especialidade médica, situação da doença, incapacidades, e causas da doença. O documento (Shortliffe et al. 2000) a seguir apresenta a doença *Bacterial pneumonias* codificada em CID-9 de forma simplificada, procurando exemplificar os conceitos relacionados mais específicos.

- 1 481 Pneumococcal Pneumonia
- 2
- 3 482 Other Bacterial Pneumonia
- 4 482.0 Pneumonia due to Klebsiella Pneumoniae
- 5 482.1 Pneumonia due to Pseudomonas
- 6 482.2 Pneumonia due to Haemophilus Influenzae
- 7 482.3 Pneumonia due to Streptococcus
- 8 482.4 Pneumonia due to Staphylococcus
- 9 482.8 Pneumonia due to Other Specified Bacteria
- 10 482.9 Bacterial pneumonia unspecified
- 11
- 12 484 Pneumonia in Infectious Disease Classified Elsewhere
- 13 484.3 Pneumonia in Whooping Cough
- 14 484.4 Pneumonia in Tularemia
- 15 484.5 Pneumonia in Anthrax

Existem diagnósticos que contêm informações tanto sobre a doença geral subjacente como também sobre uma manifestação num órgão ou localização particular. Assim, são necessários dois códigos para enunciá-los. O código primário ou símbolo da cruz (†ou +) é utilizado para marcar a causa básica da doença. O código adicional ou símbolo asterisco (*) é utilizado para marcar a manifestação da doença. Como exemplo, pode-se citar a doença *Tuberculous meningitis*, doença geral e código 013.0†, e *Meningitis in other bacterial diseases classified elsewhere*, manifestação num órgão, e código 320.7*. Esta convenção foi adotada em razão da codificação somente pela doença básica ser freqüentemente pouco satisfatória para a compilação de estatísticas de especialidades médicas (CID-10 1993).

Durante a utilização do CID por profissionais da área da saúde, percebeu-se que este sistema era inadequado pois não atendia um nível de detalhamento desejado para relatórios estatísticos em países como os EUA ². Assim, um conjunto de modificações clínicas foram adicionados ao CID-9, originando o sistema CID-9 CM (*Clinical Modifications*). Este sistema possui um nível extra de detalhe em alguns termos, sendo adicionado o quinto dígito em seus códigos. O documento de Shortliffe e co-autores em (Shortliffe et al. 2000) a seguir ilustra um exemplo de uso de 5 dígitos para codificar doenças relacionadas a *Salmonella Infections*, ilustrando os conceitos apresentados.

- 1 003 Other Salmonella Infections
- 2 003.0 Salmonella Gastroenteritis
- 3 003.1 Salmonella Septicemia
- 4 003.2 Localized Salmonella Infections
- 5 003.20 Localized Salmonella Infection , Unspecified
- 6 003.21 Salmonella Meningitis
- 7 003.22 Salmonella Pneumonia
- 8 003.23 Salmonella Arthritis
- 9 003.24 Salmonella Osteomyelitis
- 10 003.29 Other Localized Salmonella Infection
- 11 003.8 Other specified salmonella infections
- 12 003.9 Salmonella infection , unspecified

²Apesar dos EUA ainda usarem o CID-9 para morbidade, o CID-10 é usado para a mortalidade. Fonte: <http://www.medical-billing-coding.org/Content255.htm>

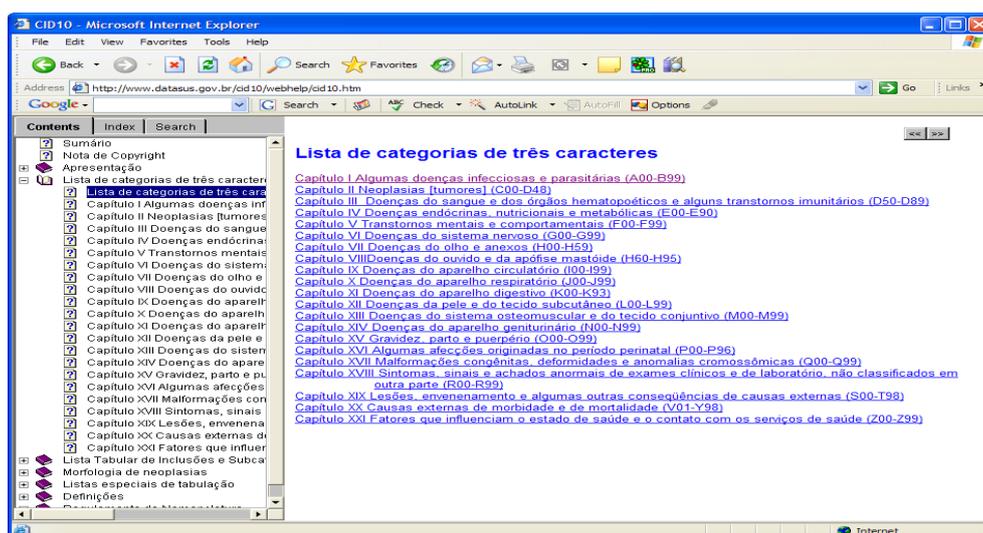


Figura 3.3: CID-10 disponível para consulta na Internet.

3.3.1.2 CID-10

A mais nova versão do CID, CID-10, manteve a tradicional estrutura da versão anterior, porém alterou o esquema de codificação. O novo esquema acrescenta caracteres alfanuméricos. Isso fez com que o sistema pudesse representar um número bem maior de códigos, deixando espaços para que em futuras revisões não haja rompimento da ordenação, como ocorria nas revisões anteriores. Como exemplo de codificação na versão 10, as doenças (*Tuberculous meningitis*) e (*Meningitis in other bacterial diseases classified elsewhere*) possuem os novos códigos A17.0†e G01*, respectivamente.

Visando aproveitar os espaços disponíveis na nova versão, alguns transtornos do mecanismo imune foram adicionados junto às doenças do sangue e dos órgãos hematopoéticos. Ainda, novos capítulos foram criados para as doenças do olho e anexos, e também para as doenças do ouvido e da apófise mastóide. Ainda, as antigas classificações suplementares de causas externas e de fatores que exercem influências sobre o estado de saúde e de oportunidades de contato com serviços de saúde passaram a fazer parte do corpo da classificação. Finalmente, o sistema duplo de classificação para alguns diagnósticos – sistema de cruz e asterisco – introduzido no CID-9 foi mantido e expandido, estando agora o eixo de asterisco contido em categorias homogêneas ao nível de três caracteres (CID-10 1993).

O CID-10 citeOMS1993 possui uma versão em português do CID³ e é utilizado desde 1998 pelo Ministério da Saúde brasileiro na codificação das doenças, nos itens de morbidade hospitalar e ambulatorial, e em mortalidade também, em todo território nacional. A Figura 3.3 ilustra a página na Internet disponibilizada pelo DATASUS para que qualquer pessoa possa navegar e consultar o CID-10. O sistema também é disponível para download.

³<http://www.datasus.gov.br/cid10/webhelp/cid10.htm>

3.3.2 SNOMED

Em 1971 o *College of American Pathologists* criou um sistema multi-eixos para descrever patologias. Cada um dos eixos representava um determinado domínio médico, como a anatomia, a morfologia, a etiologia, funções, etc. A este sistema foi dado o nome de SNOF. A evolução do sistema de descrição patológica para torná-lo mais compreensivo originou um novo nome: SNOMED. SNOMED foi primeiramente publicado em 1975 por Roger Cote e David Rothwell, e posteriormente sua revisão em 1979 culminou no sistema SNOMED II. Após isso, com a inserção de informações patológicas relacionadas a medicina veterinária, surgiu uma versão mais completa intitulada *SNOMED International*, ou *Systematized Nomenclature of Human and Veterinary Medicine*, publicada por Cote e co-autores em (Cote et al. 1993).

O sistema *SNOMED International* possui um conjunto de 150.000 registros, organizados em 12 eixos. Os conceitos atuais incluem dados do tipo sinais e sintomas de doenças, substâncias químicas, drogas, enzimas, proteínas, organismos vivos, agentes físicos, relacionamento espacial, ocupações, contextos sociais, doença/diagnóstico e procedimentos, além de informações anatômicas (topográficas), morfológicas (estrutura patológica), e funções normais e irregulares. O artigo ou livro de Shortliffe e colaboradores em (Shortliffe et al. 2000) ilustra uma parte do *SNOMED International* para a pneumonia. Veja um excerto deste abaixo. As colunas da esquerda indicam o eixo das doenças relacionadas a *Bacterial Infectious*. O símbolo NOS significa (*Not Otherwise Specified*). As colunas da direita representam códigos da SNOMED que, quando agrupados, são equivalentes aos termos à sua esquerda. Como exemplo, considere a linha 8. O termo *Pneumococcal pneumonia* DE-13510 pode também ser representado com os termos *Lung, NOS(T-28000)*, *Inflammation, NOS(M-40000)*, e *Streptococcus pneumoniae(L-25116)*. O símbolo T representa topografia ou anatomia, o símbolo M significa morfologia ou estrutura patológica, e o símbolo L representa os organismos vivos.

| | | | |
|---|----------|-----------------------------------|------------------------------|
| 1 | DE-10000 | Bacterial infectious disease, NOS | (L-10000) |
| 2 | DE-10100 | Bacterial pneumonia, NOS | (T-28000)(M-40000)(L-10000) |
| 3 | DE-11205 | Pneumonia in anthrax | (T-28000)(M-40000) |
| 4 | DE-13212 | Pneumonia in pertussis | (T-28000)(M-40000) |
| 5 | DE-13430 | Pneumonic plague, NOS | (T-28000)(L-1E401)(DE-01750) |
| 6 | DE-13431 | Primary Pneumonic plague | (T-28000)(L-1E401)(DE-01750) |
| 7 | DE-13432 | Secondary Pneumonic plague | (T-28000)(L-1E401)(DE-01750) |
| 8 | DE-13510 | Pneumococcal pneumonia | (T-28000)(M-40000)(L-25116) |
| 9 | DE-13934 | Salmonella pneumonia | (T-28000)(L-17100) |

A possibilidade de representar um termo clínico de diferentes maneiras no SNOMED, ou seja, utilizando-se de conjuntos formados por diferentes códigos, torna-o muito expressivo. Outro exemplo é a representação de *acute appendicitis*, que pode ser codificado como um termo único de doença, como também a combinação do modificador *acute* e o termo da doença *appendicitis*, ou ainda a combinação do modificador *acute*, o termo morfológico *inflammation* e o termo topográfico *vermiform appendix* (Shortliffe et al. 2000). Cada uma destas representações são corretas, apesar de não haver no *SNOMED International* uma forma de verificar se estas combinações possuem o mesmo significado, o que torna um problema para os desenvolvedores de aplicações.

3.3.2.1 SNOMED RT

SNOMED RT apresenta-se como um conjunto de conceitos e relacionamentos que fornece um ponto de referência comum para comparação e agregação de dados envolvidos nos processos em saúde, e gravados por diferentes usuários, instituições e

sistemas (Spackman et al. 1997). Neste contexto, SNOMED RT procura solucionar os problemas apresentados no SNOMED *International*, relacionando os termos únicos ou uma combinação de termos a conceitos. Assim, para saber se o termo único *acute appendicitis*, o termo combinado *acute e appendicitis*, e ainda os termos combinados *acute, inflammation e vermiform appendix* possuem o mesmo significado, é necessário consultar se estes estão relacionados ao mesmo conceito. De modo a representar estes relacionamentos, SNOMED RT possui uma linguagem lógica de descrição chamada KRSS.

Ainda, o SNOMED RT possui uma nova hierarquia explícita de representação dos termos, como mostra o artigo de Spackman e colaboradores em (Spackman et al. 1997). Veja o excerto a seguir. Como exemplo, a linha 3 do documento significa que a doença representada pelo código DD-22111 é filha ou possui o mesmo significado da doença representada por DD-22100.

| | Parent Term | Child Term | Relationship |
|---|-------------|------------|--------------|
| 1 | DD-00000 | DD-20000 | Is-A |
| 2 | DD-20000 | DD-22100 | Is-A |
| 3 | DD-22100 | DD-22111 | Is-A |
| 4 | D2-00000 | D2-80000 | Is-A |
| 5 | D2-80000 | D2-80300 | Is-A |
| 6 | D2-80300 | DD-22111 | Is-A |
| 7 | | | |

3.3.2.2 SNOMED CT

SNOMED CT é uma combinação do conteúdo e da estrutura do SNOMED RT com os termos clínicos do sistema *Read Codes* (Read and Benson 1986), também conhecido como *Clinical Terms Version 3* (CTV3). Stearns e co-autores (Stearns et al. 2001) apresentam os desafios da união destes dois sistemas de terminologias independentes, e discute o processo desenvolvido para identificar as similaridades e diferenças entre eles. Esta combinação caracteriza o SNOMED CT como contendo 366.170 conceitos com significados únicos e com definições lógicas formais, organizados em hierarquias (SNOMED 2007):

Clinical Finding : contém conceitos das descobertas e doenças, como inchaço da mão (descoberta clínica) e pneumonia (doença);

Procedure/intervention : conceitos que representam as atividades intencionais para desempenhar a promoção da saúde, como a biopsia no pulmão, uma endoscopia diagnóstica, ou uma manipulação fetal;

Observable entity :conceitos que representam uma questão ou procedimento, que quando combinados com um resultado, constituem uma descoberta. Exemplos são tamanho de um tumor e identidade sexual;

Body structure : conceitos relacionados a estrutura anatômica do corpo e anormalidades morfológicas. Tireóide lingual (estrutura do corpo) e neoplasma (anormalidade morfológica) são dois exemplos;

Organism : Incluem animais, fungos, bactérias e plantas. Exemplos: vírus da hepatite C, *streptococcus pyogenes* e *Felis silvestris* (gato);

Substance : cobre as substâncias biológicas e químicas. Pó, estrogênio, anti-corpus da hemoglobina e metano são exemplos;

Physical object : objetos utilizados para tratar de ferimentos médicos, como próteses, bolsa para colotomia, e drenos;

Physical force : inclui movimentação, fricção, eletricidade, som, radiação, força térmica e pressão do ar;

Events : representam ocorrências que resultam em ferimento. Acidente com veículo motorizado e raios;

Environments/geographical locations : Representam países, estados e regiões, como hospital do cancer, Brasil e centro de queimados;

Social context : inclui estados econômicos e familiares, religião, estilo de vida e ocupação;

Context-dependent categories : informação adicional necessária em um determinado contexto, como histórico familiar de derrame;

Staging and scales : contém informações de nomes de escalas de avaliação e estágios de doença, como a escala Glasgow para avaliar estado de coma, e a Dukes para avaliar estágios de um tumor.

O SNOMED CT contém ainda mais de 993.420 sinônimos na língua inglesa para flexibilizar a expressão dos conceitos. Um exemplo é o conceito *Pain in throat*, que possui os seguintes sinônimos: *Sore throat*, *Throat pain*, *Pain in pharynx*, *Throat discomfort*, *Pharyngeal pain*, e *Throat soreness*⁴. Esta expressividade pode muitas vezes tornar-se um problema para o sistema de terminologia caso não existam relacionamentos que indiquem que todos estes sinônimos estão associados ao mesmo conceito, o que não ocorre com o SNOMED CT. Este sistema possui aproximadamente 1.46 milhões de relacionamentos semânticos de modo a permitir uma recuperação dos dados de maneira confiável e consistente.

3.3.3 MeSH

O MeSH (Lipscomb 2000), mantido pela NLM é um vocabulário médico pelo qual a literatura médica é indexada. MeSH organiza os termos em uma estrutura que contraria as estruturas utilizadas por outros sistemas de terminologia, como CID e SNOMED CT, apresentadas nas seções 3.3.1 e 3.3.2.2, respectivamente. No sistema MeSH os termos são organizados em hierarquias e podem aparecer em vários lugares nesta. Sendo assim, este sistema geralmente não é utilizado como um esquema de codificação direta de informações de pacientes e é mais utilizado para a recuperação de informações como em pesquisas bibliográficas.

O documento a seguir apresenta o resultado de uma consulta pelo termo *Pneumonia* realizada no sistema on-line MeSH disponível na Web⁵. Existe uma versão do MeSH disponível também para download. Pode-se perceber nas linhas de 1 a 3 que *Pneumonia* foi classificado como inflamação no pulmão, e também nas linhas 10 e 11 foi classificado como 'infecção no pulmão'. Sendo assim, o mesmo termo pode possuir mais de um significado.

- 1 Respiratory Tract Diseases [C08]
- 2 Lung Diseases [C08.381]
- 3 Pneumonia [C08.381.677]
- 4 Bronchopneumonia [C08.381.677.127]
- 5 Pleuropneumonia [C08.381.677.473]

⁴What is SNOMED CT. SNOMED Clinical Terms Core Content. Disponível em http://www.snomed.org/snomedct/what_is.html

⁵<http://www.nlm.nih.gov/mesh/2006/MBrowser.html>

| | |
|----|---|
| 6 | Pneumonia, Aspiration [C08.381.677.529] + |
| 7 | Pneumonia, Bacterial [C08.381.677.540] + |
| 8 | Pneumonia, Pneumocystis [C08.381.677.675] |
| 9 | Pneumonia, Viral [C08.381.677.807] |
| 10 | Respiratory Tract Infections [C08.730] |
| 11 | Pneumonia [C08.730.610] |
| 12 | Bronchopneumonia [C08.730.610.127] |
| 13 | Pleuropneumonia [C08.730.610.473] |
| 14 | Pneumonia, Aspiration [C08.730.610.529] + |
| 15 | Pneumonia, Bacterial [C08.730.610.540] + |
| 16 | Pneumonia, Pneumocystis [C08.730.610.675] |
| 17 | Pneumonia, Viral [C08.730.610.763] |

A estrutura do sistema MeSH é organizada em 16 categorias. Cada uma delas é dividida em sub-categorias. Em cada sub-categoria, descritores são listados hierarquicamente, do mais geral para o mais específico, de cima para baixo na hierarquia, como mostra o documento anterior. Em razão de sua ramificação, esta hierarquia é também chamada de árvore. Os descritores são seguidos por um número que indica uma posição na árvore, não possuindo valor intrínseco algum.

Os sistemas de terminologia MeSH, SNOMED e CID apresentados anteriormente caracterizam-se por organizarem e representarem seus termos clínicos em estruturas particulares, o que conseqüentemente resulta em soluções individuais de armazenamento, representação e recuperação de termos médicos. Estes sistemas apresentam-se como muito úteis em casos de serem utilizados de maneira isolada e individual. Porém, no cenário explorado por este projeto de troca de informações para apoiar a decisão diagnóstica em ambiente computacional colaborativo, faz-se necessário a utilização de um sistema que possa unificar estes sistemas terminológicos na área da saúde.

3.4 UMLS Metathesaurus – Um sistema que promove a compatibilidade terminológica

3.4.1 Introdução ao UMLS

O objetivo do sistema UMLS, desenvolvido e organizado pela NLM, *National Library of Medicine*, é apoiar o desenvolvimento de sistemas computacionais que realizam atividades de recuperação e integração de informações eletrônicas biomédicas. Com este fim, a NLM produz e distribui bases de dados da UMLS, chamadas de UMLSKS (UMLS *Knowledge Sources*) e softwares associados de apoio. Com os softwares de apoio e a UMLSKS, desenvolvedores podem criar ou mesmo melhorar programas que criam, armazenam e recuperam informações relacionadas a biomedicina.

A NLM disponibiliza três sistemas para promover a compatibilidade terminológica utilizando o UMLSKS: o UMLS Metathesaurus, descrito nesta seção, a UMLS Semantic Network, descrita no próximo capítulo na seção 4.3.1, e a SPECIALIST Lexicon, um conjunto de Strings formado por variações léxicas em vários idiomas dos termos clínicos que compõem os sistemas anteriores. A SPECIALIST Lexicon não será discutida neste trabalho pois está fora do escopo do mesmo.

Estas fontes de conhecimento são distribuídas com vários softwares de apoio, como a UMLSKS API, e sistemas de consulta on-line pelo navegador da Internet.

3.4.2 Introdução ao UMLS Metathesaurus

O UMLS Metathesaurus é uma base de dados de vocabulários que contém informações sobre conceitos biomédicos e relacionados a saúde, seus vários nomes, e o relacionamento entre eles. Projetado para ser utilizado por desenvolvedores de software, este tesouro é constituído de mais de 100 fontes de vocabulários, dentre eles CID, SNOMED e MeSH apresentados na seção 3.3. Dentre eles, alguns possuem versões em vários idiomas, e são utilizados em vários contextos, como no cuidados de pacientes, prestação de contas, estatísticas médicas, e catalogação da literatura médica.

Essencialmente, o Metathesaurus organiza estes vários vocabulários por conceitos ou significados. Ainda, possui ligações de nomes alternativos do mesmo conceito, e também identifica relacionamentos entre diferentes conceitos. Como o tesouro está associado ao UMLS *Semantic Network*, todo conceito da UMLS Metathesaurus está associado a pelo menos um tipo semântico da UMLS *Semantic Network*. Isso fornece uma categorização consistente de todos os conceitos do tesouro.

Como o Metathesaurus é constituído de várias fontes de vocabulário, seu escopo é determinado por estes. Essencialmente, todos os conceitos são provenientes de uma ou mais fontes. Geralmente, se um conceito não aparece em nenhuma fonte, ele também não aparecerá no tesouro. Em função disto, quando uma instituição de saúde possui seu próprio sistema de terminologia, e este não for contemplado pela UMLS, a compatibilidade terminológica poderá não existir.

O Metathesaurus reflete e preserva o significado, o nome dos conceitos, e os relacionamentos vindos das fontes de vocabulários da seguinte maneira ([Thesaurus 2007](#)):

- Quando duas diferentes fontes utilizam o mesmo nome para diferentes conceitos, o Metathesaurus representa ambos os significados e indica qual significado está presente em cada vocabulário. Um exemplo que ilustra esta situação é apresentada na seção a seguir;
- Quando o mesmo conceito aparece em diferentes contextos hierárquicos em diferentes fontes de vocabulário, o Metathesaurus procura incluir todas as hierarquias;
- Quando relacionamentos conflitantes entre dois conceitos aparecem em diferentes fontes, ambas as visões são incluídas no tesouro;
- Apesar de alguns nomes de conceitos ou relacionamentos específicos de algumas fontes possam ser exclusivos e desprovidos de validade, eles ainda são incluídos no MetaThesaurus.

3.4.3 Conceitos na UMLS Metathesaurus

Como já apresentado, o Metathesaurus conecta diferentes nomes do mesmo Conceito de diferentes fontes de vocabulários. O documento a seguir apresenta os diferentes nomes do Conceito *Intestinal Obstruction* obtidos a partir de uma consulta na ferramenta UMLSKS disponível na Internet ⁶. Para cada um destes nomes são armazenados seus identificadores e características chaves, como a fonte do vocabulário e idioma. Como mais de uma fonte pode ter o mesmo nome de Conceito, este

⁶http://umlsks.nlm.nih.gov/kss/servlet/Turbine/template/admin%2Cuser%2CKSS_login.vm

pode ter mais de um identificador, chamado de Átomo, e explicado mais adiante na seção 3.4.5.

- 1 Intestinal Obstruction
- 2 Bowel obstruction
- 3 Intestinal obstruction NOS (disorder)
- 4 Intestinal occlusion
- 5 IO – Intestinal obstruction
- 6 Obstruction of intestine NOS
- 7 Obstruction of intestine NOS (disorder)
- 8 Obstruction of intestine or colon
- 9 occlusion; bowel
- 10 Occlusion of intestine NOS
- 11 Occlusion of intestine NOS (disorder)
- 12 Occlusion of intestine or colon
- 13 Unspecified intestinal obstruction

Um ‘Conceito’ é um significado. Um significado pode ter muitos diferentes nomes, como visto anteriormente. O objetivo do tesauro é entender o significado de cada nome em cada fonte de vocabulário e ligar estes nomes que significam o mesmo conceito (os sinônimos). Portanto, o documento anterior também apresenta os sinônimos do Conceito *Intestinal Obstruction*. Cada Conceito ou significado possui um identificador chamado de CUI (*Concept Unique Identifier*).

3.4.4 Strings na UMLS Metathesaurus

Cada nome de Conceito ou sinônimo de um Conceito ou *String* de um Conceito em cada linguagem no tesauro possui um SUI (*String Unique Identifier*). Portanto, cada linha do documento da seção 3.4.3 possui um SUI. Cada variação no conjunto de caracteres, seja caracteres maiúsculos ou minúsculos, ou ainda pontuação, são *Strings* separadas, com um SUI separado. A mesma *String* em diferentes linguagens, por exemplo, inglês e espanhol, possuem um SUI para cada linguagem. Se a mesma *String*, por exemplo *Cold*, possui mais de um significado, haverá um SUI ligado a mais de um CUI.

3.4.5 Átomos na UMLS Metathesaurus

Os Átomos do Metathesaurus são os nomes dos Conceitos ou as *Strings* de cada uma das fontes de vocabulário. Cada ocorrência da *String* em cada fonte de vocabulário é associado um AUI (*Atom Unique Identifier*). Se exatamente a mesma *String* aparece duas vezes no mesmo vocabulário, um único AUI é associado para cada ocorrência. Quando a mesma *String* aparece em várias fontes de vocabulário, são associados AUI para cada aparição, ou seja, uma mesma *String* pode ter vários AUI associados. Todos estes AUIs serão associados a um único SUI, desde que representem ocorrências da mesma *String*. Diferentemente dos SUIs, um AUI está sempre associado a um único CUI porque cada ocorrência da *String* em uma fonte somente pode ter um significado.

3.4.6 Termo Léxico na UMLS Metathesaurus

Cada *String* é associada a um único Termo Léxico ou LUI (*Lexical Unique Identifier*), que representa todas as suas variações léxicas. Da mesma forma que um SUI pode estar associado a vários CUIs, um LUI pode ser associado a mais de um Conceito.

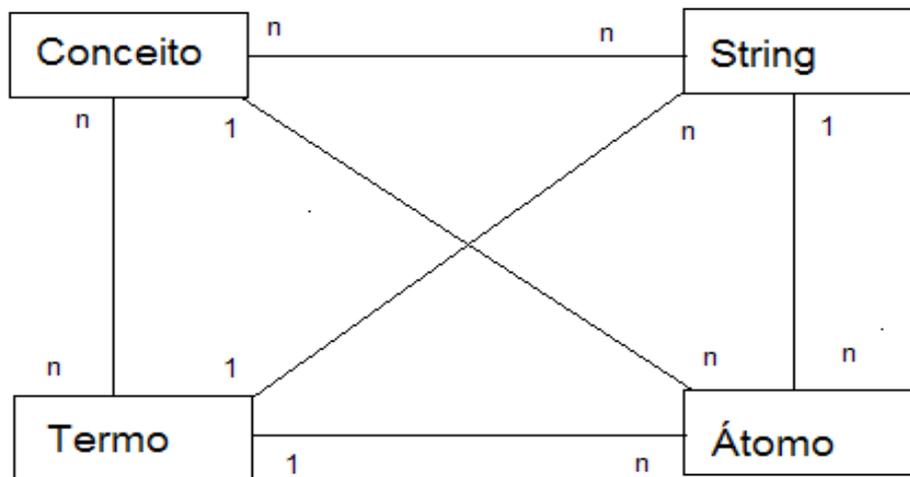


Figura 3.4: Relacionamento entre as entidades da UMLS Metathesaurus.

Isso ocorre quando as *Strings* que ele representa também estão associados a mais de um CUI. Entretanto, cada SUI e cada AUI somente podem estar associados a apenas um único LUI.

Com o objetivo de representar o que foi informado até agora a respeito da relação entre Conceitos, *Strings*, Átomos e Termos, a Figura 3.4 ilustra o relacionamento entre estas entidades. Pela figura, um Conceito pode ter muitas *Strings*, e muitos Átomos e Termos. Uma *String* pode estar associada a mais de um Conceito, a vários átomos e a um único Termo. Um Átomo está associado a um único Conceito, *String* e Termo. E finalmente, um mesmo Termo pode estar ligado a mais de um Conceito, *String* ou Átomo.

3.4.7 Átomos diferentes para representar um mesmo conceito

A Tabela 3.1 (Thesaurus 2007) apresenta a relação entre o conceito *Atrial Fibrillation*, suas variações léxicas, *Strings* e átomos. Este conceito possui 4 nomes ou sinônimos ou ainda *Strings* relacionadas: *Atrial Fibrillation*, *Atrial Fibrillations*, *Auricular Fibrillation* e *Auricular Fibrillations*. As duas primeiras *Strings* estão relacionadas ao termo léxico *Atrial Fibrillation* e as duas últimas ao termo léxico *Auricular Fibrillation*. Existem duas ocorrências da *String Atrial Fibrillation*: uma no sistema MeSH (A0027665) e outra no sistema PSY (A0027667). Ainda, existe uma ocorrência da *String Atrial Fibrillations* em MeSH (A0027668). Continuando, ocorre uma ocorrência da *String Auricular Fibrillation* em PSY (A0027930) e uma ocorrência da *String Auricular Fibrillations* no sistema MeSH (A0027932).

Este exemplo da Tabela 3.1 mostrou como o Metathesaurus da UMLS representa fielmente todas as formas diferentes encontradas nas fontes de vocabulários de se representar um mesmo conceito. Assim, este projeto faz uso da UMLS Metathesaurus através de um serviço Web durante o apoio à decisão diagnóstica compartilhada para que os dados sejam trocados carregando sempre consigo informações unificadas da UMLS. Com isso, pode-se por exemplo, encontrar todas as ocorrências de um conceito em particular, ou ainda saber se duas ocorrências com AUIs diferentes estão relacionados ao mesmo LUIs ou CUIs.

Tabela 3.1: Átomos diferentes para representar um mesmo conceito

| Conceito(CUI) | Termos(LUIs) | Strings(SUIs) | Átomos(AUIs) |
|---|--|--|---|
| C0004238 <i>Atrial Fibrillation</i> (preferred) <i>Atrial Fibrillations</i> <i>Auricular Fibrillation</i> <i>Auricular Fibrillati- ons</i> | L0004238 <i>Atrial Fibrillation</i> (preferred) <i>Atrial Fibrillations</i> | S0016668 <i>Atrial Fibrillation</i> (preferred) | A0027665 <i>Atrial Fibrillation</i> (from MSH) |
| | | S0016669 <i>Atrial Fibrillations</i> | A0027667 <i>Atrial Fibrillation</i> (from PSY) A0027668 <i>Atrial Fibrillations</i> (from MSH) |
| | L0004327 (synonym) <i>Auricular Fibrillation</i> <i>Auricular Fibrillati- ons</i> | S0016899 <i>Auricular Fibrillation</i> (preferred) | A0027930 <i>Auricular Fibrillation</i> (from PSY) |
| | | S0016900 (plural variant) <i>Auricular Fibrillati- ons</i> | A0027932 <i>Auricular Fibrillations</i> (from MSH) |

3.4.8 Átomos iguais para representar conceitos diferentes

Existem situações em que uma mesma *String* pode representar vários conceitos. Por exemplo, a *String Cold* é o nome para representar temperatura em um determinado vocabulário. Em outro vocabulário, *Cold* é o nome para resfriado, e em um terceiro vocabulário, *Cold* é um acrônimo para *Chronic obstructive lung disease*. Para representar e resolver esta ambiguidade, a UMLS os representa como mostra a Tabela 3.2 (Thesaurus 2007) a seguir.

De acordo com a Tabela, o léxico *Cold* (L0009264) está associado a três conceitos diferentes, *cold temperature*(C0009264), *Common Cold* (C0009443) e *Chronic Obstructive Airway Disease*. Especificamente em relação às *Strings* Cold e COLD relacionadas ao léxico em questão:

- a *String* S0026353 possui duas ocorrências, uma em MeSH com o átomo A0040712 representando temperatura fria, e a outra em COSTAR com o átomo A0040708 representando resfriado;
- a *String* S0007170 possui uma ocorrência em MTH com o átomo A0016032 para representar temperatura fria;
- a *String* S0007171 possui uma ocorrência em MTH com o átomo A0016033 para representar resfriado;
- a *String* S0829315 possui uma ocorrência em MTH com o átomo A0887858 para representar *Chronic Obstructive Airway Disease*;

De acordo com os exemplos apresentados na tabela anterior, pode-se saber, portanto, com o uso da UMLS MetaThesaurus se termos iguais fornecidos de fontes de terminologia diferentes possuem mesmos significados. Sendo assim, este projeto utiliza o Metathesaurus da UMLS através de um serviço Web para que os termos clínicos

Tabela 3.2: Átomos iguais para representar conceitos diferentes

| Conceitos(CUIs) | Termos(LUIs) | Strings(SUIs) | Átomos(AUIs) |
|---|---|---|---|
| C0009264 <i>cold temperature</i> | L0215040 <i>cold temperature</i> | S0288775 <i>cold temperature</i> | A0318651 <i>cold temperature</i> (from CSP) |
| | L0009264 <i>Cold</i> <i>Cold</i> | S0007170 <i>Cold</i> | A0016032 <i>Cold</i> (from MTH) |
| | | S0026353 <i>Cold</i> | A0040712 <i>Cold</i> (from MSH) |
| C0009443 <i>Common Cold</i> | L0009443 <i>Common Cold</i> | S0026747 <i>Common Cold</i> | A0041261 <i>Common Cold</i> (from MSH) |
| | L0009264 <i>Cold</i> <i>Cold</i> | S0007171 <i>Cold</i> | A0016033 <i>Cold</i> (from MTH) |
| | | S0026353 <i>Cold</i> | A0040708 <i>Cold</i> (from COSTAR) |
| C0024117 <i>Chronic Obstructive Airway Disease</i> | L0498186 <i>Chronic Obstructive Airway Disease</i> | S0837575 <i>Chronic Obstructive Airway Disease</i> | A0896021 <i>Chronic Obstructive Airway Disease</i> (from MSH) |
| | L0008703 <i>Chronic Obstructive Lung Disease</i> | S0837576 <i>Chronic Obstructive Lung Disease</i> | A0896023 <i>Chronic Obstructive Lung Disease</i> (from MSH) |
| | L0009264 <i>COLD</i> <i>COLD</i> | S0829315 <i>COLD</i> | A0887858 <i>COLD</i> (from MTH) |
| | | S0474508 <i>COLD</i> | A0539536 <i>COLD</i> (from SNMI) |

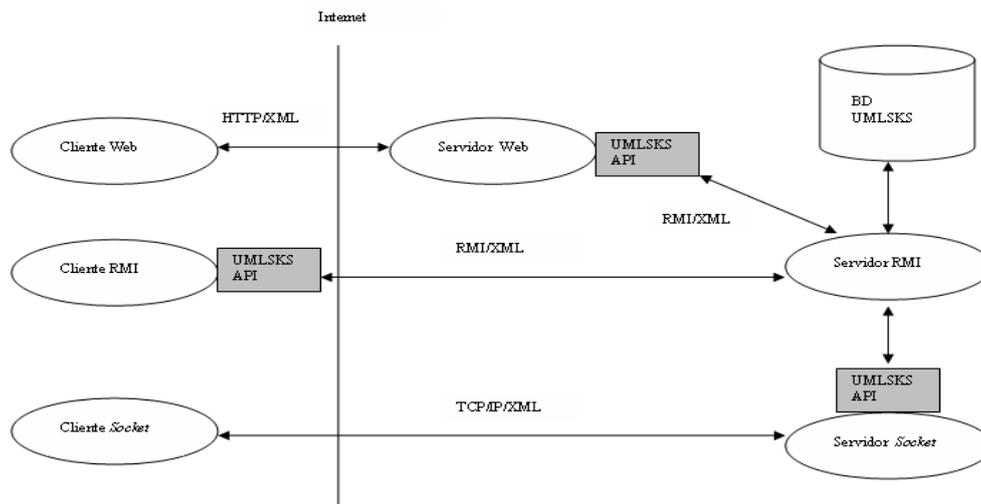


Figura 3.5: Formas de acesso a UMLSKS.

sejam trocados durante uma discussão diagnóstica carregando consigo identificadores da UMLS, como os identificadores do termo léxico, da String, do átomo, e do conceito.

3.4.9 UMLSKS API

A NLM fornece uma API Java que permite a manipulação e o acesso de aplicativos computacionais na base de dados da UMLSKS, especificamente a UMLS Metathesaurus. A forma de acesso pode ocorrer de três maneiras, como mostra a Figura 3.5:

1. Através de um navegador da Internet, que permite que o usuário faça uma requisição via HTTP ao servidor Web da NLM. Em seguida, um programa Java no servidor utiliza a UMLSKS API para fazer uma invocação remota de métodos em um servidor RMI, que por sua vez acessa a base de dados Oracle que contém os dados da UMLS;
2. A partir de um programa em Java que faça uso da UMLSKS API e do protocolo RMI, pode-se invocar o programa Java hospedado no servidor RMI, que por sua vez armazena e recupera dados da base da UMLS; e finalmente
3. Uma aplicação *socket* cliente pode estabelecer comunicação com um programa *socket* servidor, que por sua vez utiliza a UMLSKS API para chamar métodos remotos no servidor RMI, que por sua vez acessa o banco de dados.

Todas as três formas de acesso requerem um usuário e senha da UMLSKS.

O estudo de caso criado neste projeto e apresentado no Capítulo 6 faz uso da UMLSKS API através de um serviço Web para acessar a base de dados da UMLSKS. Para isso, a solução de número 2 foi escolhida em função do estudo de caso ter sido implementado em Java. Detalhes desta solução são apresentados no Capítulo 6.

A Tabela 3.3 apresenta alguns dos métodos da API UMLSKS utilizados em um dos Serviços Web implementados neste trabalho, mostrando uma descrição e exemplo de cada um deles. Especificamente, este Serviço Web, chamado de UMLSWs e

Tabela 3.3: Métodos da API UMLSKS

| Método | Descrição | Exemplo |
|--|---|--|
| <i>getConceptName(CUI, vocabulários)</i> | Retorna o nome do conceito relacionado ao identificador CUI encontrado nos vocabulários especificados | <i>getConceptName(C0009264, MSH)</i> Resp: <i>cold temperature</i> |
| <i>getTerminology(CUI, vocabulário)</i> | Retorna as informações de LUI e AUI dos termos encontrados de um CUI em uma fonte em particular | <i>getTerminology(C0009264, SNOMEDCT)</i> Cold; low temperature; Cold - thermal agent; low temperature (physical force) |
| <i>findCUI(nome do termo ou do conceito, vocabulários)</i> | Retorna o(s) identificador(es) CUI que possuem ocorrência nos vocabulários especificados | <i>findCUI(cold temperature)</i> Resp: C0009264 |
| <i>getConceptProperties(CUI, vocabulários)</i> | Retorna as informações de LUI e AUI dos termos, a definição, e o tipo semântico de um nome de conceito ou <i>String</i> | <i>getConceptProperties(C0009264, SNOMEDCT)</i> Resp: <i>cold temperature, Cold, Cold - thermal agent, Low temperature, Low temperature (physical force); An absence of warmth or heat or a temperature notably below an accustomed norm; Natural Phenomenon or Process</i> |
| <i>findTerms(termo, vocabulários)</i> | Retorna informações das ocorrências do termo nas fontes de vocabulário especificadas | <i>findTerm(Low temperature)</i> Resp: Read Codes (A1293673); SNOMEDCT (A3292554); Crisp Thesaurus (A1307780); ICPC2-CID10 Thesaurus (A4431983) e (A4461394) |
| <i>getAUIDetails(AUI)</i> | Retorna o valor da <i>String</i> , o valor do termo, e o valor do conceito de um AUI | <i>getAUIDetails(A0539536)</i> Resp: SUI - S0474508; COLD; SNOMED International |
| <i>findConcept(termo, vocabulários, idioma)</i> | Retorna um <i>Concept Vector</i> contendo vários objetos <i>Concept</i> , de acordo com o resultado da busca. Estes objetos trazem várias informações necessárias para uma troca de dados clínicos entre sistemas, como CUI e o tipo semântico. Este método é equivalente a chamar o método <i>findCUI</i> e em seguida chamar o método <i>getConceptProperties</i> . | <i>findConcept(cold temperature, SNOMEDCT, english)</i> Resp: CUI -C0009264; <i>Natural Phenomenon or Process</i> ; |

apresentado no Capítulo 5, pode ser utilizado por aplicativos médicos que desejam obter informações da UMLS referentes a termos clínicos ligados a sintomas, patologias e diagnósticos provenientes de uma fonte de vocabulário, por exemplo, MeSH, CID-10 e SNOMED CT. São exemplos destas informações o CUI, o LUI, o SUI, o AUI, e o tipo semântico. Os métodos apresentados pertencem à interface *KSSRetriever 5.0* da UMLSKS API.

Todos os métodos da UMLSKS API retornam a resposta na forma de um documento XML. Para que este documento possa ser manipulado, o pacote *UMLSKS Object Model* contém classes que permitem acessar valores específicos do XML, como a recu-

Tabela 3.4: Classes do pacote *UMLSKS Object Model*

| Nome da classe | Descrição |
|-------------------------|---|
| <i>ConceptVector</i> | Contém informações associadas a um conceito, como os termos léxicos, as definições, e os tipos semânticos |
| <i>DefinitionVector</i> | Contém definições de um conceito, bem como em qual fonte de vocabulário aquela definição foi criada |
| <i>SemTypeVector</i> | Contém tipos semânticos de um conceito, incluindo o identificador e o nome do conceito |
| <i>SourceInfo</i> | Contém informação de uma <i>String</i> em uma fonte de vocabulário, incluindo o AUI |
| <i>StringInfo</i> | Contém informações sobre as variações léxicas de um termo, incluindo SUI, a fonte do vocabulário, e o nome da <i>String</i> |
| <i>Term</i> | Contém informações de um termo, como o LUI, o nome, e seu idioma |
| <i>DefinitionVector</i> | Contém definições de um conceito, bem como em qual fonte de vocabulário aquela definição foi criada |

peração de informações de um conceito, de um termo, ou de um átomo. A Tabela 3.4 apresenta alguns exemplos destas classes juntamente com uma descrição destas.

Portanto, o uso da UMLKS API e conseqüentemente dos tipos de dados apresentados na tabela anterior para o envio e o recebimento das mensagens clínicas evita que os desenvolvedores necessitem entender o formato de dados usado pela UMLS. Ainda, não é necessário que o desenvolvedor domine XML visto que a própria API abstrai o processo de codificação e decodificação do mesmo.

3.5 Considerações Finais

Baseado nos estudos realizados em alguns sistemas de terminologias, especificamente CID, SNOMED e MeSH, verificou-se que estas soluções são interessantes para aplicações clínicas isoladas, mas não suficientes para aplicações computacionalmente distribuídas. Essa insuficiência é justificada principalmente em razão dos problemas de interoperabilidade apresentados neste capítulo.

Nesse contexto, foi necessário a investigação de uma solução de unificação terminológica. A UMLS MetaThesaurus mostra-se na literatura e particularmente neste estudo como uma solução de compatibilidade terminológica, já que consegue na maioria das vezes diferenciar termos clínicos diferentes a um mesmo conceito, ou relacionar termos iguais a conceitos diferentes. Como já foi mencionado, quando uma instituição de saúde possui seu próprio sistema de terminologia, e este não for contemplado pela UMLS, a compatibilidade terminológica ficará prejudicada.

A partir do estudo da UMLS Metathesaurus, verificou-se também que existem informações essenciais que devem ser também transferidas juntamente e relacionados com o termo clínico para promover a compatibilidade terminológica: o identificador único UMLS do átomo, o identificador único UMLS da *String*, o identificador único UMLS do termo léxico, o identificador único UMLS do conceito, o conceito UMLS, o tipo semântico UMLS, e a definição UMLS. Com este conjunto de informações este projeto criou um tipo de dados chamado de DDSOntADT, e está presente na ontologia DDSOnt, apresentada no capítulo 5. Todas as instâncias médicas criadas, como sintoma, diagnóstico ou patologia, devem ser do tipo DDSOntADT.

O próximo capítulo discute os problemas da incompatibilidade ontológica no cenário de apoio à decisão diagnóstica compartilhada na Web. Ainda, são apresentados alguns sistemas de ontologia em saúde, e justificativas da impossibilidade da utilização de alguns destes sistemas na solução proposta por este projeto. Por outro lado, o capítulo mostra por que a ontologia DDSOnt proposta para apoiar a decisão clínica precisa reutilizar dois tipos semânticos do sistema ontológico UMLS *Semantic Network*. Por último, são apresentados conceitos e recursos da Web Semântica.

Incompatibilidade Ontológica em Sistemas de Informação em Saúde Distribuídos

4.1 *Considerações Iniciais*

Este Capítulo tem como objetivo apresentar os problemas relacionados a incompatibilidade ontológica durante a troca de informações clínicas entre softwares médicos, especificamente para trocar informações de apoio à decisão diagnóstica em ambiente colaborativo. Em seguida, é apresentada uma discussão a respeito dos sistemas ontológicos em saúde, como a UMLS *Semantic Network* (UMLS SN), a Galen CRM e a HL7 RIM. Ainda, são discutidas como as especificações RDF e OWL da Web Semântica conseguem representar, instanciar e consultar informações ontológicas. Finalmente, este Capítulo apresenta de que forma o estudo dos sistemas ontológicos e dos padrões da Web Semântica contribuíram para alcançar os objetivos deste trabalho.

4.2 *Cenário Médico com Problemas de Incompatibilidade Ontológica*

É habitual aplicações clínicas fazerem uso de diferentes sistemas de ontologia médica em razão principalmente da dificuldade de se criar uma só ontologia padrão para todas as necessidades médicas. Consequentemente, as diferentes formas de se relacionar e de se representar a semântica, assim como os diferentes tipos semânticos tornam a compatibilidade ontológica entre as aplicações médicas muito difícil.

A Figura 4.1 ilustra um exemplo de softwares médicos utilizando sistemas ontológicos diferentes, tais como: HL7 RIM, Galen CRM, e UMLS SN. Pode-se perceber pela

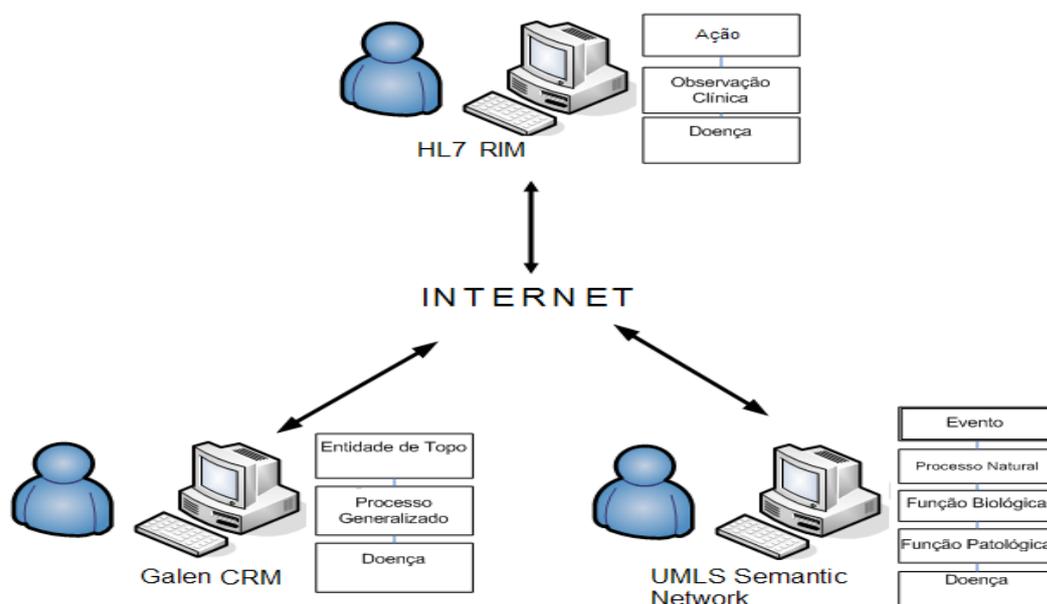


Figura 4.1: Aplicações utilizando sistemas ontológicos diferentes

figura que o conceito 'doença' é representado diferentemente por cada sistema. No sistema UMLS SN, a doença é classificada como da classe Função Patológica, que por sua vez é uma Função Biológica, que por sua vez é um Fenômeno ou Processo Natural, que é um Evento. No sistema Galen CRM, uma doença é classificada como da classe Processo Generalizado, que por sua vez é uma classe Fenômeno, que por sua vez é uma Entidade de Topo. Finalmente, no sistema ontológico HL7 RIM, uma doença é classificada como da classe Observação Clínica, que por sua vez é uma classe chamada Ação.

Especificamente no contexto deste trabalho, faz-se necessário estudar e projetar um mecanismo que procura permitir que aplicações médicas que trocam dados clínicos objetivando apoio diagnóstico colaborativo possam fazê-la de forma que não ocorra o problema apresentado no cenário anterior, ou seja, que as aplicações troquem suas informações usando a mesma semântica, o mesmo significado, entre si. Com este propósito, este trabalho apresenta os resultados da investigação de um estudo que verifica as características presentes em alguns dos sistemas de ontologia referenciados pela literatura. Estes sistemas ontológicos são apresentados a seguir.

4.3 Ontologias Clínicas

Na área da Computação, ontologia é uma especificação formal, ou seja, deve ser processada automaticamente, e explícita de uma conceitualização ou modelo abstrato compartilhado e de conhecimento consensual. O objetivo geral de se criar uma ontologia é descrever formalmente um domínio de aplicação de modo que o mesmo possa ser entendido por pessoas leigas no assunto.

Uma ontologia pode ser classificada em três principais categorias (Breitman 2005):

1. Ontologia de topo: descreve conceitos bastante genéricos, independentes de domínio, como eventos, espaço e tempo. São exemplos desse tipo de ontologia a

WordNet¹ e o dicionário Aurélio;

2. Ontologia de domínio: representa o vocabulário relativo a um domínio em específico, como da saúde, de educação, e de negócios. Exemplos desse tipo de ontologia são UMLS SN, a Galen CRM, e a HL7 RIM, discutidos nas próximas seções;
3. Ontologia de aplicação: os conceitos presentes nesta ontologia correspondem a papéis desempenhados por entidades de um domínio em específico durante a execução de alguma tarefa em particular. Exemplos desse tipo de ontologia são a DDSOnt, proposta neste trabalho.

Ontologias clínicas tem o propósito de fornecer uma categorização consistente de todos os conceitos representados em um tesouro controlado, e permitir a definição de um conjunto de relacionamentos entre estes conceitos, criando assim categorias e tipos semânticos. Exemplos atuais de ontologias médicas são UMLS SN, Galen CRM e HL7 RIM. Elas são descritas a seguir e nesta mesma ordem.

4.3.1 UMLS SN

O propósito da rede semântica da UMLS ([System 2006](#)) é de fornecer uma categorização consistente de todos os conceitos do UMLS *Metathesaurus*, apresentada na seção 3.4, e de prover os relacionamentos entre estes conceitos.

A rede semântica fornece informações sobre os tipos semânticos que podem ser associados aos conceitos e define um conjunto de relacionamentos entre os tipos semânticos. Essa rede contém 135 tipos semânticos e 54 relacionamentos ([SemanticNetwork 2006](#)).

Os nós da rede semântica da UMLS são representados pelos tipos semânticos, e as ligações existentes entre os nós fornecem os tipos de relacionamento existentes na rede. A categorização semântica da UMLS é muito ampla, e cobre múltiplos domínios de terminologia, como por exemplo, organismos, estruturas anatômicas, funções biológicas, químicas, eventos clínicos, objetos físicos, conceitos e idéias.

Como apresentado na seção 3.4, o UMLS *Metathesaurus* consiste de termos presentes nos mais de 100 fontes de vocabulários. Estes termos são associados aos conceitos do tesouro, que por sua vez estão associados a pelo menos um tipo semântico. Em todos os casos, os conceitos são relacionados aos tipos semânticos mais específicos possível. Por exemplo, o conceito 'Macaco' recebe o tipo semântico 'Mamífero' porque não existe um tipo mais específico como 'Primata' nesta rede semântica ([SemanticNetwork 2006](#)).

O nível de granularidade varia pela rede, o que provoca implicações importantes durante a interpretação dos significados dos conceitos da UMLS *Metathesaurus*. Por exemplo, o nó filho de 'Objeto Físico', 'Objeto Manufaturado', possui 3 nós filhos, 'Dispositivo Médico', 'Dispositivo de Pesquisa' e 'Remédio Clínico'. Aos objetos manufaturados diferentes de dispositivos médicos, dispositivos de pesquisa e remédios clínicos, são classificados como simplesmente objeto manufaturado. Esse tipo de classificação inibe a proliferação da quantidade de tipos semânticos ([SemanticNetwork 2006](#)).

Percebe-se aqui como pode ser difícil para todas as instituições em saúde usarem uma mesma ontologia, já que os sistemas ontológicos utilizam-se de diferentes granu-

¹<http://wordnet.princeton.edu/>

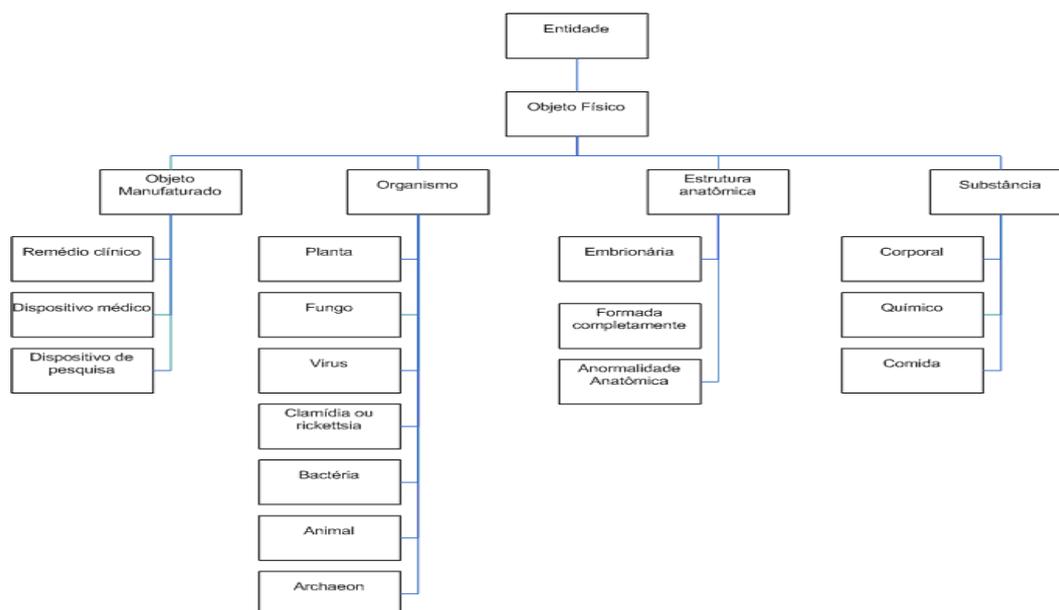


Figura 4.2: Porção da rede semântica UMLS: Entidades do tipo objeto físico.

laridades para diferentes áreas, dependendo do domínio de aplicação sendo coberto. Estas diferenças de granularidade podem impedir especificações mais detalhadas de algum tipo semântico que eventualmente seja importante para a aplicação que utiliza o sistema ontológico, mas que seja indiferente para a ontologia sendo utilizada. Por exemplo, para a UMLS SN, não é importante criar especializações para o tipo semântico ‘Mamífero’, e nem criar mais de 3 especializações para o tipo semântico ‘Objeto Manufaturado’. Se para uma aplicação clínica que faz uso da UMLS SN for importante uma especialização de ‘Mamífero’ como o tipo ‘Primata’ ou mais especializações para o tipo ‘Objeto Manufaturado’, está ficará impedida de representar semanticamente as informações associadas.

De modo a ilustrar a rede semântica da UMLS ², as Figuras 4.2 e 4.3 são apresentadas. Na Figura 4.2, são mostrados os nós filhos do tipo semântico ‘Objeto Físico’, filho do tipo semântico ‘Entidade’.

O tipo semântico ‘Objeto Físico’ é especializado em ‘Objeto Manufaturado’, ‘Organismo’, ‘Estrutura Anatômica’, e ‘Substância’. Como já citado, ‘Objeto Manufaturado’ é dividido em ‘Remédio Clínico’, ‘Dispositivo Médico’ e ‘Dispositivo de Pesquisa’. O tipo semântico ‘Organismo’ é especializado em ‘Plantas’, ‘Fungos’, ‘Vírus’, ‘Clamídia’, ‘Bactéria’, ‘Animal’, e ‘Archaeon’. Já o tipo semântico ‘Estrutura Anatômica’ é especializada em ‘Embrionária’, ‘Formada Completamente’, e ‘Anormalidade Anatômica’. Finalmente, o tipo semântico ‘Substância’ é especializado em ‘Corporal’, ‘Químico’, e ‘Comida’. Todos os relacionamentos nesta parte da rede semântica são do tipo ‘é-um’, ou seja, é uma hierarquia do tipo generalização/especialização.

Na Figura 4.3, mostra-se os nós filhos do tipo semântico ‘Entidade Conceitual’, filho do tipo semântico ‘Entidade’. O tipo semântico ‘Entidade Conceitual’ é especializado em ‘Atributo do Organismo’, ‘Descoberta’, ‘Idéia ou Conceito’, ‘Ocupação’, ‘Organização’, ‘Produto Intelectual’, ‘Grupo’ e ‘Atributo de Grupo’. O tipo semântico ‘Atributo do Organismo’ especializa-se em ‘Atributo Clínico’. A ‘Descoberta’ divide-se em ‘Resultado de Laboratório’ e ‘Sintoma’.

²As ilustrações da UMLS SN apresentadas neste trabalho foram realizadas utilizando tradução livre, pois não têm-se ainda uma rede semântica em português

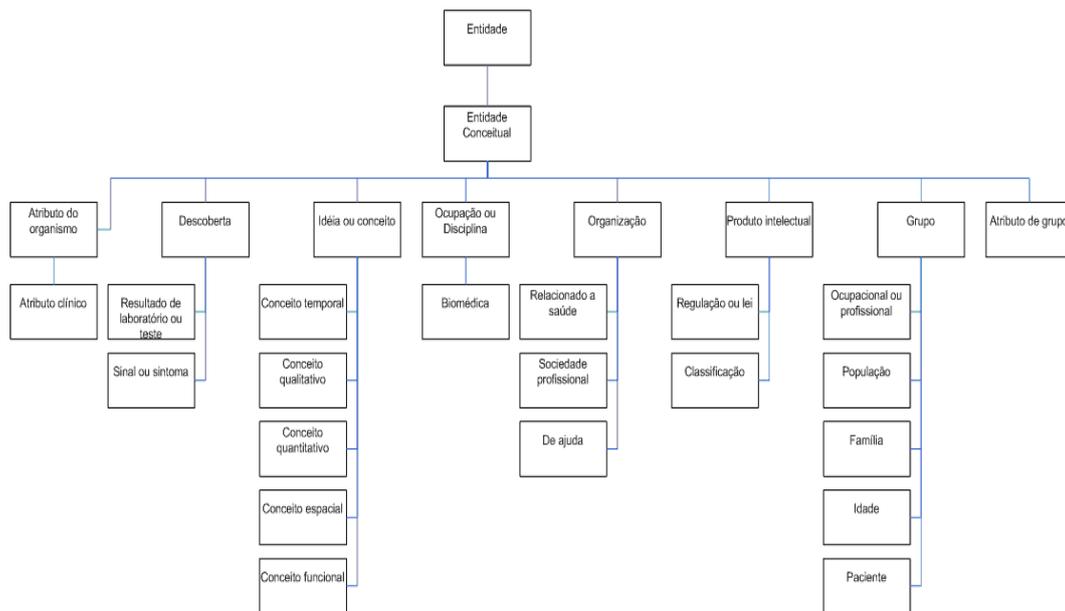


Figura 4.3: Porção da rede semântica UMLS: Entidades do tipo conceitual

O tipo semântico 'Idéia ou Conceito' especializa-se em 'Conceito Temporal', 'Qualitativo', 'Quantitativo', 'Espacial' e 'Funcional'. A 'Ocupação' é especializada em 'Biomédica'. O tipo semântico 'Organização' é dividido em 'Relacionado a Saúde', 'Sociedade Profissional' e 'De Ajuda'. 'Produto Intelectual' é dividido em 'Regulação ou Lei', e 'Classificação'. Finalmente, o tipo 'Grupo' é especializado em 'Ocupacional', 'População', 'Família', 'Idade', e 'Paciente'. Todos os relacionamentos nesta parte da rede semântica também são do tipo 'é-um'.

Além das entidades 'Objeto Físico' e 'Conceitual', a rede semântica faz a modelagem também dos tipos semânticos relacionados a eventos. A Figura 4.4 ilustra essa parte da rede semântica. Os eventos podem se dividir em 'Atividade' e 'Fenômeno ou Processo'. As atividades podem ser do tipo 'Comportamento Social' ou 'Individual', 'Atividades Ocupacionais', como 'Em Saúde', 'Em Pesquisa', 'Governamental', e 'Educativa', ou ainda 'Recreação', e 'Relacionado a Máquinas'. Os 'Fenômenos' são classificados em 'Ferimento', 'Efeitos Ambientais Causados pelo Homem', ou 'Natural'. Os fenômenos naturais do tipo biológicos são divididos em 'Função Fisiológica' e 'Função Patológica'. Os tipos de função fisiológica são 'Celular', do 'Organismo', do 'Órgão ou Tecido', e 'Molecular'. Finalmente, os tipos de função patológica são 'Modelo Experimental da Doença', 'Doença ou Síndrome', e 'Disfunção Celular'. Todos os relacionamentos nesta parte da rede semântica também são do tipo 'é-um'.

As relações que representam significado diferente do 'é-um' entre os tipos semânticos também são identificadas na rede semântica da UMLS. Estes tipos de relacionamento são agrupados em 5 categorias, que são: 'Fisicamente Relacionado A', 'Espacialmente Relacionado A', 'Temporalmente Relacionado A', 'Funcionalmente Relacionado A', e 'Conceitualmente Relacionado A'. A Figura 4.5 ilustra uma porção da rede semântica da UMLS que fornece o relacionamento entre tipos de relação. Para cada tipo de relacionamento, existem várias especializações, como mostra a Figura.

As relações são colocadas dentro do possível entre os tipos semânticos de mais alto nível na rede de modo que estas possam ser herdadas para os tipos semânticos mais abaixo na hierarquia. Assim, por exemplo, o relacionamento 'Processo De', filho da associação funcional 'Ocorre Em', relaciona os tipos semânticos 'Função Biológica'

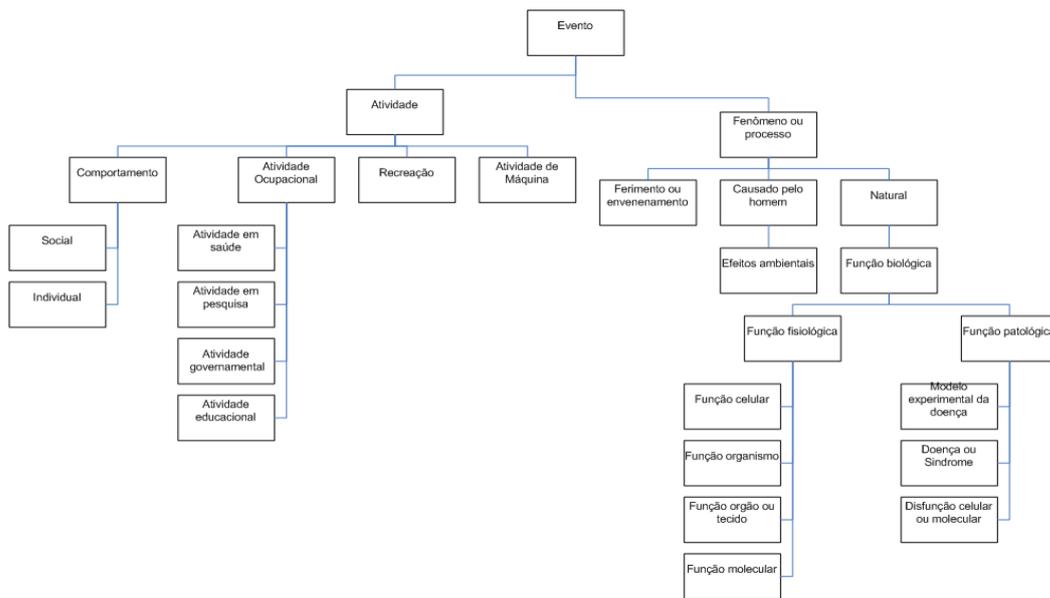


Figura 4.4: Porção da rede semântica UMLS: Tipos semânticos relacionados a evento.

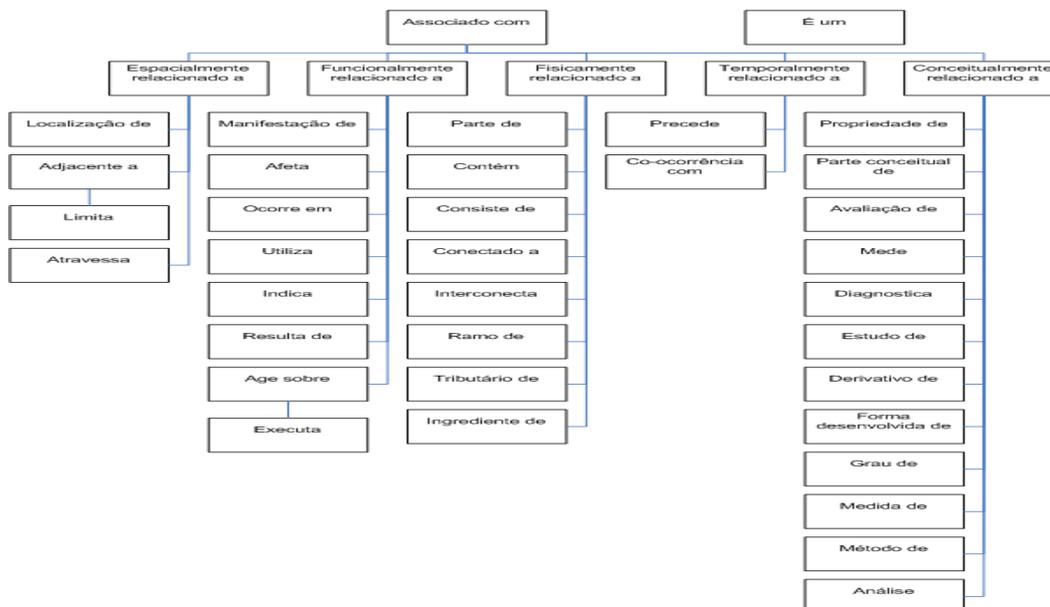


Figura 4.5: Porção da rede semântica UMLS: Relacionamento entre os tipos de relação

e 'Organismo'. Sendo assim, por herança, esse mesmo relacionamento também relaciona o tipo semântico 'Função do Órgão ou Tecido' (que é uma 'Função Fisiológica', que por sua vez é uma 'Função Biológica', veja a Figura 4.4) e 'Animal', que é um 'Organismo', como mostra a Figura 4.2.

Como apresentado anteriormente, as relações são colocadas entre os tipos semânticos de mais alto nível, mas não necessariamente se aplicam a todas as instâncias dos conceitos que foram associados àqueles tipos semânticos. Sendo assim, a relação pode ou não relacionar qualquer par de conceitos. O relacionamento 'Avaliação De' mostrado nas Figuras 4.5 e 4.6 relaciona os tipos semânticos 'Sintoma' e 'Atributo do Organismo', presentes na Figura 4.3. Um sintoma ou um atributo em particular podem não estar associados a esta relação. Como exemplo, os sinais 'Excesso de Peso' e 'Febre' são avaliações respectivamente de 'Peso Corporal' e 'Temperatura do Corpo'. Entretanto, 'Excesso de Peso' não é uma avaliação de 'Temperatura do Corpo', e nem 'Febre' é uma avaliação de 'Peso Corporal'.

A Figura 4.6 apresenta uma parte da rede semântica que relaciona hierarquicamente e associativamente os tipos semânticos da UMLS. Nesta figura, os três ramos ('Objeto Físico', 'Entidade Conceitual' e 'Evento') estão presentes. Em alguns casos, existem conflitos entre os tipos semânticos na rede e as relações que são herdadas. Nestes casos, a herança da relação é explicitamente bloqueada. Como exemplo, por herança, o tipo semântico 'Processo Mental', uma função fisiológica do organismo, como mostra a Figura 4.4, seria um 'Processo de Planta', como mostra a Figura 4.6. Como plantas não possuem sentido humano, essa relação é bloqueada.

Em outros casos, a natureza da relação mostra-se de tal forma que a mesma não deve ser herdada para os filhos dos tipos semânticos que ela relaciona. Neste caso, a relação é definida para os dois tipos semânticos que ela relaciona, mas é bloqueada para todos os filhos destes dois tipos. Por exemplo, o relacionamento 'Parte Conceitual De' relaciona 'Sistema Corporal' e 'Estrutura Anatômica Formada Completamente', como ilustra a Figura 4.6, mas a relação não vale para todos os filhos de 'Estrutura Anatômica Formada Completamente', como 'Célula' ou 'Tecido', como mostra a Figura 4.6.

Outros exemplos de relacionamento entre tipos semânticos podem ser retirados da Figura 4.6. Um 'Ferimento' ou 'Envenenamento' interrompe uma 'Estrutura Anatômica Formada Completamente' e também um 'Função Fisiológica'. Uma 'Descoberta' é uma avaliação de um 'Atributo do Organismo' e também de uma 'Função Fisiológica'. Um 'Gene' é parte de um 'Componente da Célula', que por sua vez é parte de uma 'Célula', que por sua vez é parte de um 'Tecido', e que por sua vez é parte do 'Corpo'.

No contexto deste trabalho, a UMLS SN pode ser utilizada para apoiar a decisão diagnóstica e promover a compatibilidade ontológica. Esta promoção torna-se importante pois as aplicações que precisam compartilhar dados clínicos têm de representá-los semanticamente, e ainda esta representação deve ser uniforme. Sendo assim, este trabalho procura reutilizar a UMLS SN e criar uma nova ontologia que estende esta, denominada DDSOnt, e apresentada no próximo Capítulo. A DDSOnt tem como objetivo definir um novo conjunto de tipos semânticos necessários para a criação de uma estrutura que suporte a manipulação de bases de conhecimento que são utilizadas por médicos durante a decisão diagnóstica.

A reutilização da UMLS SN através da DDSOnt visa principalmente a criação de uma solução que promova a compatibilidade ontológica. Entretanto, apesar deste trabalho se beneficiar dos tipos semânticos criados a partir dos mais de 100 sistemas

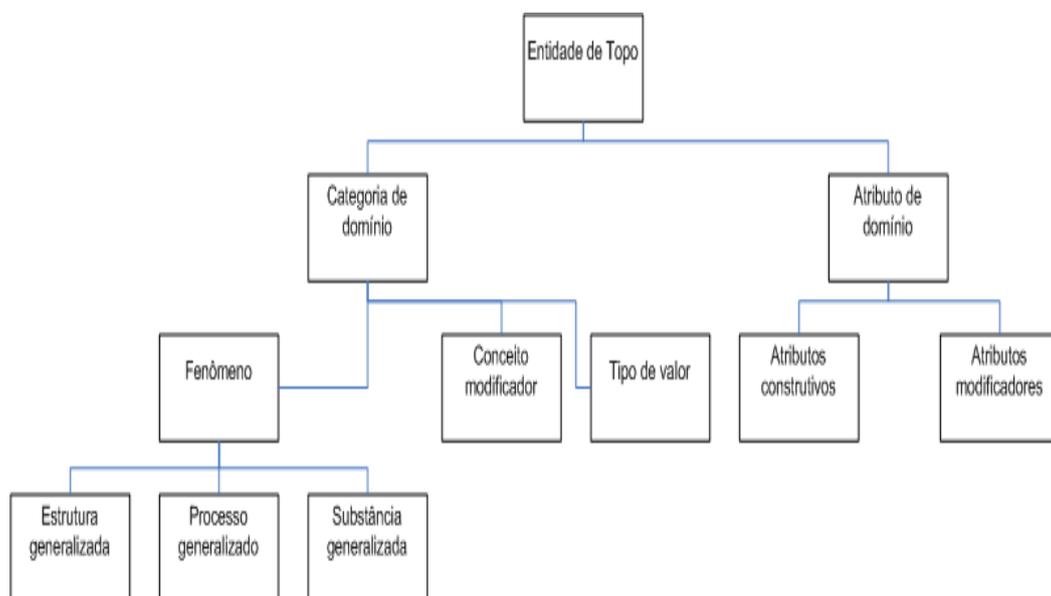


Figura 4.7: Galen CRM de alto nível.

entidades são apresentadas na sequência.

Entidades destas três categorias geralmente relacionam-se entre si utilizando-se de uma família de ligações semânticas conhecida como 'Atributos Construtivos', que fazem parte da entidade 'Atributos de Domínio'. Estas ligações incluem 'éParteDe', 'temFunção', 'éConsequênciaDe', 'temDivisãoSólida', 'temSubProcesso', 'serve' (Artéria que serve o Fígado), 'contém', 'temLocalização', 'ageEm', 'éFunçãoDe', 'temAssociação-CausalCom', e outras mais.

Existem também na ontologia CRM as propriedades dos fenômenos, encontradas na entidade 'Conceito Modificador'. Os conceitos modificadores são propriedades, ou características das estruturas, substâncias e processos. Os conceitos modificadores dos fenômenos podem ser alterados utilizando-se a família de ligações conhecida como 'Atributos Modificadores', também da entidade 'Atributos de Domínio'. Exemplos destes atributos modificadores são 'temCaracterística', 'temEstado', 'temStatus', 'temSeletor', 'temModalidade', e 'temRegra', 'temForma', 'temFrequência', 'temSeleto-*EsquerdaDireita*', e outros mais. Finalmente, a ontologia de alto nível fornece tipos de valor como valores inteiros e ordinários.

A Figura 4.8 apresenta as especializações dos fenômenos. A 'Estrutura Generalizada' divide-se em 'Estrutura Abstrata' e 'Estrutura Física'. São exemplos de estruturas abstratas as 'Construções Psico-Sociais Clínica e Hospital', e as 'Estruturas Lógicas Protocolo e Plano'. São exemplos de estrutura física: a 'Estrutura Linear Rota', a 'Estrutura Planar Triângulo', a 'Estrutura Microscópica Microorganismo', a 'Estrutura Sólida e Inerte Dispositivo', a 'Parte do Corpo Perna', a 'Parte Genérica do Corpo Rim', o 'Organismo Bactéria', a 'Região Sólida' de um Pedaco do Fígado e o 'Espaço Cavidade'. O 'Processo Generalizado' é dividido em 'Processo Específico' e 'Processo Genérico'. Um exemplo de 'Processo Genérico' é transporte. Exemplos de 'Processo Específico' são o 'Processo do Corpo' respiração e coagulação, o 'Processo Comportamental' de ato voluntário ou clínico, o 'Processo Físico' de irradiação, e o 'Processo Químico' de mancha histológica. Finalmente, a 'Substância Generalizada' divide-se em 'Energia' e 'Substância'. Exemplos de 'Energia' são radiação e efeitos do som. São exemplos de 'Substância' o tecido do músculo, substâncias nomeadas

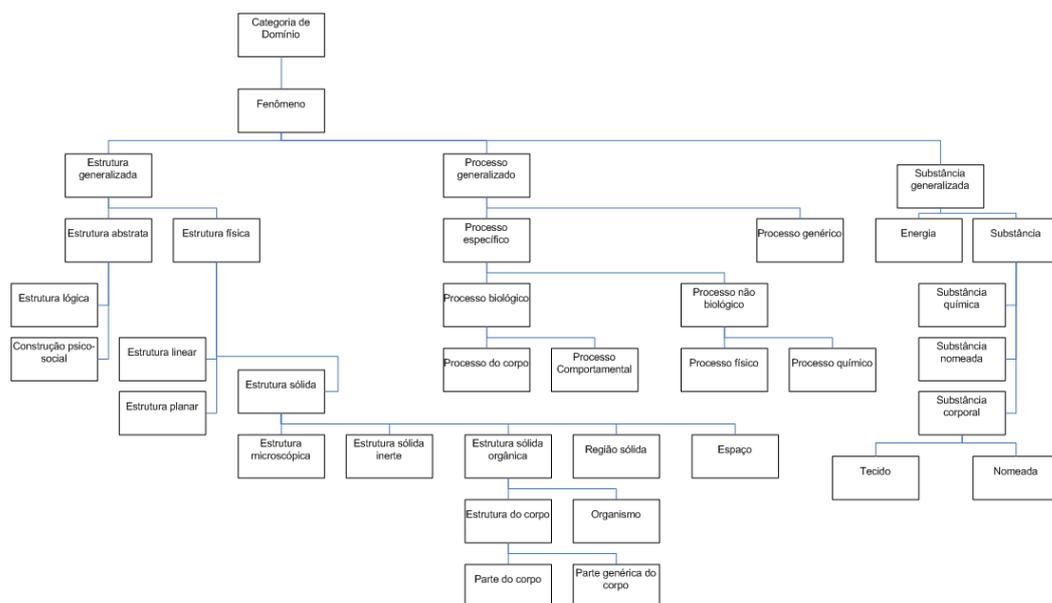


Figura 4.8: Porção da Galen CRM: Especialização dos fenômenos

como urina e bile, substância química com sódio, e substâncias nomeadas como ar e madeira.

Os conceitos modificadores dizem respeito às propriedades dos fenômenos. A Figura 4.9 ilustra como os modificadores são divididos. Os 'Aspectos', que refinam o significado da categoria, dividem-se em 'Característica', 'Estado', 'Seletor' e 'Status'. São exemplos de 'Característica': o sexo, a cronicidade e a forma. São exemplos de 'Estado': masculino/feminino e agudo/crônico. São exemplos de 'Seletor': o seletor para direita ou esquerda. E exemplos de 'Status' são: patológico/fisiológico e normal/anormal. Continuando a divisão dos modificadores, existem as 'Unidades', como segundos e metros, as 'Modalidades', como histórico familiar estar presença/ausente, as 'Regras', como dos hormônios e das drogas, as 'Coleções', como múltiplas manchas constituindo uma brotoeja, e finalmente os 'Níveis de Especificação', que relacionam às instâncias os requisitos linguísticos para se determinar quando utilizar um artigo definido ou indefinido (Rector et al. 1996).

Existem três formas de se representar os aspectos dos fenômenos: utilizando pares 'Característica-Estado', como 'Seletores e, como 'Status'. Independente do tipo de representação, todas são formalmente descritas na linguagem GRAIL. A linguagem de representação de conhecimento GRAIL é utilizada para formalizar a ontologia CRM, e ainda criar documentos clínicos com conteúdo relacionado semanticamente à CRM. Um exemplo de par 'Fenômeno-Estado' escrito em GRAIL é *Patient which (hasFeature Sex which (hasState Masculino))*. Dois bons exemplos com seletores são *Hand which hasLaterality Left* e *Hand which hasPosition Left*.

Finalmente, dois exemplos de 'Status' são '*Secretion which hasPathologicalStatus pathological*' e '*Ventilation which hasAbnormalityStatus nonNormal*' (Rector et al. 1996). Os documentos clínicos criados com estes exemplos apresentados possuem uma estrutura e organização estabelecida conforme as regras léxicas e sintáticas formais da linguagem GRAIL⁴.

Comparando a ontologia CRM da Galen com a ontologia UMLS SN, verifica-se

⁴<http://www.opengalen.org/index.html>

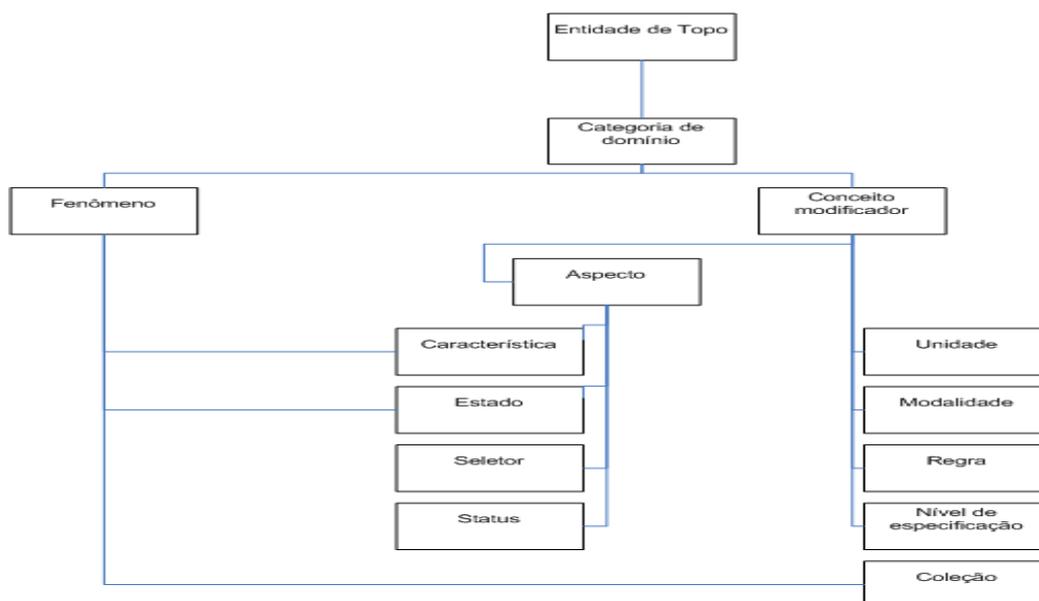


Figura 4.9: Porção da Galen CRM: Hierarquia dos modificadores.

que a primeira cria centenas de tipos semânticos a partir das experiências do grupo que a criou ⁵. Já a UMLS SN define 135 tipos semânticos a partir de conceitos provenientes de mais de 100 sistemas de terminologia. Mesmo que as diferenças de granularidade sejam em geral maior na rede semântica da UMLS SN, acredita-se que a compatibilidade ontológica seja mais frequentemente conseguida no cenário deste projeto com o uso da UMLS SN, já que a compatibilidade terminológica é promovida neste trabalho com a UMLS MetaThesaurus.

A API JDDSOnt que será descrita no Capítulo 5 pode ser comparada à linguagem GRAIL da ontologia Galen CRM, uma vez que ambas possibilitam a criação de documentos clínicos contendo informações que estão associadas diretamente a suas respectivas ontologias. Entretanto, ao contrário da GRAIL, a JDDSOnt permite a criação de documentos clínicos em XML e RDF, suportando e promovendo assim a interoperabilidade sintática entre as aplicações médicas. Ainda, de acordo com o estudo realizado na GRAIL, não existe uma regra formal específica para transferir documentos para apoiar a decisão diagnóstica como propõe a DDSOnt.

4.3.3 Modelo HL7 RIM

4.3.3.1 Introdução a HL7

HL7 é uma organização que desenvolve padrões e especificações na área de saúde com o objetivo de promover a integração entre sistemas de informações médicas. Nesse sentido, aplicações médicas podem trocar, gerenciar e compartilhar dados clínicos através de metodologias especificadas e padronizadas por essa organização (Seven 2006).

HL7 é um padrão ANSI e reconhecido por um órgão que regulamenta padrões (SDO) na área de saúde, especificamente relacionado a dados clínicos e administra-

⁵<http://www.opengalen.org/faq/faq1.html#1.9>

tivos. O termo HL7 é uma referência ao termo mais alto do modelo de comunicação ISO, que é a camada 7 do modelo OSI, a camada de aplicação. Este nível tem como objetivo definir o mecanismo de troca de dados, e principalmente, formalizar a estrutura da troca destes.

A primeira versão aprovada como um padrão ANSI foi a HL7 V2, ou HL7 *Version 2.5* em junho de 2003. Orientada à definição e à estruturação de mensagens clínicas trocadas entre instituições de saúde, também conhecidas como CDA, essa versão, assim como as anteriores, define mensagens com conteúdo relativos a, segundo o documento (Seven 2006):

- Controle das mensagens;
- Administração de pacientes, como a admissão, o pagamento, e a transferência desses;
- Entrada de serviços e observações clínicas, farmacêuticos, alimentares e de suplementos;
- Definição de regras aplicadas a consulta clínica e suas respostas;
- Gerenciamento financeiro das contas e pagamento dos pacientes;
- Relatórios de observação clínica;
- Marcação de exames, de consultas e de outros procedimentos clínicos;
- Registros orientados a problemas

Apesar da versão 2.5 abordar um grande conjunto de assuntos clínicos, três problemas básicos ainda eram evidentes:

1. a falta de uma linguagem de marcação que auxiliasse a estruturar os documentos ou mensagens clínicas;
2. a falta de um modelo de informação que pudesse associar as informações contidas nas mensagens à conceitos e tipos semânticos clínicos mais abrangentes; e
3. ainda por não ser uma solução orientada a objetos.

A versão HL7 V3, ou HL7 *version 3.0*, datada de setembro de 2003, contempla os três problemas citados anteriormente. Nesta versão, existe um documento XML Schema, pertencente à arquitetura CDA, explicada mais adiante, que formaliza e define como deve ser a estrutura dos documentos clínicos. A versão V3 também inclui um modelo orientado a objetos denominado RIM que permite que as mensagens clínicas possam ser derivadas do modelo orientado a objetos. Esse modelo é apresentado a seguir.

4.3.3.2 *Reference Information Model*

O modelo de referência de informação HL7 RIM (Model 1994) é uma parte essencial da metodologia de desenvolvimento da versão 3.0 do padrão HL7. O RIM fornece uma

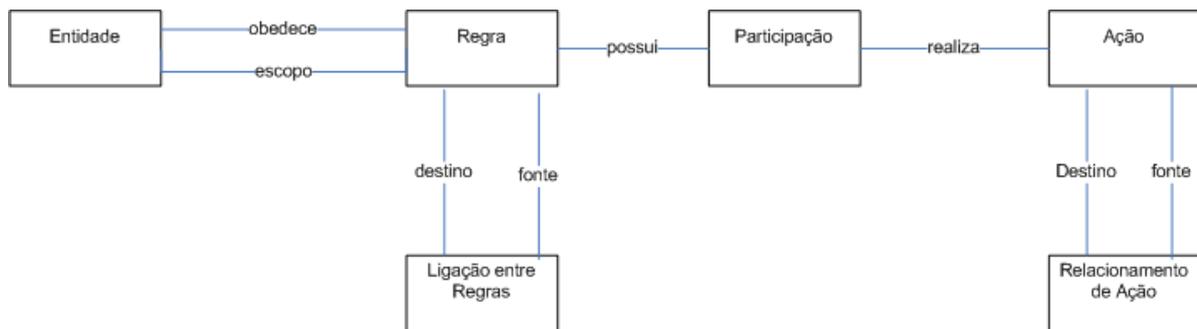


Figura 4.10: Classes principais do modelo RIM ((Beeler et al. 2003))

representação explícita das conexões semânticas que existem na informação contida nos campos da mensagem HL7.

Existem dois assuntos principais abordados pelo modelo RIM: as infra-estruturas normativa e de comunicação. A normativa descreve a gramática do HL7, enquanto que a de comunicação define as mensagens e os documentos estruturados para permitir que as aplicações em saúde comuniquem-se entre si.

As classes principais do modelo RIM e que compõem a infra-estrutura normativa são(Beeler et al. 2003):

- ‘Ação’: representa as ações executadas em um ambiente hospitalar;
- ‘Participação’: descreve quem, para que e aonde foi realizada a ação;
- ‘Entidade’: são os agentes físicos e humanos que são de interesse e participam das ações em saúde, como por exemplo seringa, paciente, grupo e substância química;
- ‘Regra’: estabelece as regras que as Entidades executam quando desempenham uma Ação;
- ‘Relacionamento de Ação’: relaciona as ações, detecta as ações geradas por outras ações. Como exemplo, mostra o relacionamento entre o pedido de uma observação e o evento ocorrido em consequência da execução da observação;
- ‘Ligação entre Regras’: relacionamento entre regras. Como exemplo, pode-se relacionar as regras ‘gerente’ e ‘analista’ através da relação ‘tem autoridade’. Assim, pode-se afirmar que o gerente tem autoridade direta sobre um analista;

A Figura 4.10 apresenta o relacionamento entre as classes explicadas anteriormente⁶. Um exemplo de uma instrução clínica sendo criada a partir destas classes e seus relacionamentos é: a ‘Entidade’ pessoa Fulano de Tal executa a ‘Regra’ Paciente enquanto ‘Participa’ como sujeito da ‘Ação’ de exame de um resfriado, ou seja, o paciente Fulano de Tal está resfriado conforme o exame realizado.

De modo a ilustrar as especializações de uma ‘Entidade’, a Figura 4.11 apresenta as classes ‘Objeto Vivo’, ‘Lugar’ (cidade, país, estado), ‘Organização’ (companhias, instituições) e ‘Material’. Um objeto vivo pode ser um ser-humano ou não-humano

⁶As ilustrações da HL7 RIM apresentadas neste trabalho foram realizadas utilizando tradução livre, pois não têm-se ainda uma versão em português

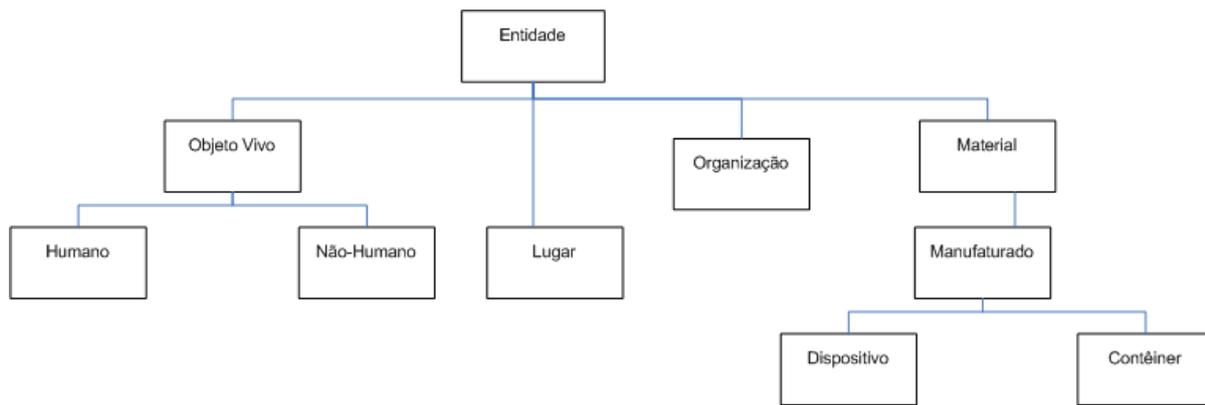


Figura 4.11: Classes Especializadas de uma 'Entidade' do modelo RIM ((Beeler et al. 2003))

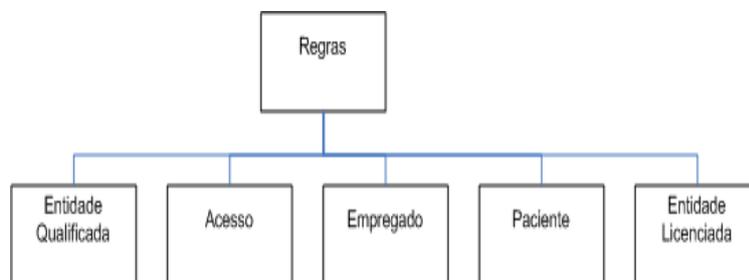


Figura 4.12: Classes Especializadas de uma 'Regra' do modelo RIM ((Beeler et al. 2003))

(bactérias, fungos). Os materiais manufaturados podem ser um dispositivo médico (equipamentos duráveis e descartáveis) ou ainda um contêiner (guarda outras entidades para fins de transporte ou proteção, como catéter e cânula).

De modo também a ilustrar as especializações de uma 'Regra', a Figura 4.12 apresenta as classes 'Acesso', 'Empregado', 'Paciente', e as 'Entidades Licenciada' e 'Qualificada'. A regra 'Acesso' é aplicada em dispositivos quando administram ações terapêuticas no corpo ou materiais de drenagem. A regra 'Empregado' é aplicada a qualquer pessoa dentro de uma organização e que recebe salário. A regra 'Paciente' é aplicada às entidades seres-humanos. A regra 'Entidade Licenciada' é aplicada a entidades que estão certificadas em desempenhar funções específicas (licença para um profissional, certificado para um equipamento). Finalmente, a 'Entidade Qualificada' é associada a entidades que sejam reconhecidas como tendo experiência para educar ou qualificar outras entidades.

A parte da ontologia HL7 RIM que mapeia as ações clínicas está ilustrada na Figura 4.13. São tipos de ações: 'Suprimento' (pedir lençóis de cama, preparar medicamentos), 'Tarefa de Dispositivo' (tarefa executada espontaneamente por um dispositivo automatizado), 'Encontro com Paciente' (estadia de um paciente, suporte familiar a saúde), 'Exame' (documento clínico contendo informações minuciosas de um assunto clínico), 'Lista de Trabalho' (listas de objetivos, lista de alergias), 'Administração de Substâncias' (regras de quimioterapia, prescrição de drogas, registro de vacinação), 'Conta' (contas de paciente), 'Contrato Financeiro' (contrato de seguros), 'Transação Financeira' (custo de um serviço, pagamento de fatura), 'Ação de Controle'

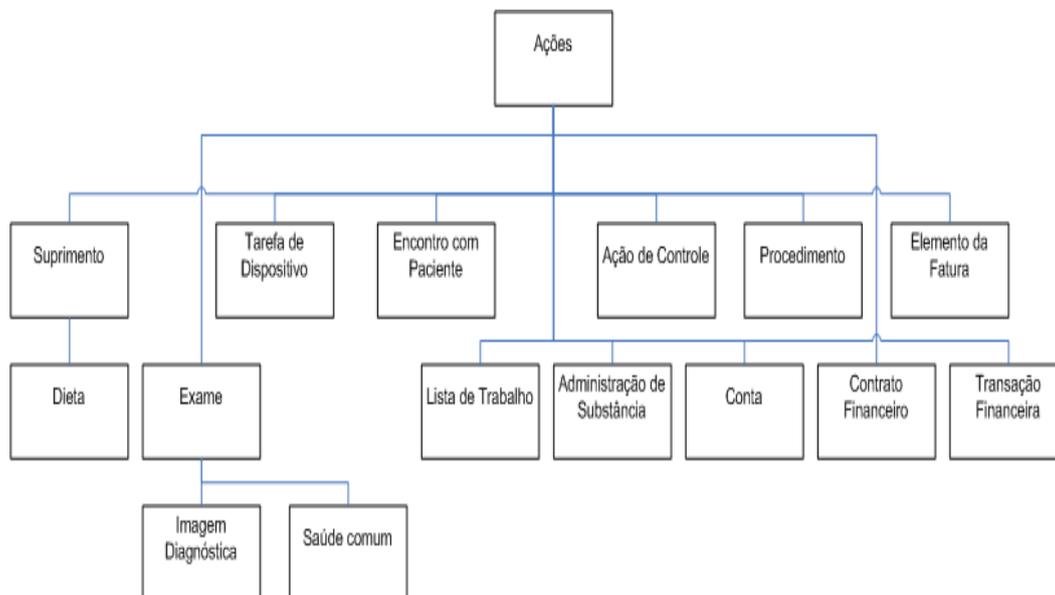


Figura 4.13: Classes Especializadas de uma 'Ação' do modelo RIM ((Beeler et al. 2003))

(interrupção de um medicamento, envio de sumários diários), 'Procedimento' (incisão em cirurgias, acupuntura, massagem), e 'Elemento da Fatura' (justificativa do valor devido na fatura).

Comparando a ontologia RIM do HL7 com a ontologia UMLS SN, verifica-se que a primeira cria seus tipos semânticos a partir das experiências do grupo que a criou. Não foram encontradas informações quantitativas a respeito dos tipos semânticos do HL7 RIM. Já a UMLS SN define 135 tipos semânticos a partir de conceitos provenientes de mais de 100 sistemas de terminologia. Espera-se portanto que, mesmo que as diferenças de granularidade sejam em geral maior na rede semântica da UMLS SN, acredita-se que a compatibilidade ontológica seja mais frequentemente conseguida no cenário deste projeto com o uso da UMLS SN, já que a compatibilidade terminológica é promovida com a UMLS MetaThesaurus.

Clinical Document Architecture

As informações clínicas trocadas entre aplicações que se utilizam do padrão HL7 devem estar associadas à ontologia RIM e estruturadas em documentos XML, obedecendo o documento XML Schema (Dolin et al. 2001). O documento XML Schema ou CDA Schema é parte do nível 1 do padrão ANSI CDA que propõe uma arquitetura de documentos clínicos em multi-camadas. Os documentos clínicos de nível 2 são construídos a partir do documento Schema do nível 1, e o documentos clínicos do nível 3 são construídos a partir dos documentos do nível 2. Assim, pressupoem-se que os documentos de nível mais baixo são mais básicos que os superiores, que por sua vez possuem elementos de marcação com com mais especificidade e granularidade que os documentos de nível inferior (Dolin et al. 2005).

O documento de nível 1 do padrão CDA HL7 consiste basicamente de três especificações técnicas sem informação semântica(Dolin et al. 2005):

1. Informações de cabeçalho, como autor do documento e sua instituição;

2. Informações do corpo da mensagem, como o código LOINC que se refere a mensagem clínica; e
3. Os tipos de dados HL7.

Até este momento os documentos de nível 2 e 3 ainda não estão prontos, e passam por fase de escolha e votação pela comunidade HL7. De acordo com a definição CDA, os documentos de nível 2 devem definir um esquema de elementos de marcação para eventos clínicos particulares, como relatos ou notas de histórico, da condição física, da operação, de um procedimento, ou da conta de um paciente. Já os documentos de nível 3 definem um esquema de elementos de marcação para que o conteúdo clínico de especialidades como cardiologia e endocrinologia possa ser expresso formalmente. Os documentos clínicos dos três níveis devem estar de acordo com a ontologia HL7 RIM. O documento a seguir ilustra uma possível hierarquia de documentos CDA ([Dolin et al. 2005](#)).

- CDA Schema

- CDA Schema :: Nota de progresso clínico com aplicação de template de documento nível 2.
 - * CDA Schema :: Nota de progresso clínico e de sinais vitais com aplicação de templates de documentos nível 2 e 3, respectivamente.
 - CDA Schema :: Nota de progresso endócrino e de sinais vitais com aplicação de templates de documentos nível 2 e 3, respectivamente.
 - CDA Schema :: Nota de progresso clínico e de sinais vitais de UTI com aplicação de templates de documentos nível 2 e 3, respectivamente.
- CDA Schema :: Nota de progresso cardiológico com aplicação de template de documentos de nível 2.
 - * CDA Schema :: Nota de progresso cardiológico e exame cardíaco com aplicação de templates de documentos nível 2 e 3, respectivamente.
- CDA Schema :: Nota de progresso endócrino com aplicação de template de documento nível 2.
 - * CDA Schema :: Nota de progresso endócrino e de sinais vitais com aplicação de templates de documentos nível 2 e 3, respectivamente.

Outras Especificações HL7

A especificação HL7 *Vocabulary* fornece uma organização e um repositório de vocabulários codificados que possam ser utilizados nos documentos clínicos CDA e em associação ao RIM. Assim, esta especificação pretende suportar o compartilhamento de dados clínicos de forma bem definida, e sem ambiguidade no significado dos dados sendo transferidos ([Boyer et al. 2007](#)). Alguns dos sistemas de vocabulário ou de terminologia utilizados pelo HL7 *Vocabulary* são: CPT, ICD-9 CM, ICD-10, UMLS Metathesaurus, LOINC, Read codes, NDC codes, SNOMED International.

A especificação HL7 *Templates* é mais um dos esforços da organização HL7 de modo a criar um *framework* que suporte a troca de dados entre aplicações clínicas através de documentos estruturados. Esta especificação define uma padrão para a criação de instâncias de documentos clínicos no contexto da arquitetura CDA. Sendo

assim, os modelos de documentos clínicos criados para situações diversas das atividades médicas devem ser projetados segundo este padrão (Greive et al. 2007).

Pode-se resumir então que o padrão HL7 oferece um *framework* para suportar a troca de informações clínicas entre aplicações em saúde através das seguintes especificações: a ontologia RIM (*Reference Information Model*), a arquitetura de documentos clínicos CDA *Clinical Document Architecture*, o padrão para construção de modelos de documentos clínicos *Templates* e o conjunto de vocabulários de termos clínicos *Vocabulary*. Apesar desta solução conjunta mostrar-se completa e abrangente para os domínios clínicos e administrativos, dois aspectos devem ser salientados no contexto deste trabalho:

1. a versão final dos documentos para situação clínica como relatos ou notas de histórico, da condição física, da operação, de um procedimento, ou da conta de um paciente ainda encontra-se em fase de votação e aceitação (*ballot*) pela comunidade HL7⁷; e
2. não é proposto um documento clínico que suporte especificamente a troca de informações visando apoiar a decisão diagnóstica em ambiente colaborativo.

Em função disto, o padrão HL7 não pôde ser utilizado neste trabalho para cumprir os propósitos e objetivos estabelecidos neste.

4.4 Web Semântica

Liderada pela W3C, a Web Semântica tem como objetivo fornecer um conjunto de padrões e especificações que pode ser utilizado para se adicionar informação semântica aos dados na Web através de ontologias. Segundo Berners-Lee e co-autores em (Berners-Lee et al. 2001), somente assim as informações na Web poderão ser automaticamente processadas por computadores, e não somente através de intervenção humana.

Um grupo de trabalho na área de Web Semântica específico da área de informática em saúde nomeado *Semantic Web Health Care and Life Sciences* (Miller et al. 2006a) foi criado com o objetivo de criar e aperfeiçoar a aplicação de tecnologias da Web Semântica na indústria da informática médica. O escopo deste grupo inclui vocabulários médicos, ontologias, guias de conduta, e integração e comunicação de dados clínicos.

Dentre as especificações estudadas pela comunidade da Web Semântica para representar e manipular ontologias, RDF e OWL destacam-se por serem recomendações da própria W3C e também por serem frequentemente exploradas por aplicações na Web que procuram a interoperabilidade semântica. Estes dois padrões são apresentados a seguir.

4.4.1 RDF

A especificação RDF tem como propósito oferecer uma linguagem de marcação XML de modo que recursos Web possam ser descritos. Um recurso na Web pode

⁷<http://www.hl7.org/v3ballot/html/welcome/environment/index.htm>

ser qualquer coisa que tenha uma URI associada, como por exemplo, um documento HTML, ou um elemento de marcação de um documento XML, ou um arquivo de áudio e vídeo, ou ainda uma impressora ou servidor de arquivos.

Para representar um recurso, RDF define um modelo e uma sintaxe que é basicamente formada por uma tripla ‘recurso – propriedade – valor da propriedade’. Essa tripla pode ser vista como uma instrução ‘sujeito – predicado – objeto’. Um recurso é descrito utilizando o elemento de marcação *Description*. Ainda, *Namespace*s são utilizados para identificar e diferenciar quais elementos de marcação são da linguagem RDF e quais são elementos da própria aplicação.

O documento RDF a seguir apresenta a definição do recurso ou sujeito *CholecistitisDiagnosis*, que possui oito propriedades ou predicados com seus respectivos valores de propriedade ou objeto: *umlsAui*(A0038480), *umlsSui*(S0024825), *umlsLui*(L0008325), *umlsCui*(C0008325), *umlsConcept*(*Cholecystitis*), *umlsSemanticType*(*Disease or Syndrome*), *umlsDefinition* (*Inflammation of the gallbladder, generally caused by impairment of bile flow, gallstones in the biliary tract, infections, or other diseases*) e *umlsSourceSystemName*(*Mesh*). Todas as propriedades são qualificadas com o *namespace* *ddsont*. A Figura 4.14 apresenta graficamente o documento RDF do recurso *Cholecistitis*. As elipses representam os sujeitos, as arestas indicam os predicados e os retângulos são os objetos.

```
1 <rdf:RDF
2   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3   xmlns:ddsont="http://143.107.220.180/DDSont">
4   <rdf:Description rdf:about="http://143.107.220.180/DDSont#CholecistitisDiagnosis">
5     <ddsont:umlsAui>A0038480</ddsont:umlsAui>
6     <ddsont:umlsSui>S0024825</ddsont:umlsSui>
7     <ddsont:umlsLui>L0008325</ddsont:umlsLui>
8     <ddsont:umlsCui>C0008325</ddsont:umlsCui>
9     <ddsont:umlsConcept>Cholecystitis</ddsont:umlsConcept>
10    <ddsont:umlsSemanticType>Disease or Syndrome</ddsont:umlsSemanticType>
11    <ddsont:umlsDefinition>
12      Inflammation of the GALLBLADDER; generally caused by impairment of BILE
13      flow, GALLSTONES in the BILIARY TRACT, infections, or other diseases
14    </ddsont:umlsDefinition>
15    <ddsont:umlsSourceSystemName>Mesh</ddsont:umlsSourceSystemName>
16  </rdf:Description>
17 </rdf:RDF>
```

As propriedades utilizadas no documento em questão formam o recurso DDSontADT da ontologia DDSont, descrita no Capítulo 5. Estas propriedades são consideradas neste projeto como obrigatórias para trocar ou compartilhar um termo clínico entre aplicações clínicas terminologicamente e ontologicamente heterogêneas e que desejam trocar informações para apoiar a decisão diagnóstica compartilhada. Os valores destas propriedades são informações contidas na UMLS Metathesaurus sobre o diagnóstico *Cholecistitis* da fonte de vocabulário Mesh. Ainda, o documento apresentado é parte dos documentos clínicos construídos em conformidade com a ontologia DDSont que serão apresentados no Capítulo 5.

O documento RDF em sua totalidade que instância a ontologia DDSont com exemplos utilizados no estudo de caso do projeto pode ser encontrada na web⁸.

4.4.1.1 RDFS

RDFS oferece uma especificação para descrição de uma ontologia e seus recursos particulares, como uma classe, suas propriedades, valores, suas sub-classes, e a

⁸<http://143.107.220.180/DDSont/api/RDF/DDSontInstance.rdf>

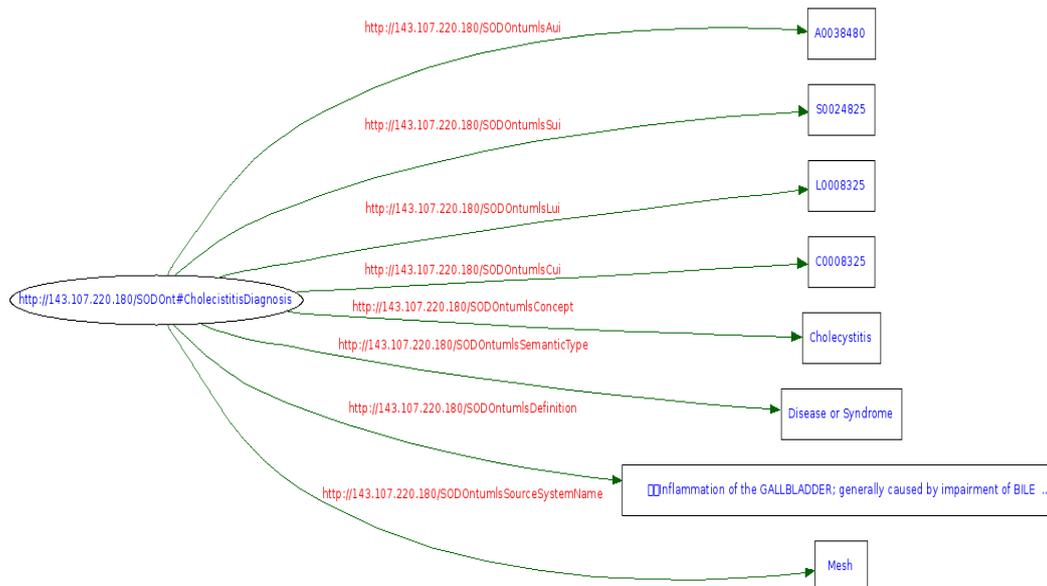


Figura 4.14: Documento RDF na forma gráfica.

relação dela com outras classes. Estas classes em RDFS se parecem com as classes do paradigma da orientação a objetos. Instâncias de classes são recursos RDF.

Com RDFS pode-se especificar e formalizar uma ontologia na forma de um documento estruturado XML. A Figura 4.15 apresenta parte da ontologia DDSOnt representada graficamente com a linguagem de modelagem UML. Nesta figura, um sintoma (*Signal*) deve ter uma descrição (*hasSignalDescription*) e um valor ou conteúdo (*hasSignalValue*), ambos do tipo DDSOntADT. DDSOntADT representa um tipo de dado contendo as informações da UMLS Metathesaurus que identificam qualquer dado clínico para apoiar a decisão diagnóstica. Uma avaliação clínica (*ClinicalEvaluation*) deve ter uma descrição de sua patologia (*hasPathologyDescription*), uma possível decisão diagnóstica *hasDiagnosisDecision* opcional e vários sinais *hasSignals*.

A ontologia apresentada pode então ser representada como um documento RDFS. Basicamente, pode-se dividir um documento RDFS em dois blocos: o bloco para definição das classes, e o bloco para definição das propriedades das classes. Uma propriedade pode ser classificada como de tipo de dados quando seu tipo é primitivo, ou então classificada como objeto, quando seu tipo é uma classe. O documento RDFS a seguir ilustra a representação documental da ontologia mostrada na Figura 4.15. Na linha 3 é definido o *namespace* rdfs utilizado em elementos de marcação específicos da linguagem RDFS. Nas linhas de 7 a 10 é definida a classe DDSOntADT através do elemento *Description*. O que indica que esse recurso é uma classe é o conteúdo do elemento *type*. As outras classes podem ser definidas da mesma forma. Das linhas 14 a 21 é definida a propriedade *umlsAui*, indicado pelo conteúdo do elemento *type*. Ainda, pode-se identificar que essa propriedade pertence a classe DDSOntADT através do elemento *domain* e de que essa propriedade é de tipo de dados, pois o conteúdo do elemento *range* é *string*, um tipo primitivo. As outras propriedades da classe DDSOntADT podem ser também assim definidas. Das linhas 24 a 31 é definida a propriedade *hasSignalDefinition*, indicado pelo conteúdo do elemento *type*. Ainda, pode-se identificar que essa propriedade pertence a classe *Signal* através do elemento *domain* e de que essa propriedade é de objetos, pois o conteúdo do elemento *range* é a classe *DDSOntADT*. A propriedade *hasSignalValue* pode também ser definida da mesma forma. Finalmente, das linhas 34 a 41 é definida a propriedade *hasDiagno-*

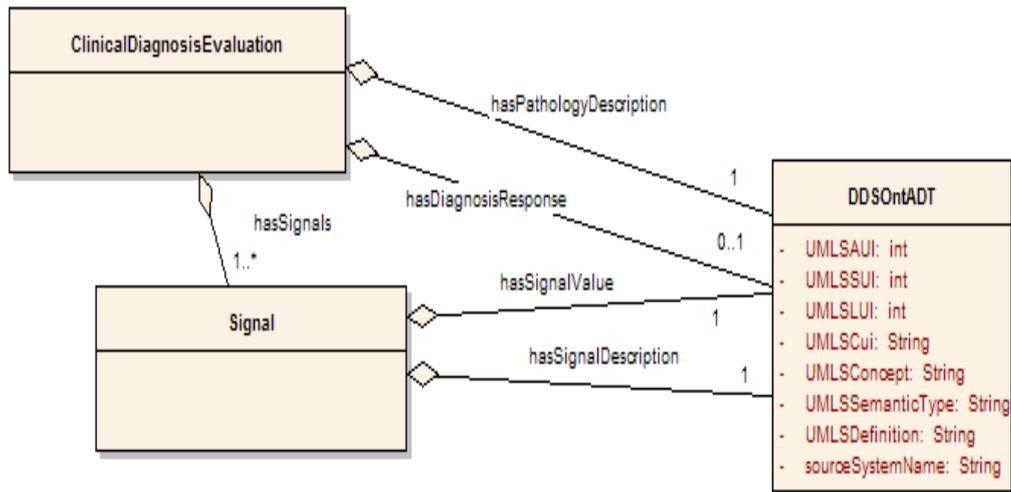


Figura 4.15: Parte da ontologia DDSOnt.

sisDecision, indicado pelo conteúdo do elemento *type*. Ainda, pode-se identificar que essa propriedade pertence a classe *ClinicalEvaluation* através do elemento *domain* e de que essa propriedade é de objetos, pois o conteúdo do elemento *range* é a classe *DDSOntADT*. As outras propriedades da classe *ClinicalEvaluation* podem ser também assim definidas.

```

1 <rdf:RDF xml:lang="en"
2   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
4   xmlns:ddsont="http://143.107.220.180/DDSont#" >
5   <!-- classes da DDSOnt -->
6   <rdf:Description rdf:about="DDSOntADT">
7     <rdf:type
8       rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class" />
9   </rdf:Description>
10  ...
11  <!-- Propriedades de tipos de dados da classe DDSOntADT -->
12  <rdf:Description rdf:about="umlsAui">
13    <rdf:type
14      rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property" />
15    <rdfs:domain rdf:resource="#DDSOntADT" />
16    <rdfs:range
17      rdf:resource="http://www.w3.org/2000/03/example/classes#string" />
18  </rdf:Description>
19  ...
20  <!-- Propriedades de objetos da classe Signal -->
21  <rdf:Description rdf:about="hasSignalDescription">
22    <rdf:type
23      rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property" />
24    <rdfs:domain rdf:resource="#Signal" />
25    <rdfs:range rdf:resource="DDSOntADT" />
26  </rdf:Description>
27  <!-- Propriedades de objetos da classe ClinicalEvaluation -->
28  <rdf:Description rdf:about="hasDiagnosisDecision">
29    <rdf:type
30      rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property" />
31    <rdfs:domain rdf:resource="#ClinicalEvaluation" />
32    <rdfs:range rdf:resource="DDSOntADT" />
33  </rdf:Description>
34  ...
35 </rdf:RDF>
  
```

O documento RDFS que formaliza a ontologia DDSOnt em sua totalidade pode ser

encontrada na web⁹.

A importância de se ter uma ontologia descrita formalmente na forma de um documento estruturado XML como o apresentado é justificada pela possibilidade de se verificar se instâncias de documentos RDF, como apresentado na seção 4.4.1, estão em conformidade com a ontologia. Ainda, é possível se criar agentes de software que façam a manipulação automática nestes documentos XML.

A linguagem RDFS para descrever esquemas e ontologias mostra-se limitada em alguns aspectos, como restrições de cardinalidade nas propriedades, especificação de uma propriedade transitiva, definição de classes diferentes que representam a mesma classe, especificação de instâncias diferentes que representam o mesmo indivíduo, e a possibilidade de se definir novas classes a partir da combinação de outras. Estas limitações motivou a criação da linguagem OWL.

No contexto deste trabalho, a não possibilidade de se especificar cardinalidades de propriedades foi o que limitou o uso de utilização da RDFS na representação da ontologia DDSOnt. A linguagem OWL tem também como objetivo descrever esquemas e ontologias sem as limitações encontradas na linguagem RDFS. Ela é descrita na seção a seguir.

4.4.2 OWL

OWL (Miller and Hendler 2006) é caracterizada por oferecer um vocabulário adicional para a formalização semântica de ontologias e portanto possui maior poder de interpretabilidade em relação a linguagens como RDF e RDFS apresentadas anteriormente (McGuinness and Harmelen 2004). A OWL é composta de 3 sub-linguagens: OWL *Light*, OWL DL(*Description Logic*) e OWL *Full*.

A OWL *Light* suporta a descrição de hierarquias de classificação de classes e restrições simples, como as cardinalidades mínimas e máximas que somente podem conter valor 0 ou 1. A OWL DL suporta a expressividade máxima de descrições lógicas assegurando a completude e decidibilidade computacional. Um exemplo é a declaração explícita de classes cujas instâncias da mesma não podem ser instâncias de outras classes simultaneamente. E finalmente a OWL *Full* suporta expressividade máxima e liberdade sintática RDF, mas sem nenhuma garantia computacional. Como exemplo, pode-se atribuir valores diferentes de apenas 0 e 1 para cardinalidades máxima e mínima. Sendo assim, foi necessário neste projeto a utilização da OWL *Full* para formalizar a ontologia DDSOnt.

Uma classe é definida na linguagem OWL com o elemento de marcação *Class*. Para garantir que uma instância da classe sendo definida não pode ser simultaneamente instância de uma outra classe, utiliza-se o elemento *disjointWith*. De acordo com o documento a seguir, uma instância da classe DDSOntADT não pode ser instância simultaneamente das classes *ClinicalEvaluation* e *Signal*.

```
1 <owl:Class rdf:about="#DDSOntADT">
2   <owl:disjointWith rdf:resource="#ClinicalEvaluation"/>
3   <owl:disjointWith rdf:resource="#Signal"/>
4 </owl:Class>
```

Caso a classe que está sendo definida possua restrições de cardinalidade em uma associação com uma outra classe, deve-se expandí-la com o reuso do elemento *subClassOf* da linguagem RDFS de modo a incluir as noções de cardinalidade. Pode-se

⁹<http://143.107.220.180/DDSOnt/api/RDF/DDSOnt.rdf>

definir cardinalidades exata, mínima e máxima. O documento a seguir apresenta a definição da classe *Signal*. Esta classe possui uma restrição de cardinalidade exata em sua propriedade *hasSignalDescription* representada pelos elementos OWL *cardinality*, *onProperty* e *ObjectProperty*. Isso significa que um sintoma deve ter exatamente uma descrição. Da mesma forma que foi definida a restrição para a propriedade *hasSignalDescription*, pode-se definir também a restrição para a propriedade *hasSignalValue* também da classe *Signal*.

```

1 <owl:Class rdf:about="#Signal">
2     ...
3     <rdfs:subClassOf>
4         <owl:Restriction>
5             <owl:cardinality
6                 rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
7                 1
8             </owl:cardinality>
9             <owl:onProperty>
10                <owl:ObjectProperty rdf:ID="hasSignalDescription"/>
11            </owl:onProperty>
12        </owl:Restriction>
13    </rdfs:subClassOf>
14 </owl:Class>

```

O documento a seguir apresenta a definição da classe *ClinicalEvaluation*. Esta classe possui restrições de cardinalidade mínima, exata e máxima. Uma avaliação clínica deve ter ao menos 1 e no máximo 40 sinais, de acordo com o estudo de caso de dor abdominal aguda desenvolvido no trabalho, e exatamente uma descrição patológica, uma ou zero possível decisão diagnóstica. A cardinalidade mínima é representada pelo elemento *minCardinality* e a cardinalidade máxima pelo elemento *maxCardinality*. Da linha 3 até a linha 11 é definida a restrição para a propriedade *hasDiagnosisDecision* com cardinalidade mínima de 0 e cardinalidade máxima de 1. Da linha 12 a 20 é definida a restrição da propriedade *hasSignals* com cardinalidade mínima de 1 e máxima de 40.

```

1 <owl:Class rdf:about="#ClinicalEvaluation">
2     ...
3     <rdfs:subClassOf>
4         <owl:Restriction>
5             <owl:onProperty>
6                 <owl:ObjectProperty rdf:ID="hasDiagnosisDecision"/>
7             </owl:onProperty>
8             <owl:minCardinality
9                 rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
10                0
11            </owl:minCardinality>
12            <owl:maxCardinality
13                rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
14                1
15            </owl:maxCardinality>
16        </owl:Restriction>
17    </rdfs:subClassOf>
18    <rdfs:subClassOf>
19        <owl:Restriction>
20            <owl:onProperty>
21                <owl:ObjectProperty rdf:ID="hasSignals"/>
22            </owl:onProperty>
23            <owl:maxCardinality
24                rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
25                40
26            </owl:maxCardinality>
27            <owl:minCardinality
28                rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
29                1
30            </owl:minCardinality>
31        </owl:Restriction>
32    </rdfs:subClassOf>
33 </owl:Class>

```

Finalmente, para a definição das propriedades, 2 elementos de marcação OWL

são utilizados: *DatatypeProperty* para criação de propriedades de tipo primitivo, e *ObjectProperty* para definição de propriedades do tipo objeto. Para apresentar a qual classe a propriedade está associada e qual o seu conteúdo, são utilizados os elementos de marcação da linguagem RDFS *range* e *domain*, respectivamente. O documento a seguir apresenta exemplos de definição de propriedades. Da linha 1 a 4 é definido a propriedade *umlsAui* da classe *DDSOntADT* e de tipo *string*. E da linha 5 a 8 é criado a propriedade *hasDiagnosisDecision* da classe *ClinicalEvaluation* e do tipo *DDSOntADT*.

```
1 <owl:DatatypeProperty rdf:ID="umlsAui">
2   <rdfs:range
3     rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
4   <rdfs:domain
5     rdf:resource="#DDSOntADT"/>
6 </owl:DatatypeProperty>
7 <owl:ObjectProperty rdf:about="#hasDiagnosisDecision">
8   <rdfs:domain rdf:resource="#ClinicalEvaluation"/>
9   <rdfs:range rdf:resource="#DDSOntADT"/>
10 </owl:ObjectProperty>
```

O documento OWL que formaliza a ontologia DDSOnt em sua totalidade pode ser encontrada em [10](#).

4.5 Considerações Finais

Promover a compatibilidade ontológica durante a troca e compartilhamento de dados clínicos é uma difícil tarefa em função principalmente dos sistemas ontológicos representarem um mesmo tipo semântico de diferentes formas, e ainda destes sistemas possuírem diferentes níveis de granularidade na rede semântica.

A partir do estudo realizado com os sistemas ontológicos UMLS SN, Galen CRM e HL7 RIM, verificou-se que enquanto o primeiro define os tipos semânticos a partir de conceitos provenientes de mais de 100 sistemas de terminologia, os outros dois criam os tipos semânticos a partir das experiências de seus autores. Apesar das diferenças de granularidade serem em geral maior na UMLS SN, acredita-se que a compatibilidade ontológica seja mais frequentemente conseguida no cenário deste projeto com o uso da UMLS SN, visto que a compatibilidade terminológica é promovida com a UMLS MetaThesaurus. É importante observar que a compatibilidade ontológica com a UMLS SN não será conseguida quando as aplicações clínicas fizerem uso de sistemas de terminologia não contemplados pelo UMLS MetaThesaurus.

Neste contexto, este projeto cria a ontologia DDSOnt que estende a UMLS SN através de seus tipos semânticos Entidade e Evento com objetivo de não redefinir tipos semânticos já definidos pela SN. A DDSOnt tem como objetivo definir um novo conjunto de tipos semânticos necessários para facilitar o suporte semântico durante a criação de uma estrutura que auxilie no compartilhamento e na recuperação de bases de conhecimento utilizadas por médicos no apoio à decisão diagnóstica em ambiente colaborativo. Estes tipos semânticos não foram encontrados em nenhuma das ontologias apresentadas neste capítulo, e nem em outras ontologias médicas presentes na literatura, como MedO [11](#), Foundation Model of Anatomy [12](#), Gene Ontology (Ashburner et al. 2000), UNSPSC ontology [13](#), dentre outras.

¹⁰<http://143.107.220.180/DDSOnt/api/OWL/DDSOnt.owl>

¹¹<http://www.ifomis.uni-saarland.de/>

¹²<http://sig.biostr.washington.edu/projects/fm/>

¹³<http://www.unspsc.org/>

O uso das linguagens OWL e RDF para formalizar a DDSOnt mostram-se importantes e necessárias em função do ambiente colaborativo que apóia a decisão diagnóstica ser suportado pela Web. Com estas linguagens é possível fazer com que os dados clínicos em conformidade à DDSOnt sejam processadas também por máquinas, e não somente pela intervenção humana.

O próximo Capítulo apresenta a ontologia proposta DDSOnt, como ela foi desenvolvida com a metodologia 101 ([Noy and McGuinness 2001](#)), e quais são seus tipos semânticos, suas propriedades, associações e restrições entre estes. Ainda, são apresentados os softwares relacionados a DDSOnt, JDDSOnt e DDSOntWs, e como estes podem apoiar o uso da DDSOnt em aplicações médicas na Web que precisam trocar informações para apoiar a decisão diagnóstica.

Ontologia DDSOnt e Softwares Relacionados

5.1 Considerações Iniciais

A partir dos estudos realizados nas ontologias clínicas existentes na literatura, verificou-se a inexistência de um conjunto de tipos semânticos que possa facilitar a criação de uma estrutura que auxilie no compartilhamento e na recuperação de bases de conhecimento utilizadas por médicos no apoio à decisão diagnóstica em ambiente colaborativo. Sendo assim, este trabalho se propõe a criar uma ontologia que defina esses tipos semânticos com a preocupação de reutilizar uma ontologia clínica existente de modo a promover a compatibilidade ontológica.

Este Capítulo tem como objetivo apresentar a ontologia DDSOnt (*Diagnosis Decision Support Ontology*), bem como expor o seu processo de desenvolvimento com a metodologia 101 (Noy and McGuinness 2001). Ainda, é explicado como foram construídos os softwares JDDSOnt e DDSOntWS relacionados a ontologia, e de que forma estes devem ser utilizados por aplicações médicas que precisam fazer uso da DDSOnt de modo a serem apoiadas na decisão diagnóstica de forma colaborativa.

5.2 Definição da Ontologia DDSOnt

5.2.1 Escolha da Metodologia

Várias são as metodologias que tem a função de guiar a construção e desenvolvimento de ontologias. São exemplos na literatura a TOVE (Gruninger and Fox 2002), Ushold (Osterwalder and Pigneur 2002), Methontology (Fernández-López et al. 1997), Onto Clean (Guarino and Welty 2002) e 101 (Noy and McGuinness 2001). Estas metodologias não são especificações ou recomendações, mas sim modelos empíricos

baseados na experiência dos projetos onde foram criadas e defendidas pelos seus criadores.

Dentre as metodologias pesquisadas, a metodologia 101 mostrou-se a mais objetiva, prática, didática e capaz de atender e satisfazer as características da ontologia DDSOnt. Sendo assim, a 101 foi utilizada para defini-la. Esta metodologia foi criada por Noy e McGuinness na Universidade de Stanford, nos EUA, e os resultados do trabalho foram publicados em (Noy and McGuinness 2001). Segundo os autores, a 101 tem o propósito de ser um guia inicial para ajudar os desenvolvedores iniciantes de ontologia a desenvolver a sua própria.

A metodologia 101 basicamente define 7 passos que devem ser seguidos para a construção de uma ontologia. Como mostra Noy e McGuinness em (Noy and McGuinness 2001), estes passos são:

- Passo 1: O primeiro passo aborda 4 aspectos. No primeiro deles, deve-se definir o domínio e o escopo coberto pela ontologia. No segundo aspecto, deve-se definir o propósito de sua utilização. No terceiro aspecto, deve-se definir quais os tipos de perguntas esta deve responder. E finalmente no quarto aspecto, deve-se discutir quem irá utilizar e manter a ontologia;
- Passo 2: O objetivo do segundo passo é analisar a possibilidade de reutilizar outras ontologias existentes na literatura;
- Passo 3: O terceiro passo consiste em enumerar os termos importantes da ontologia;
- Passo 4: O objetivo do quarto passo é definir os tipos semânticos a partir dos termos criados no passo anterior, e determinar o relacionamento entre estes tipos semânticos;
- Passo 5: O quinto passo consiste em definir as propriedades de cada um dos tipos semânticos definidos no passo anterior;
- Passo 6: O objetivo de sexto passo é apresentar as restrições de cada uma dessas propriedades definidas no passo anterior;
- Passo 7: O sétimo e último passo consiste em definir e criar instâncias ou exemplos dos tipos semânticos da ontologia.

Os resultados da execução dos 7 passos da 101 para a construção da DDSOnt são apresentados a seguir.

5.2.2 Execução do Passo 1

Define-se que a DDSOnt possui o domínio e o escopo de sistemas eletrônicos ou prontuários eletrônicos em saúde que precisam ser apoiados durante a tomada de decisão diagnóstica em ambiente computacional colaborativo e orientado a serviços. Afirma-se que o propósito da DDSOnt é definir um novo conjunto de tipos semânticos necessários para a criação de uma estrutura que suporte a manipulação de bases de conhecimento de modo que possam ser utilizadas por médicos no apoio à decisão clínica em ambiente colaborativo.

A ontologia DDSOnt está habilitada a responder os seguintes tipos de perguntas:

- Um termo clínico qualquer da área médica como sinal clínico, doença, conceito qualitativo, substância do corpo, sinal ou resultado laboratorial, funções do organismo, descoberta clínica, funções patológicas, funções fisiológicas, e outras mais, é uma instância individual do tipo semântico Evento ou Entidade da ontologia UMLS SN? Esta pergunta torna-se importante de ser respondida para garantir a compatibilidade ontológica no contexto deste trabalho;
- Um termo clínico como os citados anteriormente carregam consigo durante uma troca de dados, informações relacionadas ao identificador único UMLS do átomo, o identificador único UMLS da *String*, o identificador único UMLS do termo léxico, o identificador único UMLS do conceito, o conceito UMLS, o tipo semântico UMLS, a definição UMLS, e o nome do sistema de terminologia de origem? Como descrito no Capítulo 3, estas informações da UMLS *MetaThesaurus* são consideradas essenciais para se promover a compatibilidade terminológica. E portanto, devem estar presentes na transferência de qualquer informação entre as aplicações que se utilizam da proposta DDSOnt;
- Uma avaliação clínica contém uma descrição da síndrome sendo discutida? Ainda, a descrição patológica contém as informações apresentadas anteriormente da UMLS *MetaThesaurus*?
- Uma avaliação clínica contém uma ou nenhuma resposta diagnóstica? E ainda, a resposta diagnóstica contém as informações da UMLS *MetaThesaurus* descritas anteriormente?
- Uma avaliação clínica contém um ou mais sinais clínicos e/ou laboratoriais?. Ademais, o sinal clínico ou laboratorial contém as informações apresentadas anteriormente da UMLS *MetaThesaurus*?
- Um sinal clínico ou laboratorial possui um valor ou resultado? Adicionalmente, este valor ou resultado contém informações da UMLS *MetaThesaurus*?
- Finalmente, um sinal contém uma auto-descrição, e ainda esta auto-descrição contém associado informações da UMLS *MetaThesaurus*?

O quarto e último aspecto do Passo 1 é discutir quem irá utilizar e manter a ontologia. Os potenciais usuários da DDSOnt são juntas médicas especializadas e médicos que desejam compartilhar suas experiências com a comunidade para auxiliar outros médicos durante a decisão diagnóstica. Por outro lado, outros tipos de usuários da DDSOnt são os médicos que precisam ser apoiados durante a tomada de decisão clínica em ambiente colaborativo.

O laboratório ImagCom, do Departamento de Física e Matemática – Universidade de São Paulo – Ribeirão Preto está mantendo a DDSOnt.

5.2.3 Execução do Passo 2

Como já mencionado nos Capítulos 3 e 4, a ontologia DDSOnt reutiliza a ontologia UMLS SN em função de dois motivos: o primeiro é que a UMLS SN define seus tipos semânticos a partir dos conceitos provenientes de mais de 100 sistemas de terminologia, enquanto que outras ontologias como a Galen CRM e HL7 RIM definem seus tipos semânticos a partir da experiência de seus próprios autores; e o segundo motivo é que a proposta deste trabalho utiliza a UMLS *MetaThesaurus* para promover

Tabela 5.1: Primeiro conjunto de termos presentes na DDSOnt

| Termo | Descrição Textual |
|------------------------------------|--|
| <i>Severities</i> | Conceito qualitativo utilizado para medição de algo não desejado, por exemplo, a intensidade da dor |
| <i>Moderate(severity modifier)</i> | Conceito qualitativo que indica que a condição pode resultar em consequências adversas, mas é improvável um tratamento ou dano |
| <i>Vomitus</i> | Substância do corpo que foi expelido do estômago pela boca |
| <i>Vomit:bilious</i> | Resultado de laboratório ou de teste indicando substância amarelo-esverdeada secretada pelo fígado pelos vertebrados |
| <i>Diverticulitis</i> | Doença causada por um inflamação de um divertículo que pode apresentar perfuração e abscesso |
| <i>Anorexia</i> | Sinal relacionado a perda de apetite acompanhado de aversão a comida e a falta de habilidade para comer |
| <i>Defecation</i> | Função Orgânica que representa o processo normal de eliminação do material fecal do reto |
| <i>Dyspareunia</i> | Descoberta associada a dor ou dificuldade durante o coito |
| <i>Right upper quadrant pain</i> | Sinal relacionado a uma dor no quadrante superior direito da região abdominal |
| <i>Icterus</i> | Função patológica caracterizada pela coloração amarela dos tecidos e das secreções orgânicas |

a compatibilidade terminológica, o que naturalmente provocou a aproximação com a UMLS.

A DDSOnt possui um tipo semântico denominado DDSOntADT (DDSOnt *Abstract Data Type*), descrito mais adiante neste capítulo. Este tipo semântico representa qualquer termo clínico da área médica como sinal clínico, doença, conceito qualitativo, substância do corpo, sinal ou resultado laboratorial, funções do organismo, descoberta clínica, funções patológicas, funções fisiológicas, e outras mais. Portanto, as propriedades deste tipo semântico são exatamente as informações da UMLS *MetaThesaurus* considerados neste trabalho como essenciais para a promoção da compatibilidade terminológica. Novamente, estas informações são: o identificador único UMLS do átomo, o identificador único UMLS da *String*, o identificador único UMLS do termo léxico, o identificador único UMLS do conceito, o conceito UMLS, o tipo semântico UMLS, a definição UMLS, e o nome do sistema de terminologia de origem.

O tipo semântico DDSOntADT é uma especialização dos tipos semânticos relacionados com as Entidades e Eventos da UMLS SN. Esta reutilização tem o propósito de não redefinir tipos semânticos genéricos e abstratos já definidos pela UMLS SN.

5.2.4 Execução do Passo 3

O primeiro conjunto de termos identificados para a criação da DDSOnt é formado por qualquer termo clínico na área médica, como conceito qualitativo, sinal, doença, substância do corpo, resultado laboratorial ou teste, funções do organismo, descoberta, funções patológicas, funções fisiológicas, e outras mais. A Tabela 5.1 mostra exemplos de termos do primeiro conjunto presentes na DDSOnt com suas respectivas descrições textuais.

Os termos apresentados na Tabela 5.1 estão todos relacionados a um conceito da UMLS *MetaThesaurus*, e conseqüentemente a algum tipo semântico da UMLS SN.

Tabela 5.2: Termos do segundo conjunto presentes na DDSOnt

| Termo | Descrição Textual |
|--|--|
| (<i>severities; moderate(severity modifier)</i>) | A medição da intensidade da dor está moderada |
| (<i>Vomitus; Vomit:bilious</i>) | A substância do corpo vômito apresenta-se como um líquido viscoso amarelo-esverdeado |
| (<i>Vomitus; Vomit:feculant</i>) | A substância do corpo vômito apresenta-se fétido, com odor desagradável |
| (<i>Defecation;Obstipation</i>) | O processo de eliminar o material fecal está constipado |
| (<i>Defecation;Diarrhea</i>) | O processo de eliminar o material fecal está gerando fluidos fecais |
| (<i>Relieved by;Antiacids</i>) | O fator de alívio da dor é com o uso de anti-ácido |
| (<i>Relieved by; food</i>) | O fator de alívio da dor é com a ingestão de substâncias que nutrem |
| (<i>Auscultation; Hypokinesia</i>) | Auscultação mostra movimento muscular do corpo devagar ou diminuído |
| (<i>Auscultation; Hyperactive behavior</i>) | Auscultação mostra movimento muscular do corpo rápido ou aumentado |
| (<i>Dyspareunia; Yes - Yes/no indicator</i>) | Descoberta associada a dor ou dificuldade durante o coito com resultado positivo |

Os tipos semânticos da UMLS SN podem ser todos representados através dos tipos semânticos Entidade e Eventos, pois todos herdam ou são especializações destes.

Ainda sobre os termos apresentados, todos devem carregar consigo, durante uma troca de dados em ambiente compartilhado, as informações essenciais da UMLS *MetaThesaurus* apresentadas anteriormente para promover a compatibilidade terminológica.

O segundo conjunto de termos é formado pelos pares (auto-descrição do sinal clínico ou laboratorial; valor ou resultado do sinal clínico ou laboratorial), em que a auto-descrição e o valor do sinal são termos do primeiro conjunto de termos, apresentados na Tabela 5.1. Sendo assim, tanto a descrição quanto o valor de um sinal carregam consigo as informações essenciais da UMLS *MetaThesaurus* para promover a compatibilidade terminológica. A Tabela 5.2 mostra termos do segundo conjunto com suas respectivas descrições textuais.

O último e terceiro conjunto de termos é formado pela tripla ('descrição patológica'; 'sinais'; 'resposta diagnóstica'), onde 'descrição patológica' e 'resposta diagnóstica' são termos do primeiro conjunto, e sinais é um conjunto de termos do segundo conjunto. Sendo assim, tanto a descrição patológica como a possível decisão diagnóstica também carregam consigo as informações da UMLS *MetaThesaurus*. A Tabela 5.3 mostra termos do terceiro conjunto com suas respectivas descrições textuais ¹.

Este terceiro conjunto de termos é o que de fato compõe uma unidade de transferência para apoiar a decisão diagnóstica em ambiente colaborativo. No contexto do estudo de caso criado neste projeto, pode-se dizer sobre a tripla ('descrição patológica'; 'sinais'; 'resposta diagnóstica') que:

- A 'descrição patológica' refere-se sempre a dor abdominal aguda;
- O termo 'sinais' refere-se sempre a um dos 40 sinais clínicos ou laboratoriais relacionados a dor abdominal aguda;

¹Os termos presentes na tabela podem não condizer com a realidade clínica.

Tabela 5.3: Termos do terceiro conjunto presentes na DDSOnt

| Termo | Descrição Textual |
|---|---|
| <i>(acute abdominal pain; {(severities, moderate(severity modifier)),(vomitus, vomit:bilious)}; cholecystitis)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica colecistite de acordo com os resultados da medição da intensidade da dor e da substância do corpo vômito |
| <i>(acute abdominal pain; {(Vomitus, Vomit:bilious),(defecation, obstipation)}; appendicitis)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica apendicite de acordo com os resultados da substância do corpo vômito e do processo de eliminação fecal |
| <i>(acute abdominal pain; {(severities, moderate(severity modifier)),(defecation, obstipation)}; Peptic Ulcer)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica úlcera gástrica de acordo com os resultados da medição da intensidade da dor e do processo de eliminação fecal |
| <i>(acute abdominal pain; {(Defecation, Diarrhea),(Relieved by, Antiacids)}; Gastroenteritis)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica inflamação gástrica de acordo com os resultados do processo de eliminação fecal e fator de alívio da dor |
| <i>(acute abdominal pain; {(Defecation, Diarrhea),(Relieved by, food)}; diverticulitis)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica diverticulite de acordo com os resultados do processo de eliminação fecal e fator de alívio da dor |
| <i>(acute abdominal pain; {(Vomitus, Vomit:bilious),(Auscultation, Hypokinesia)}; pancreatitis)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica pancreatite de acordo com os resultados da substância do corpo vômito e processo de auscultação |
| <i>(acute abdominal pain; {(Auscultation, Hypokinesia),(Dyspareunia, Yes - Yes/no indicator)}; hepatitis)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica hepatite de acordo com os resultados do processo de auscultação e resultado da dificuldade durante coito |
| <i>(acute abdominal pain; {(Auscultation, Hyperactive behavior),(Dyspareunia, Yes - Yes/no indicator)}; Intestinal Obstruction)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica obstrução intestinal de acordo com os resultados do processo de auscultação e resultado da dificuldade durante coito |
| <i>(acute abdominal pain; {(Auscultation, Hyperactive behavior),(Relieved by, food)}; pelvic inflammatory disease)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica inflamação pélvica de acordo com os resultados do processo de auscultação e fator de alívio da dor |
| <i>(acute abdominal pain; {(Dyspareunia, No - yes/no indicator),(Relieved by, Antiacids)}; Thrombosis of mesenteric vein)</i> | A avaliação clínica da síndrome dor abdominal aguda tem resposta diagnóstica trombose na veia do mesentério de acordo com os resultados da dificuldade durante coito e do fator de alívio da dor |

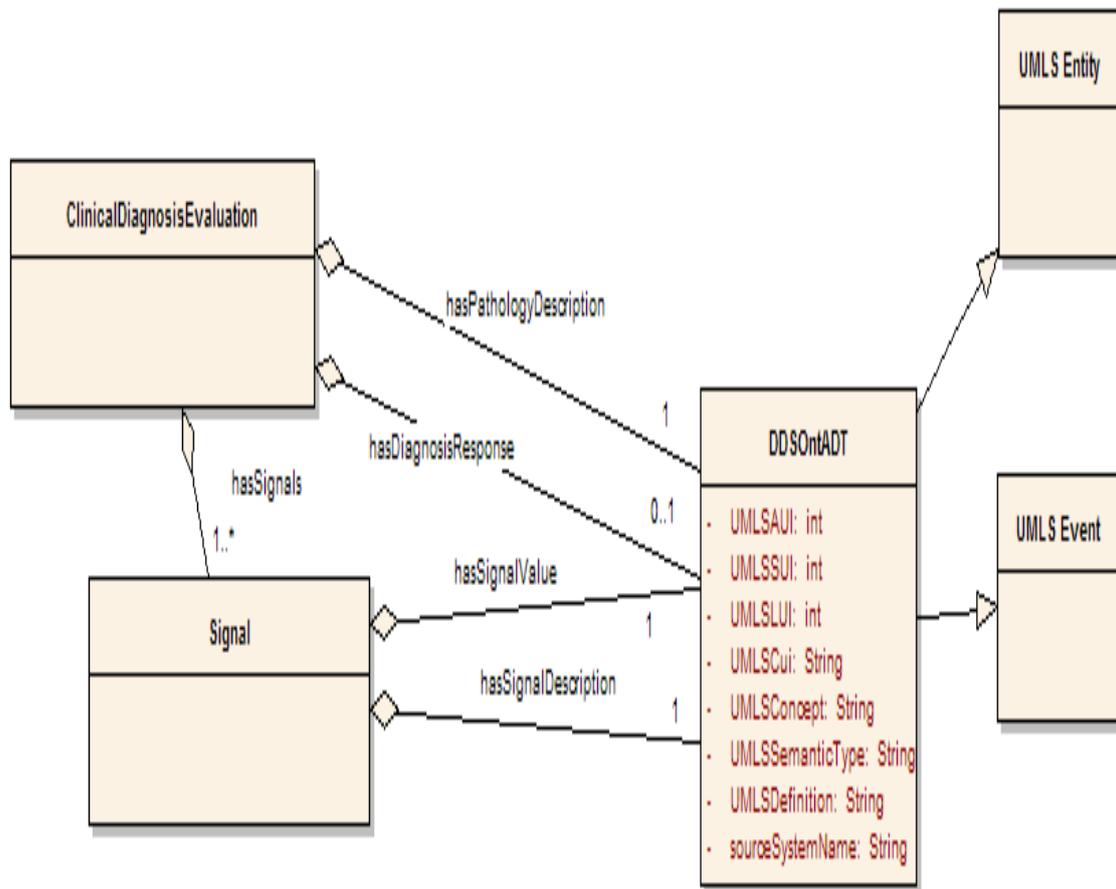


Figura 5.1: Definição dos tipos semânticos da DDSOnt

- São 14 os valores possíveis de resposta diagnóstica, e todos são do tipo semântico *Disease or Syndrome* da UMLS SN;
- No total, foram criados 237 registros semelhantes aos da Tabela 5.3. Estes 237 registros podem ser utilizadas, por exemplo, para se criar uma base de conhecimento;
- O uso de uma única tripla sem a resposta diagnóstica pode ser utilizada, por exemplo, para consultar por uma resposta diagnóstica a partir de uma base de conhecimento, como a criada no item anterior;

5.2.5 Execução do Passo 4

A ontologia DDSOnt é composta basicamente de 5 tipos semânticos relacionados conforme ilustra a Figura 5.1 e descritas como:

1. O tipo semântico *UMLS Entity* representa todos os tipos semânticos criados na UMLS SN relacionados ao tipo semântico *Entidade*, como Substância, Organismo, Estrutura Anatômica, Descoberta Clínica, entre outras, conforme descrito no Capítulo 4;

2. O tipo semântico *UMLS Event* representa todos os tipos semânticos criados na UMLS SN relacionados ao tipo semântico *Event*, como Função Fisiológica, Função Patológica, Doença, Ferimento, entre outras, conforme descrito no Capítulo 4;
3. O tipo semântico *DDSOntADT* representa os termos do primeiro conjunto definidos no Passo 3 da metodologia 101. Os termos do primeiro conjunto são provenientes da UMLS MetaThesaurus. Portanto, pode-se dizer que o tipo semântico *DDSOntADT* é um tipo especializado da *UMLS Entity* e subtipo do tipo semântico *UMLS Event*. Dessa forma, a *DDSOnt* reutiliza a ontologia UMLS SN;
4. O tipo semântico *Signal* identifica os termos do segundo conjunto definidos no Passo 3. Este tipo semântico se relaciona e é composto pelo tipo semântico *DDSOntADT* em duas situações representadas pelas associações *hasSignalDescription* e *hasSignalValue*, conforme ilustra a Figura 5.1. A primeira situação é para a auto-descrição do sinal sendo transferido, e a segunda situação é para representar o valor desse sinal;
5. O tipo semântico *ClinicalDiagnosisEvaluation* representa os termos do terceiro conjunto também definidos no Passo 3. Este tipo se relaciona e é composto pelos tipos *Signal* em uma situação representada pela associação *hasSignals*, e pelo tipo *DDSOntADT* em duas situações representadas pelas associações *hasPathologyDescription* e *hasDiagnosisResponse*, conforme ilustra a Figura 5.1. A primeira situação de relacionamento com o tipo semântico *DDSOntADT* é para a auto-descrição da síndrome sendo diagnosticada. A situação para relacionamento com o tipo *Signal* é a representação de cada sinal sendo transferido. Caso estejam sendo transferidos 'n' sinais, o relacionamento ocorre nestes 'n' momentos. Finalmente, a segunda situação de relacionamento com o tipo *DDSOntADT* é para a representação da resposta diagnóstica. Este relacionamento pode não ocorrer quando se está consultando por uma resposta diagnóstica, porém precisa ocorrer quando se está criando uma base de conhecimento a partir de um conjunto de tipos *ClinicalDiagnosisEvaluation*.

5.2.6 Execução dos Passos 5 e 6

Os objetivos dos passos 5 e 6 da metodologia 101 são, respectivamente, definir as propriedades de cada um dos tipos semânticos e, apresentar as restrições de cada uma dessas propriedades. Como os tipos UMLS *Entity* e UMLS *Event* são reutilizadas da UMLS SN, não cabe aos autores da *DDSOnt* definir suas propriedades, já que estas também serão reutilizadas. Sendo assim, as propriedades dos tipos semânticos *DDSOntADT*, *Signal* e *ClinicalEvaluation*, juntamente com seus tipos e descrições, são listadas a seguir.

- *DDSOntADT*
 - *UMLSAui* (literal): identificador único de um átomo na UMLS *MetaThesaurus*;
 - *UMLSSui* (literal): identificador único de um *string* na UMLS *MetaThesaurus*;
 - *UMLSLui* (literal): identificador único de um termo léxico na UMLS *MetaThesaurus*;
 - *UMLSCui* (literal): identificador único de um conceito na UMLS *MetaThesaurus*;

- UMLSConcept (literal): conceito de um termo clínico na UMLS *MetaThesaurus*;
- UMLSSemanticType (literal): tipo semântico de um termo clínico na UMLS *MetaThesaurus*;
- UMLSDefinition (literal): definição de um termo clínico na UMLS *MetaThesaurus*;
- SourceSystemName (literal): nome do sistema de terminologia de origem do termo clínico;
- Signal
 - *hasSignalDescription* (UMLSADT): auto-descrição de um sinal contendo as propriedades do tipo semântico *DDSOntADT*;
 - *hasSignalValue* (UMLSADT): valor de um sinal contendo as propriedades do tipo semântico *DDSOntADT*;
- ClinicalDiagnosisEvaluation
 - *hasPathologyDescription* (DDSOntDT): auto-descrição de uma síndrome contendo as propriedades do tipo semântico *DDSOntADT* ;
 - *hasSignals* (Signal): sinais de uma síndrome contendo as propriedades do tipo semântico *Signal* ;
 - *hasDiagnosisResponse* (DDSOntADT): possível resposta diagnóstica de uma síndrome contendo as propriedades do tipo semântico *DDSOntADT* ;

5.2.7 Execução do Passo 7

O último e sétimo passo da 101 tem como objetivo definir e criar instâncias dos tipos semânticos da DDSOnt. Sendo assim, pode-se citar como exemplos de instâncias do tipo semântico DDSOntADT os termos clínicos *Vomitus*, *Vomit:Bilious Finding*, *Severities*, *Moderate(severity modifier)*, *Acute abdominal pain* e *Cholecistitis*, ilustrados na Figura 5.2. Como instâncias do tipo semântico *Signal*, pode-se exemplificar os termos *SeveritySignal* e *VomitingSignal*, ilustrados na Figura 5.2. O primeiro termo é composto pelas instâncias *Severities* e *Moderate(severity modifier)*, e o segundo termo é composto pelas instâncias *Vomitus* e *Vomit:Bilious Finding*. Finalmente, como instância do tipo semântico *ClinicalDiagnosisEvaluation*, pode-se citar o termo *Acute-AbdominalPainEvaluation*, composto pelas instâncias *SeveritySignal*, *VomitusSignal*, *Cholecistitis* e *Acute Abdominal Pain*, também ilustradas na Figura 5.2.

Após a execução dos 7 passos da metodologia 101, pode-se considerar de posse dos artefatos definidos que:

- os tipos semânticos *DDSOntADT*, *Signal* e *ClinicalDiagnosisEvaluation* definidos na DDSOnt visam facilitar o entendimento semântico, por parte de aplicações computacionais, de informações sendo transferidos durante o apoio de uma decisão diagnóstica em ambiente colaborativo;
- as ontologias investigadas e estudadas na literatura procuram representar um domínio médico com grande abrangência e de forma generalizada. Porém, estes sistemas não fornecem tipos semânticos específicos, como faz a DDSOnt, para suportar uma estrutura que auxilie no compartilhamento e na recuperação de

bases de conhecimento utilizadas por médicos no apoio à decisão diagnóstica em ambiente colaborativo;

A próxima seção apresenta os softwares JDDSOnt e DDSOntWs propostos neste trabalho para auxiliar a utilização da ontologia DDSOnt por parte de aplicações clínicas.

5.3 Softwares Relacionados a DDSOnt

Procurando permitir e facilitar o uso da ontologia DDSOnt por parte de aplicações em saúde que desejam trocar informações para apoiar a decisão diagnóstica em ambiente colaborativo, este trabalho projetou a criação de uma API Java nomeada JDDSOnt e de um conjunto de Serviços Web nomeados DDSOntWs. A JDDSOnt auxilia os programas em saúde a gerarem documentos semânticos de acordo com a ontologia DDSOnt, enquanto que DDSOntWs permite que aplicações possam criar, publicar e procurar por uma base de conhecimento, e ainda obter uma resposta diagnóstica a partir de uma base. Estes softwares relacionados a DDSOnt são apresentados a seguir.

5.3.1 Uma Introdução aos Documentos Semânticos da DDSOnt

Para formalizar e representar a ontologia DDSOnt em termos de documentos semânticos, permitindo assim que agentes de software possam interpretá-la, documentos RDF Schema e OWL foram criados e estão disponíveis na Web no endereço do laboratório IMAGCOM em ² e ³, respectivamente. A especificação OWL, diferentemente do RDF Schema como mostrado na seção 4.4.2, permite a criação de linhas de instrução que definem cardinalidades mínimas e máximas. O documento OWL parcial apresentado a seguir mostra exemplos dessas linhas de instrução. A classe *ClinicalDiagnosisEvaluation* precisa ter pelo menos 1 *Signals* (linha 3) e no máximo 40 *Signals* (linha 4), de acordo com o estudo de caso criado, e ainda pelo menos 0 ou 1 *DiagnosisDecision* (linhas 10 e 11).

```
1 <owl:Class rdf:about="#ClinicalDiagnosisEvaluation">
2 <owl:Restriction>
3   <owl:minCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int"> 1 </owl:minCardinality>
4   <owl:maxCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int"> 40 </owl:maxCardinality>
5   <owl:onProperty>
6     <owl:ObjectProperty rdf:ID="hasSignals"/>
7   </owl:onProperty>
8 </owl:Restriction>
9 <owl:Restriction>
10  <owl:minCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int"> 0 </owl:minCardinality>
11  <owl:maxCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int"> 1 </owl:maxCardinality>
12  <owl:onProperty>
13    <owl:ObjectProperty rdf:ID="hasDiagnosisResponse"/>
14  </owl:onProperty>
15 </owl:Restriction>
16 </owl:Class>
```

Suponha que uma aplicação médica está trocando informação clínica de modo a ser apoiada na tomada de decisão diagnóstica. A *ClinicalDiagnosisEvaluation* é relacionada a síndrome de dor abdominal aguda. Os sinais *Vomitus* e *Severities* são um dos importantes sinais a serem considerados e trocados neste caso, e *Vomit:Bilious Finding* e *Moderate(severity modifier)* são os seus valores, respectivamente. Ainda neste exemplo, *Cholecystitis* é a resposta diagnóstica. De modo a trocar essas informações

²<http://143.107.220.180/DDSOnt/api/RDF>

³<http://143.107.220.180/DDSOnt/api/OWL>

utilizando documentos estruturados e a Web, uma instância de documento RDF é adequada neste caso. O documento parcial a seguir ilustra uma instância RDF bem-formada e válida em relação ao documento RDF Schema da ontologia DDSOnt. A instância RDF completa está disponível no endereço do laboratório ImagCom na Web em ⁴.

Analisando o documento, é importante notar que os recursos *Acute Abdominal Pain*, *Vomitus*, *Vomit:Bilious Finding*, *Severities*, *Moderate*, e *Cholecystitis* são instâncias de DDSOntADT, significando que a aplicação que receber estes dados podem entendê-los independentemente de sistema terminológico e ontológico, pois carregam consigo informações da UMLS MetaThesaurus e da DDSOnt.

A linha 1 do documento RDF a seguir indica o *namespace* da DDSOnt. Da linha 2 até a linha 6 o recurso *AcuteAbdominalPainEvaluation* do tipo semântico *Clinical-DiagnosisEvaluation* é definido sendo composto dos recursos *AcuteAbdominalPain* do tipo semântico *DDSOntADT*, *signals1* do tipo semântico *Signal*, e *Cholecystitis* do tipo semântico *DDSOntADT*. Da linha 7 até a linha 16 a instância *AcuteAbdominalPainEvaluation* é criada. Da linha 17 até a linha 20 os sinais *signals1* são definidos, sendo compostos pelas instâncias *VomitusSignal* e *SeveritiesSignal*, ambos do tipo semântico *Signal*. Da linha 21 até a linha 24 a instância *VomitusSignal* é definida com uma auto-descrição (*VomitusDescription*) e um valor(*BiliousValue*), ambos do tipo semântico *DDSOntADT*. A instância *VomitusDescription* está sendo criada da linha 25 a 34. Da linha 35 até a linha 44, a instância *BiliousValue* está sendo definida. Da linha 45 até a linha 48 a instância *SeveritySignal* é definida com uma auto-descrição (*SeverityDescription*) e um valor(*ModerateValue*), ambos do tipo semântico *DDSOntADT*. A instância *SeverityDescription* está sendo criada da linha 49 a 60. Da linha 61 até a linha 73, a instância *ModerateValue* está sendo definida. É importante notar que todos os valores contidos nas instâncias são obtidos a partir da UMLS MetaThesaurus e o sistema de vocabulário origem é o SNOMED CT.

```

1 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:DDSOnt="http://143.107.220.180/DDSOnt">
2 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/AcuteAbdominalPainEvaluation">
3 <DDSOnt:hasPathologyDescription rdf:resource="http://143.107.220.180/DDSOnt/AcuteAbdominalPain"/>
4 <DDSOnt:hasSignals rdf:resource="http://143.107.220.180/DDSOnt/signals1"/>
5 <DDSOnt:hasDiagnosisResponse rdf:resource="http://143.107.220.180/DDSOnt/Cholecystitis"/>
6 </rdf:Description>
7 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/AcuteAbdominalPain">
8 <DDSOnt:UMLSAui> A2922390 </DDSOnt:UMLSAui>
9 <DDSOnt:UMLSSui> S0354326 </DDSOnt:UMLSSui>
10 <DDSOnt:UMLSLui> L0276335 </DDSOnt:UMLSLui>
11 <DDSOnt:UMLSCui> C0740577 </DDSOnt:UMLSCui>
12 <DDSOnt:UMLSCconcept> Acute Abdominal Pain </DDSOnt:UMLSCconcept>
13 <DDSOnt:UMLSSemanticType> Sign or Symptom </DDSOnt:UMLSSemanticType>
14 <DDSOnt:UMLSDefinition>None found</DDSOnt:UMLSDefinition>
15 <DDSOnt:sourceSystemName> SNOMED CT </DDSOnt:sourceSystemName>
16 </rdf:Description>
17 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/signals1">
18 <DDSOnt:hasSignals rdf:resource="http://143.107.220.180/DDSOnt/VomitusSignal"/>
19 <DDSOnt:hasSignals rdf:resource="http://143.107.220.180/DDSOnt/SeveritySignal"/>
20 </rdf:Description>
21 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/VomitusSignal">
22 <DDSOnt:hasSignalDescription rdf:resource="http://143.107.220.180/DDSOnt/VomitusDescription"/>
23 <DDSOnt:hasSignalValue rdf:resource="http://143.107.220.180/DDSOnt/BiliousValue"/>
24 </rdf:Description>
25 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/VomitusDescription">
26 <DDSOnt:UMLSAui> A2874452 </DDSOnt:UMLSAui>
27 <DDSOnt:UMLSSui> S0002380 </DDSOnt:UMLSSui>
28 <DDSOnt:UMLSLui> L0042965 </DDSOnt:UMLSLui>
29 <DDSOnt:UMLSCui> C0042965 </DDSOnt:UMLSCui>
30 <DDSOnt:UMLSCconcept> Vomitus </DDSOnt:UMLSCconcept>
31 <DDSOnt:UMLSSemanticType> Body Substance </DDSOnt:UMLSSemanticType>
32 <DDSOnt:UMLSDefinition>None found</DDSOnt:UMLSDefinition>
33 <DDSOnt:sourceSystemName> SNOMED CT </DDSOnt:sourceSystemName>
34 </rdf:Description>
35 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/BiliousValue">
36 <DDSOnt:UMLSAui> A3783782 </DDSOnt:UMLSAui>
37 <DDSOnt:UMLSSui> S3616945 </DDSOnt:UMLSSui>
38 <DDSOnt:UMLSLui> L2883434 </DDSOnt:UMLSLui>
39 <DDSOnt:UMLSCui> C0427815 </DDSOnt:UMLSCui>
40 <DDSOnt:UMLSCconcept> Vomit: Bilious Finding </DDSOnt:UMLSCconcept>
41 <DDSOnt:UMLSSemanticType> Laboratory or Test Result </DDSOnt:UMLSSemanticType>
42 <DDSOnt:UMLSDefinition>None Found</DDSOnt:UMLSDefinition>
43 <DDSOnt:sourceSystemName> SNOMED CT </DDSOnt:sourceSystemName>
44 </rdf:Description>
45 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/SeveritySignal">

```

⁴<http://143.107.220.180/DDSOnt/api/RDF>

```

46 <DDSOnt:signalDescription rdf:resource="http://143.107.220.180/DDSOnt/SeverityDescription"/>
47 <DDSOnt:signalDescription rdf:resource="http://143.107.220.180/DDSOnt/ModerateValue"/>
48 </rdf:Description>
49 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/SeverityDescription">
50 <DDSOnt:umlsAui>A3200095</DDSOnt:umlsAui>
51 <DDSOnt:umlsSui>S1025782</DDSOnt:umlsSui>
52 <DDSOnt:umlsLui>L1224559</DDSOnt:umlsLui>
53 <DDSOnt:umlsCui>C0439793</DDSOnt:umlsCui>
54 <DDSOnt:umlsConcept>Severities</DDSOnt:umlsConcept>
55 <DDSOnt:umlsSemanticType>Qualitative Concept</DDSOnt:umlsSemanticType>
56 <DDSOnt:umlsDefinition>
57 Used of the degree of something undesirable e.g. pain or weather; also, strictness
58 </DDSOnt:umlsDefinition>
59 <DDSOnt:umlsSourceSystemName>SNOMED CT</DDSOnt:umlsSourceSystemName>
60 </rdf:Description>
61 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/ModerateValue">
62 <DDSOnt:umlsAui>A3565834</DDSOnt:umlsAui>
63 <DDSOnt:umlsSui>S3418839</DDSOnt:umlsSui>
64 <DDSOnt:umlsLui>L3107281</DDSOnt:umlsLui>
65 <DDSOnt:umlsCui>C0205081</DDSOnt:umlsCui>
66 <DDSOnt:umlsConcept>Moderate (severity modifier)</DDSOnt:umlsConcept>
67 <DDSOnt:umlsSemanticType>Qualitative Concept</DDSOnt:umlsSemanticType>
68 <DDSOnt:umlsDefinition>
69 Indicates the condition may result in noticeable adverse adverse consequences but is unlikely to be
70 life-threatening or cause permanent injury
71 </DDSOnt:umlsDefinition>
72 <DDSOnt:umlsSourceSystemName>SNOMED CT</DDSOnt:umlsSourceSystemName>
73 </rdf:Description>
74 <rdf:Description rdf:about="http://143.107.220.180/DDSOnt/Cholecystitis">
75 <DDSOnt:umlsAui>A2879849</DDSOnt:umlsAui>
76 <DDSOnt:umlsSui>S0024825</DDSOnt:umlsSui>
77 <DDSOnt:umlsLui>L0008325</DDSOnt:umlsLui>
78 <DDSOnt:umlsCui>C0008325</DDSOnt:umlsCui>
79 <DDSOnt:umlsConcept>Cholecystitis</DDSOnt:umlsConcept>
80 <DDSOnt:umlsSemanticType>Disease or Syndrome</DDSOnt:umlsSemanticType>
81 <DDSOnt:umlsDefinition>Inflammation of the GALLBLADDER; generally caused by impairment of BILE flow,
82 GALLSTONES in the BILIARY TRACT, infections, or other diseases
83 </DDSOnt:umlsDefinition>
84 <DDSOnt:umlsSourceSystemName>SNOMED CT</DDSOnt:umlsSourceSystemName>
85 </rdf:Description>
86 </rdf:RDF>

```

Para que as aplicações possam ter suporte na geração e manipulação de documentos XML como o mostrado anteriormente, um conjunto de classes em Java foi criado na forma de uma API. A essa API deu-se o nome JDDSOnt. Ela é apresentada a seguir.

5.3.2 JDDSOnt

As classes da JDDSOnt (DDSOnt Java API) foram construídas utilizando-se a JENA (*Semantic Web Framework for Java*) RDF API⁵, que fornece um conjunto de classes e interfaces para a leitura e escrita de documentos RDF de maneira simples. Especificamente, a JDDSOnt permite que aplicações que precisam trocar informações de apoio à decisão diagnóstica possam gerar e manipular documentos clínicos como o documento RDF explicado anteriormente. A DDSOnt Java API, ilustrada pela Figura 5.3 que detalha a interface dos métodos e a estrutura de pacotes, é composta pelos seguintes componentes de software:

1. *DDSOntADT*: contém métodos *getters()* e *setters()* para manipulação das propriedades da classe *DDSOntADT*. Ainda, contém variáveis e métodos para manipular a *DDSOntADT* na forma de um recurso (*DDSOntADTResource*) e de um modelo (*DDSOntADTModel*) RDF, conforme a especificação da JENA. São exemplos de partes de documentos clínicos criados ou acessados pela classe *DDSOntADT* as linhas do documento clínico RDF explicado anteriormente: de 7 a 16, de 25 a 34, de 35 a 44, de 49 a 60, de 61 a 71, e de 74 a 85;
2. *Signal*: contém métodos *getters()* e *setters()* para manipulação das propriedades relativas ao valor e a auto-descrição de um *Signal*. Ainda, contém variáveis e métodos para manipular a *Signal* na forma de um recurso (*SignalResource*) e de um modelo (*SignalModel*) RDF. São exemplos de partes de documentos clínicos

⁵<http://jena.sourceforge.net/>

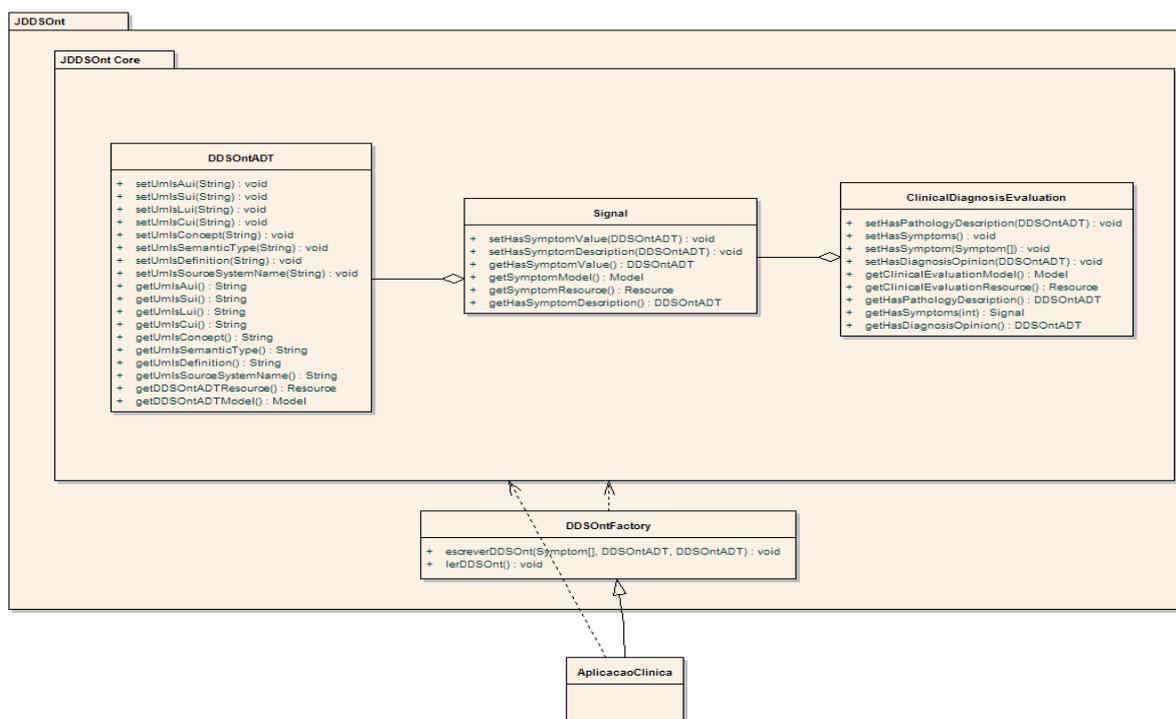


Figura 5.3: DDSOnt Java API - Auxilia na construção de documentos semânticos em conformidade a DDSOnt

criados ou acessados pela classe *Signal* as linhas do documento clínico RDF explicado anteriormente: de 21 a 24, e de 45 a 48;

3. *ClinicalDiagnosisEvaluation*: contém métodos *getters()* e *setters()* para manipulação das propriedades relativas aos vários sinais, a uma descrição patológica e a uma resposta diagnóstica da classe *ClinicalDiagnosisEvaluation*. Tem também uma variável que guarda a quantidade de sinais, e ainda, variáveis e métodos que manipulam a classe *ClinicalDiagnosisEvaluation* na forma de um recurso (*ClinicalDiagnosisEvaluationResource*) e de um modelo (*ClinicalDiagnosisEvaluationModel*) RDF. São exemplos de partes de documentos clínicos criados ou acessados pela classe *ClinicalDiagnosisEvaluation* as linhas de 2 a 6 do documento clínico RDF explicado anteriormente;
4. *DDSOnt Factory*: contém a implementação dos métodos *escreverDDSOnt()* e *lerDDSOnt()*:
 - *escreverDDSOnt*: método que gera um documento RDF como o do documento clínico explicado anteriormente a partir de um conjunto de sinais do tipo *Signal*, e de uma descrição patológica e uma resposta diagnóstica, ambas do tipo *DDSOntADT*;
 - *lerDDSOnt*: método que manipula um documento RDF como o do documento clínico explicado anteriormente. O método cria um objeto que representa um *ClinicalDiagnosisEvaluationModel*. A partir deste objeto do tipo modelo, cria-se um outro objeto que representa um *ClinicalDiagnosisEvaluationResource*. Deste objeto do tipo recurso, pode-se recuperar os recursos relacionados à síndrome, à resposta diagnóstica, e aos sinais, e através deste último recupera-se os recursos da auto-descrição e do valor dos sinais. Assim, de posse de todos estes recursos, pode-se acessar e listar o

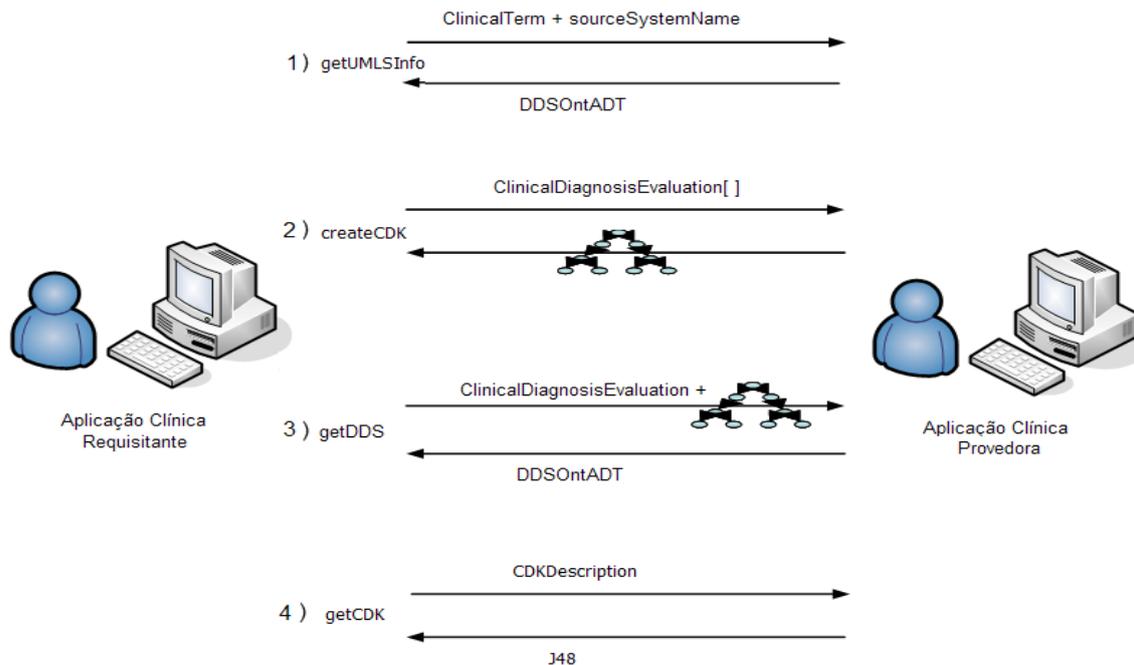


Figura 5.4: Implementação e uso dos serviços Web da DDSOntWs

conteúdo de uma avaliação clínica diagnóstica.

A Figura 5.3 ilustra também as classes *DDSOntADT*, *Signal* e *ClinicalDiagnosisEvaluation* compondo o pacote *JDDSOnt Core*. Este pacote também é sub-pacote do pacote *JDDSOnt*, que contém a classe *DDSOntFactory*, que por sua vez contém os métodos *DDSOntWriteExample()* e *DDSOntReadExample()*. Ainda na Figura 5.3, mostra-se uma aplicação clínica herdando os métodos *escreverDDSOnt()* e *lerDDSOnt()* da classe *DDSOntFactory*, e realizando as classes do pacote *JDDSOntCore*.

Todos estes componentes de software estão compactados no arquivo *JDDSOnt.jar* e disponíveis na Web em ⁶. Também foi gerado uma documentação Java (javadoc) da API para auxiliar os programadores no uso das classes e métodos da mesma. Esta documentação está disponível na Web em ⁷.

5.3.3 DDSOntWs

Para permitir a troca de informações de modo a apoiar a decisão diagnóstica em ambiente colaborativo na Web, um conjunto de 4 serviços Web foram definidos. Estes serviços permitem basicamente criar, compartilhar e descobrir bases de conhecimento, assim como obter uma resposta diagnóstica a partir destas bases. Este conjunto de serviços é detalhado a seguir, e um cenários de funcionamento destes é ilustrado pela Figura 5.4.

1. *Get UMLS Metathesaurus Information (getUMLSInfo)*: Este serviço tem como objetivo recuperar informações da classe *DDSOntADT* presentes no servidor da UMLS Metathesaurus. Para se chamar o serviço, é necessário passar como

⁶<http://143.107.220.180/DDSOnt/api/classes>

⁷<http://143.107.220.180/DDSOnt/api/javadoc>

parâmetro um termo clínico do vocabulário fonte sendo utilizado, bem como o nome do vocabulário, como SNOMED CT, por exemplo. Este serviço pode ser oferecido por uma entidade reguladora explicada mais adiante, e ser utilizado por entidades provedoras ou requisitantes de bases de conhecimento. No estudo de caso apresentado no Capítulo 6, este serviço Web foi implementado utilizando a API Java UMLS da NLM explicada no Capítulo 3. Assim, a disponibilização deste serviço promove a compatibilidade terminológica.

2. *Create Clinical Database Knowledge (createCDK)*: Este serviço tem o propósito de criar uma base de conhecimento clínico a partir de um conjunto de casos clínicos representados por um vetor de objetos da classe *ClinicalDiagnosisEvaluation*. Este serviço pode ser oferecido por uma entidade provedora de base de conhecimento, e utilizado por uma entidade requisitante de base de conhecimento para obtenção de uma resposta diagnóstica. No estudo de caso apresentado no capítulo 6, o createCDK foi implementado com o algoritmo C4.5 presente na API WEKA⁸, especificamente o algoritmo J48. Este algoritmo permite o projeto e a execução da técnica de aprendizado de máquina Árvore de Decisão. Em função disto, a base de conhecimento sendo retornada é do tipo J48, como mostra a Figura 5.4.
3. *Get Diagnosis Decision Support (getDDS)*: Este serviço é empregado na obtenção de uma resposta diagnóstica sempre representada na forma de um objeto da classe *DDSOntADT*. Quando da chamada desse serviço, deve-se passar como parâmetro a base de conhecimento representada por um objeto da classe J48, e um caso clínico representado por um objeto da classe *ClinicalDiagnosisEvaluation*. O getDDS pode ser oferecido por uma entidade provedora de resposta diagnóstica, e utilizado por uma entidade que requisita por uma resposta diagnóstica. No estudo de caso apresentado no capítulo 6, este serviço Web foi implementado utilizando a API WEKA e o algoritmo C4.5;
4. *Get Clinical DataBase Knowledge (getCDK)*: Este serviço é aplicado para a publicação e o compartilhamento de uma base de conhecimento com a comunidade médica. Para fazer uso deste serviço, deve-se passar como parâmetro informações representadas por um objeto da classe *CDKDescription*, que descreve a base de conhecimento sendo publicada. A classe *CDKDescription* é explicada adiante. O getCDK pode ser oferecido por uma entidade provedora que deseja disponibilizar sua base de conhecimento, e pode ser utilizada por uma entidade que requisita uma base de conhecimento para obtenção de uma resposta diagnóstica;

A Figura 5.5 ilustra a classe de uma aplicação clínica fazendo uso da classe *DDSOntWs* que implementa todos os serviços Web apresentados. Ainda, a figura mostra que a classe *DDSOntWs* faz uso da classe *CDKDescription*. Esta classe, como explicado anteriormente, é utilizada quando se publica ou compartilha uma base de conhecimento. As informações que a compõe estão presentes na Figura 5.5, e descritas como:

- *PathologyName*: nome da síndrome clínica que a base de conhecimento se refere na forma de um objeto *UMLSADT*;
- *SignalsQuantity*: quantidade de sinais utilizados para a criação de base de conhecimento;

⁸<http://www.cs.waikato.ac.nz/ml/weka/index.html>

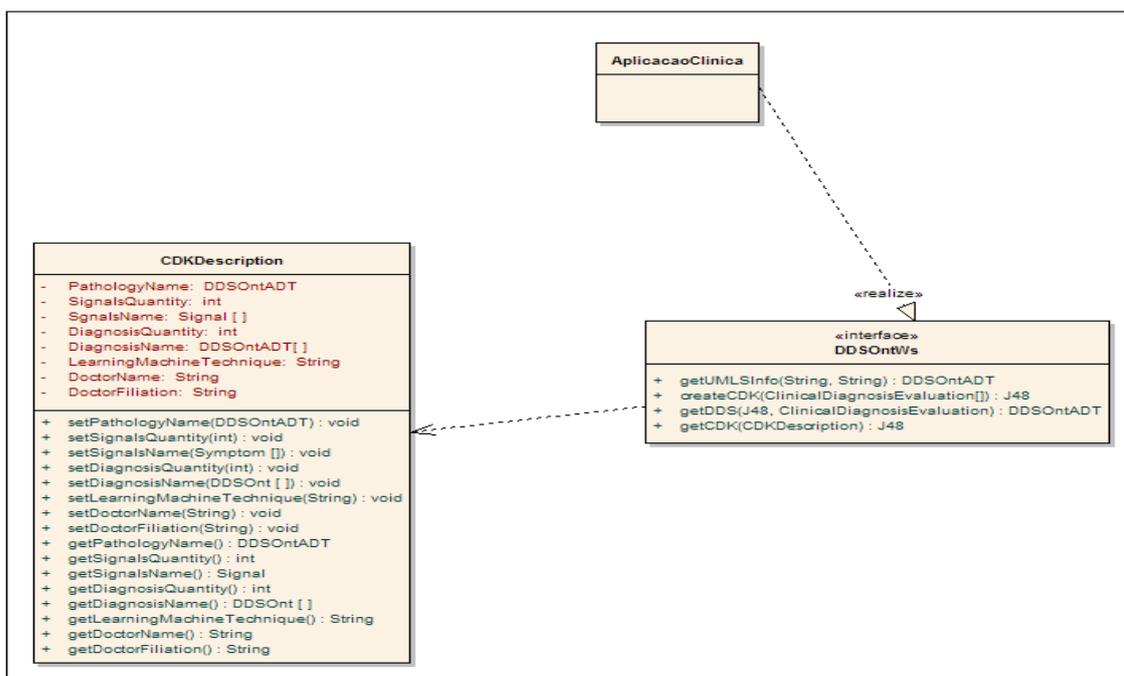


Figura 5.5: Serviços Web relacionados a DDSOnt e a classe CDKDescription

- *SignalsName*: valor e descrição na forma de objetos *Signal* de cada um dos sinais utilizados na construção da base de conhecimento;
- *DiagnosisQuantity*: quantidade de diagnósticos possíveis utilizados na criação da base de conhecimento;
- *DiagnosisName*: nome dos diagnósticos na forma de objetos *DDSOntADT* utilizados na construção da base de conhecimento;
- *LearningMachineTechnique*: técnica de aprendizado de máquina utilizada para construção da base de conhecimento;
- *DoctorName*: nome do profissional médico responsável clinicamente pela construção de base de conhecimento;
- *DoctorFiliation*: nome da instituição que o médico é filiado;

Com a proposta de se criar serviços Web semânticos neste projeto, a linguagem OWL-S, descrita na seção 2.4, foi utilizada para a descrição semântica destes. As 3 facetas, o perfil do serviço, o serviço na forma de um processo, e o fundamento do serviço, são representadas na forma de 3 arquivos OWL, *DDSOntPerfil.owl*, *DDSOntProcesso.owl* e *DDSOntFundamento.owl*, disponíveis na Web em ⁹. A utilização e referência destes arquivos pode permitir que o médico possa melhor representar o significado semântico dos seus serviços Web sendo publicados, e um outro médico possa obter melhores resultados na busca por serviços de interesse.

Durante a descrição do perfil dos serviços, foram informados o nome do serviço, uma descrição textual do serviço, quais devem ser os dados de entrada, como exemplo, os sintomas que devem ser informados para criação de uma base de conhecimento ou consulta por uma segunda opinião diagnóstica, e quais são os dados de

⁹<http://143.107.220.180/DDSOnt/Semantic>

saída, como exemplos os possíveis diagnósticos como resposta. Além disso, foram informados a categoria do serviço pelo sistema NAICS, um indicador de qualidade, e de que o serviço é válido por enquanto apenas para médicos iniciantes e residentes.

Durante a descrição dos serviços na forma de processos, procurou-se mostrar todas as maneiras que um cliente pode interagir com os mesmos, tentando facilitar assim o uso. Foram utilizados os 3 tipos de processos, atômico, simples e composto. Por exemplo, o serviço relacionado a criação de uma base de conhecimento segue a sequência de receber todos os casos clínicos segundo uma ordem especificada de sintomas, faz-se uma chamada à API Weka que implementa o algoritmo C4.5 para criar a base de conhecimento, e o retorno da base criado ao cliente.

Finalmente, durante a criação do fundamento dos serviços, objetivou-se especificar os detalhes de como acessá-los, como por exemplo apresentar o protocolo de comunicação, o formato de mensagens, a forma de serialização dos objetos, e o meio de transporte. Nesta fase, foram definidas as mensagens de entrada e/ou de saída de cada um das operações relativas a um serviço, relacionando-as ao protocolo SOAP.

De modo a promover a disponibilização e recuperação dos serviços Web relacionados a DDSOntWs, estes precisam ser publicados e descobertos de maneira universal e padronizada. Como mostrado na seção 2.3.5 do Capítulo 2, a especificação UDDI deve ser utilizada para o cumprimento destas promoções. Sendo assim, o conjunto de serviços Web da DDSOntWs precisam ser registrados em um servidor UDDI na forma de tModels e através de uma entidade reguladora ou controladora, especificamente, um conselho regional ou uma associação médica de uma especialidade, ou ainda uma junta médica especializada. A partir do registro de tModels em um UDDI, médicos provedores ou mesmo juntas médicas provedoras podem consultar o servidor UDDI e posteriormente registrar suas implementações de modo que possam ser descobertos por outras aplicações médicas. Assim, softwares clínicos de médicos requisitantes podem encontrar estas implementações no servidor UDDI e fazer uso destas.

A Figura 5.6 ilustra duas situações de troca de informações para apoiar a decisão diagnóstica em ambiente colaborativo e suportado pela DDSOntWs e também por um servidor UDDI. A figura mostrada não está contemplando a entidade reguladora, responsável por registrar os tModels dos serviços Web da DDSOntWs. Esta tarefa é realizada anteriormente às duas situações que serão apresentadas. A primeira delas apresenta softwares de médicos provedores publicando a implementação dos serviços *getUMLSInfo*, *createCDK*, *getDDS* e *getCDK* em um servidor UDDI (passo 1a da Figura 5.6). Na sequência, softwares de médicos requisitantes procuram no servidor UDDI por implementações dos serviços Web da DDSOntWs (passo 1b da Figura 5.6), e recebe como resposta os endereços das implementações relacionados à sua consulta. A situação 1c da Figura 5.6 indica os softwares dos médicos requisitantes fazendo uso do serviço Web *getUMLSInfo* encontrado para recuperar informações da UMLS Metathesaurus relacionadas aos seus sistemas de terminologia. Na sequência, a situação 1d da Figura 5.6 ilustra os softwares dos médicos requisitantes utilizando o serviço Web *createCDK*. De posse de uma base de conhecimento criada a partir de seus casos clínicos, os médicos requisitantes podem utilizar-se do serviço Web *getDDS* para solicitar por um apoio à decisão diagnóstica (passo 1e da Figura 5.6). Os médicos requisitantes podem ainda obter a base de conhecimento dos médicos provedores executando o serviço Web *getCDK* (passo 1f da Figura 5.6) e posteriormente solicitar pela segunda vez por um apoio à decisão diagnóstica a partir desta última base de conhecimento executando novamente o serviço Web *getDDS* (passo 1g da Figura 5.6).

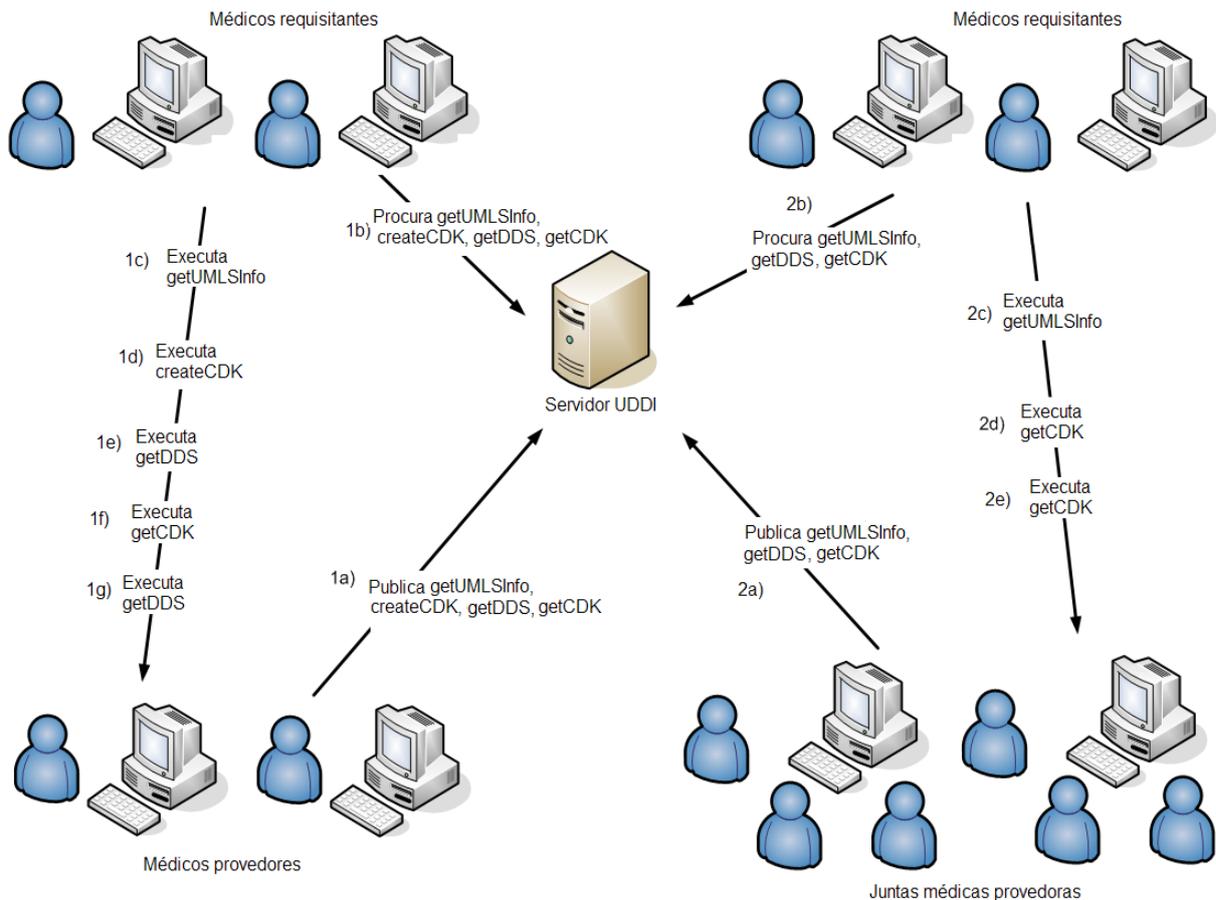


Figura 5.6: DDSOntWs com suporte UDDI

A segunda situação ilustrada pela Figura 5.6) mostra os softwares clínicos de juntas médicas especializadas publicando em um servidor UDDI as implementações dos serviços Web *getUMLSInfo*, *getDDS* e *getCDK* (passo 2a). Na situação 2b, os softwares de médicos requisitantes conectados à Internet estão procurando no servidor UDDI pelas implementações registradas dos serviços Web da DDSOntWs, recebendo como resposta os endereços das implementações relacionados à sua consulta. A situação 2c da Figura 5.6 indica os softwares dos médicos requisitantes fazendo uso do serviço Web *getUMLSInfo*. Na sequência, a situação 2d da Figura 5.6 ilustra os softwares dos médicos requisitantes utilizando o serviço Web *getCDK* para obterem bases de conhecimento das juntas médicas provedoras. Posteriormente, estes médicos requisitantes podem solicitar apoio à decisão diagnóstica a partir destas bases de conhecimento recuperadas executando o serviço Web *getDDS* (passo 2f da Figura 5.6).

As duas situações apresentadas anteriormente ilustram como é possível, através do conjunto de serviços Web da DDSOntWs, que médicos provedores e requisitantes de bases de conhecimento e de respostas diagnósticas, possam trabalhar de forma colaborativa para uma decisão diagnóstica na Web. Em especial, médicos requisitantes podem recuperar bases de conhecimento de vários médicos provedores para obtenção de várias respostas diagnósticas, utilizando uma solução que promove a compatibilidade técnica e terminológica. Os serviços Web da DDSOntWs fazem uso das classes disponibilizadas pela API JDDSOnt, que são relacionadas aos tipos semânticos definidos pela DDSOnt. Assim, a solução proposta também promove a compatibilidade ontológica.

5.4 Considerações Finais

Este capítulo apresentou e discutiu a ontologia DDSOnt proposta neste trabalho, e ainda mostrou dois softwares relacionados a DDSOnt que oferecem suporte ao uso desta por aplicações clínicas: a JDDSOnt e a DDSOntWs. A ontologia DDSOnt mostra-se na literatura como uma ontologia original no sentido de fornecer tipos semânticos específicos para suportar uma estrutura que auxilie no compartilhamento e na recuperação de bases de conhecimento utilizadas por médicos no apoio à decisão diagnóstica em ambiente colaborativo. A DDSOnt reutiliza a ontologia UMLS SN com a justificativa de não redefinir tipos semânticos genéricos e abstratos já definidos por esta.

A JDDSOnt é uma API Java desenvolvida com a JENA RDF API que auxilia as aplicações clínicas a criarem e manipularem documentos semânticos relacionados à ontologia DDSOnt. Finalmente, DDSOntWs é um conjunto de serviços Web que permite que as aplicações clínicas na Web possam ser apoiadas na decisão diagnóstica, e ainda permite a criação, a busca e a disponibilização de bases de conhecimento clínico de acordo com a DDSOnt para geração de respostas diagnósticas.

O próximo capítulo apresenta os resultados obtidos com a construção de um estudo de caso relacionado a síndrome de dor abdominal aguda, que faz uso da ontologia DDSOnt, e dos softwares relacionados JDDSOnt e DDSOntWs apresentados neste capítulo.

Resultados da Implementação do Estudo de Caso

6.1 *Considerações Iniciais*

Este capítulo tem como objetivo apresentar os resultados e experiências obtidas durante a criação de um estudo de caso que faz uso da ontologia DDSOnt, da API JDDSOnt e do conjunto de serviços Web DDSOntWs para apoiar a decisão diagnóstica em ambiente colaborativo relacionada a síndrome de dor abdominal aguda. Basicamente, o estudo de caso é composto de três módulos que suportam as atividades realizadas por três atores diferentes, definidos neste estudo como:

- uma entidade reguladora;
- uma entidade provedora; e
- uma entidade requisitante.

Cada uma destes modos é explicado neste Capítulo.

6.2 *Considerações Iniciais do Estudo de Caso*

6.2.1 *A Síndrome, os Sinais Clínicos e Laboratoriais, e os Diagnósticos*

O estudo de caso descrito neste Capítulo foi desenvolvido com o apoio médico do Dr. Ricardo Aprille Halah, proctologista e proprietário da clínica médica SH Clínicas, situada na cidade de Ribeirão Preto, SP. A definição da utilização da síndrome de dor

abdominal aguda neste estudo de caso foi liderada pelo próprio médico que possui interesse clínico e amplo conhecimento nesta patologia. Ainda, a dor abdominal aguda é uma síndrome comum, com muitos sinais e diagnósticos e portanto, clinicamente rica.

A partir do estudo que o Dr. Halah realizou na obra de Porto em (Porto 1991), verificou-se os sinais clínicos e laboratoriais, e os diagnósticos mais frequentes na síndrome de dor abdominal aguda. A partir deste estudo, foram definidos 33 sinais clínicos, 7 sinais laboratoriais, e também 13 diagnósticos. Os 33 sinais clínicos foram divididos em 3 categorias, que são:

1. identificação do paciente;
2. queixa-duração; e
3. exame físico.

Para cada um dos sinais, foram definidos valores discretos ou contínuos. A primeira categoria dos sinais clínicos tem como sinal apenas *Sex*, que possui os valores discretos *Male Gender* e *Female Gender*. Outros valores clínicos que identificam o paciente mas não são considerados sinais clínicos são a identificação dele no sistema, o nome, e a idade. A categoria de queixa/duração dos sinais clínicos apresenta 22 sinais, dos quais 21 possuem valores discretos, como *Location* e *Radiation*, e apenas 1 sinal com valor contínuo, o (*Menstruation*), como informa a Tabela 6.1. Esta tabela lista todos os sinais clínicos de queixa/duração juntamente com seus possíveis valores.

A terceira categoria dos sinais clínicos é formada pelos sinais físicos mostrados na Tabela 6.2. Estes sinais são compostos de apenas um elemento com valor contínuo, a *temperature*, e o restante dos sinais possuem valores discretos, como *Tachycardia* e *Auscultation*.

Os 7 sinais laboratoriais do estudo de caso em dor abdominal aguda são formados por 2 sinais discretos, *Left shifted white blood cells* e *Hematuria*, e 5 sinais contínuos, como *Leukocytes* e *SGOT - Glutamate oxaloacetate transaminase*. Estes 7 sinais são representados na Tabela 6.3.

Os 13 possíveis diagnósticos definidos no estudo de caso da patologia de dor abdominal aguda são *Appendicitis*, *Cholecystitis*, *Peptic Ulcer*, *Gastroenteritis*, *Diverticulitis*, *Pancreatitis*, *Hepatitis*, *Intestinal Obstruction*, *Pelvic Inflammatory Disease*, *Thrombosis of mesenteric vein*, *Dysmenorrhea*, *Pregnancy*, *Ectopic*, e *Ovarian Pain*.

Todos os nomes dos sinais clínicos e laboratoriais, os valores discretos que estes podem assumir, assim como os diagnósticos apresentados nesta seção são propositalmente conceitos da UMLS *MetaThesaurus*. Estes valores da UMLS serão efetivamente trocados e compartilhados por médicos no estudo de caso para promover o apoio à decisão diagnóstica em ambiente colaborativo e com compatibilidade terminológica. Ainda, cada um destes valores carrega sempre consigo informações das propriedades do tipo semântico *DDSOntADT* da ontologia *DDSOnt*, que é uma especialização dos tipos semânticos *Entidade* e *Eventos* da UMLS *SN (Semantic Network)*.

A composição do conjunto de termos ('auto-descrição do sinal clínico', 'valor discreto do sinal clínico') ou ('auto-descrição do sinal clínico', 'valor contínuo do sinal clínico') ou ainda ('auto-descrição do sinal laboratorial', 'valor contínuo do sinal laboratorial') ou ('auto-descrição do sinal laboratorial', 'valor discreto do sinal laboratorial')

Tabela 6.1: Sinais clínicos de Queixa/Duração Definidos Para o Estudo de Caso

| Sinais clínicos | Possíveis valores do sinal |
|--|--|
| <i>Abdominal Pain</i> | <i>Yes - yes/no indicator ou No - yes/no indicator</i> |
| <i>Location</i> | <i>Right Lower Quadrant Pain ou Right Upper Quadrant Pain ou [D]Epigastric Pain ou Left Lower Quadrant Pain ou [D]Suprapubic pain ou Mesogastric Region ou Pelvis</i> |
| <i>Physiologic Migration</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |
| <i>Firstly</i> | <i>Right Lower Quadrant Pain ou Right Upper Quadrant Pain ou [D]Epigastric Pain ou Left Lower Quadrant Pain ou [D]Suprapubic pain ou Mesogastric Region ou Pelvis ou No Information - not applicable</i> |
| <i>Initially</i> | <i>No Pain ou Colicky sensation quality ou Cramping sensation quality ou Burning sensation ou Gnawing sensation quality ou Sustained ou Intermittent ou Diffuse</i> |
| <i>Changing</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |
| <i>End-stage</i> | <i>No Pain ou Colicky sensation quality ou Cramping sensation quality ou Burning sensation ou Gnawing sensation quality ou Sustained ou Intermittent ou Diffuse ou No Information - not applicable</i> |
| <i>Severities</i> | <i>Sensory Discomfort ou Moderate (severity modifier) ou Severe</i> |
| <i>Changed status</i> | <i>Increased ou Decreased ou Sustained ou No Information - not applicable</i> |
| <i>Radiation</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |
| <i>Location</i> | <i>Absent ou Back ou Entire right shoulder region ou Entire left shoulder region ou Left Scapula ou Right Scapula ou Pelvis ou Genitalia ou No Information - not applicable</i> |
| <i>Symptom aggravating factors</i> | <i>Absent ou Standard meal ou Movement ou Fatty meal ou Sexual Intercourse - finding ou No Information - not applicable</i> |
| <i>Relieved by</i> | <i>Absent ou Food ou Antacids ou Sitting position ou Limitation of joint movement ou Fatty Meal ou Vomiting ou No Information - not applicable</i> |
| <i>Defecation</i> | <i>Obstipation ou Diarrhea ou Hematemesis ou Melena ou Hematochezia ou Normal ou No Information - not applicable</i> |
| <i>Icterus</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |
| <i>Alcoholic Intoxication, Chronic</i> | <i>Yes - yes/no indicator ou No - yes/no indicator</i> |
| <i>Anorexia</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |
| <i>Vomiting</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |
| <i>Vomit</i> | <i>Translucent ou Vomit: bilious finding ou Feculent vomit O/E ou No Information - not applicable</i> |
| <i>Menstruation</i> | <i>real</i> |
| <i>Dyspareunia</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |
| <i>Vaginal discharge</i> | <i>Yes – yes/no indicator ou No - yes/no indicator</i> |

Tabela 6.2: Sinais clínicos de Exame Físico Definidos Para o Estudo de Caso

| Sinais clínicos | Possíveis valores do sintoma |
|--------------------------------|--|
| <i>Tachycardia</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |
| <i>Auscultation</i> | <i>Normal</i> ou <i>Mettalic Sounds</i> ou <i>Hypokinesia</i> ou <i>Hyperactive behavior</i> ou <i>No Information - not applicable</i> |
| <i>Hydration status</i> | <i>Hypokinesia</i> ou <i>Normal</i> ou <i>Hyperactive behavior</i> ou <i>No Information - not applicable</i> |
| <i>Pathological Dilatation</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |
| <i>Temperature</i> | <i>real</i> |
| <i>Murphy's sign</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |
| <i>Decompression, external</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |
| <i>ABDOMEN RLQ PAIN</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |
| <i>Rebound tenderness</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |

Tabela 6.3: Sinais Laboratoriais Definidos Para o Estudo de Caso

| Sintoma | Possíveis valores do sintoma |
|---|---|
| <i>Leukocytes</i> | <i>real</i> |
| <i>Left shifted white blood cells</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |
| <i>Hematuria</i> | <i>Yes - yes/no indicator</i> ou <i>No - yes/no indicator</i> |
| <i>Amylases</i> | <i>real</i> |
| <i>Alkaline Phosphatase</i> | <i>real</i> |
| <i>SGOT - Glutamate oxaloacetate transaminase</i> | <i>real</i> |
| <i>SGPT - Glutamate pyruvate transaminase</i> | <i>real</i> |

gera as instâncias dos tipos semânticos *Signals* da DDSOnt. E finalmente, a composição da entidade (Descrição patológica, *Signals**, resposta diagnóstica?) irá compor o tipo semântico *ClinicalDiagnosisEvaluation* da DDSOnt. O símbolo ‘?’ pós-fixado em resposta diagnóstica significa que este pode ou não aparecer, e o símbolo ‘*’ pós-fixado em *Signals* significa que este pode aparecer 0 ou várias vezes.

6.2.2 Entrada Estruturada dos Dados

Quando médicos fazem uso de aplicações que possuem recursos de apoio à decisão diagnóstica clínica, é comum que a entrada de dados seja do tipo estruturada, principalmente quando o valor de um sinal clínico ou de um sinal laboratorial é do tipo discreto. Isso significa que o médico não necessariamente digita o valor referente ao sinal, mas sim escolhe um dos valores permitidos em uma lista predefinida de nomes ou valores. Por exemplo, quando o médico informa na aplicação a localização da dor abdominal, o mesmo não pode digitar um valor qualquer, ou seja, não pode usar texto livre. Sendo assim, o médico deve escolher dentre os valores possíveis, como exemplo *Right Lower Quadrant Pain*, *Right Upper Quadrant Pain*, *[D]Epigastric Pain*, *Left Lower Quadrant Pain*, *[D]Suprapubic pain*, *Mesogastric Region* ou ainda *Pelvis*. Esta característica limita a entrada de dados. Por outro lado, a entrada estrutura de dados permite que as técnicas de aprendizado de máquina possam funcionar mais adequadamente, sem que os diferentes termos léxicos de uma palavra possam atrapalhar ou confundir o aprendizado da máquina.

6.2.3 Requisitos de Software do Estudo de Caso

Para a implementação do estudo de caso sendo descrito neste Capítulo, foram instalados dois softwares servidores no laboratório ImagCom. O primeiro software é um servidor de serviços Web da Apache ¹ chamado AXIS 1.X ², que implementa o protocolo SOAP especificado pela W3C, e é executado sobre o servidor de WWW Apache Tomcat ³. O segundo software instalado também é um servidor da Apache, este chamado jUDDI⁴, uma implementação da especificação UDDI e com suporte a vários sistemas gerenciadores de banco de dados relacionais, como DB2, SyBase, HSQLDB e MySQL⁵. Este último é utilizado para gerenciar e armazenar as informações clínicas do estudo de caso.

Três módulos de software foram criados no estudo de caso:

- o primeiro apóia as atividades realizadas por uma entidade reguladora nas definições dos serviços da DDSOntWs;
- o segundo módulo apóia as atividades de uma junta médica ou médico provedor nas implementações dos serviços definidos pela entidade reguladora; e finalmente,
- o terceiro módulo apóia as atividades de um médico requisitante na utilização dos serviços implementados pela entidade provedora.

A Figura 6.1 ilustra os requisitos de software do estudo de caso criado, e ainda resume como ocorre o relacionamento entre os módulos de software das entidades participantes do estudo de caso. De acordo com a figura:

- As entidades reguladora, provedora e requisitante fazem uso da ontologia DDSOnt, da API JDDSOnt e do conjunto de serviços Web da DDSOntWs para que possa haver apoio à decisão diagnóstica em ambiente colaborativo. Ainda, estas utilizam o banco de dados MySQL e as versões cliente e servidor do AXIS;
- As entidades reguladora, provedora, requisitante, e o servidor UDDI comunicam-se através do protocolo SOAP;
- A entidade reguladora é responsável por definir e registrar no servidor UDDI um *tModel* para cada cada serviço Web da DDSOntWs. É através deste *tModel* que as entidades provedoras e requisitantes tomam conhecimento sobre qual síndrome, e quais os sinais e os diagnósticos clínicos são utilizados nos serviços Web. Ainda, a entidade reguladora deve implementar o serviço *getUMLSInfo* com o auxílio da JDDSOnt, e posteriormente registrá-la no servidor UDDI;
- A entidade provedora deve consultar por *tModels* de interesse e registrados no servidor UDDI, e a partir da definição destes, implementar os serviços *createCDK*, *getCDK* e *getDDS* com o auxílio da JDDSOnt, e do serviço Web *getUMLSInfo* oferecido pela entidade reguladora. Após as implementações, a entidade provedora deve registra-las no UDDI;

¹<http://www.apache.org>

²<http://ws.apache.org/axis/>

³<http://tomcat.apache.org/>

⁴<http://ws.apache.org/juddi/>

⁵<http://www.mysql.com>

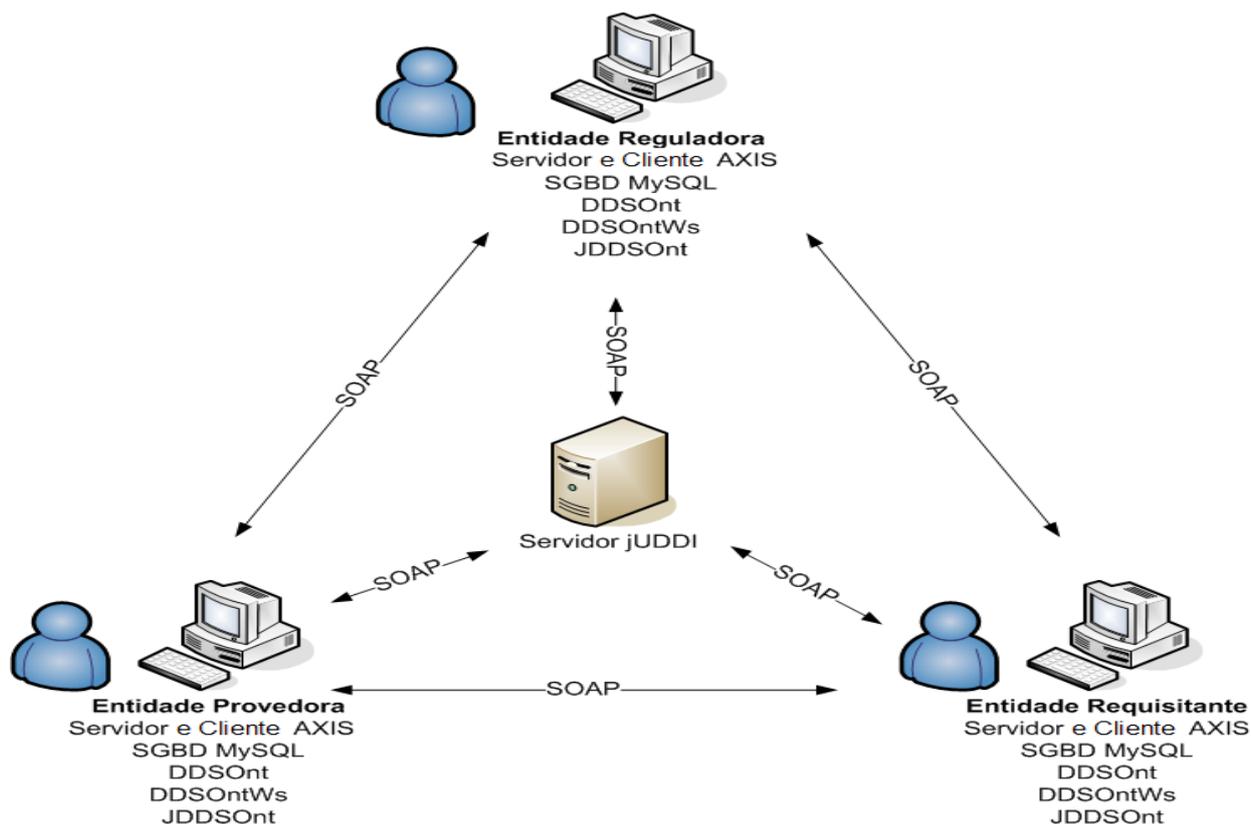


Figura 6.1: Diagrama de Requisitos de Software do Estudo de Caso

- A entidade requisitante precisa consultar no servidor UDDI os tModels de seu interesse, e a partir do resultado consultar as implementações do mesmo. Assim, a entidade requisitante pode utilizar os serviços implementados pela entidade provedora, e ser apoiado por esta nas atividades de consulta por uma base de conhecimento, de obtenção por uma resposta diagnóstica, e eventualmente de criação de uma própria base de conhecimento. Durante a utilização dos serviços, a entidade requisitante é apoiada pela JDDSOnt e pelo serviço *Web getUMLSInfo* oferecido pela entidade reguladora.

Os módulos das entidades reguladoras, provedoras e requisitantes são apresentados mais detalhadamente a seguir:

6.3 Módulo de Apoio a uma Entidade Reguladora

O primeiro módulo criado no estudo de caso tem como objetivo apoiar uma entidade reguladora. Uma entidade reguladora pode na realidade representar um conselho regional ou mesmo um conselho nacional de uma especialidade médica, ou então um grupo de trabalho de um centro médico avançado e de referência. Neste estudo de caso, a entidade reguladora é representada pelo próprio Dr. Halah, que baseou suas ações no livro de Porto (Porto 1991) para simular este papel. Este módulo apóia a entidade reguladora na realização das seguintes atividades ilustradas no Diagrama de Caso de Uso em UML da Figura 6.2:

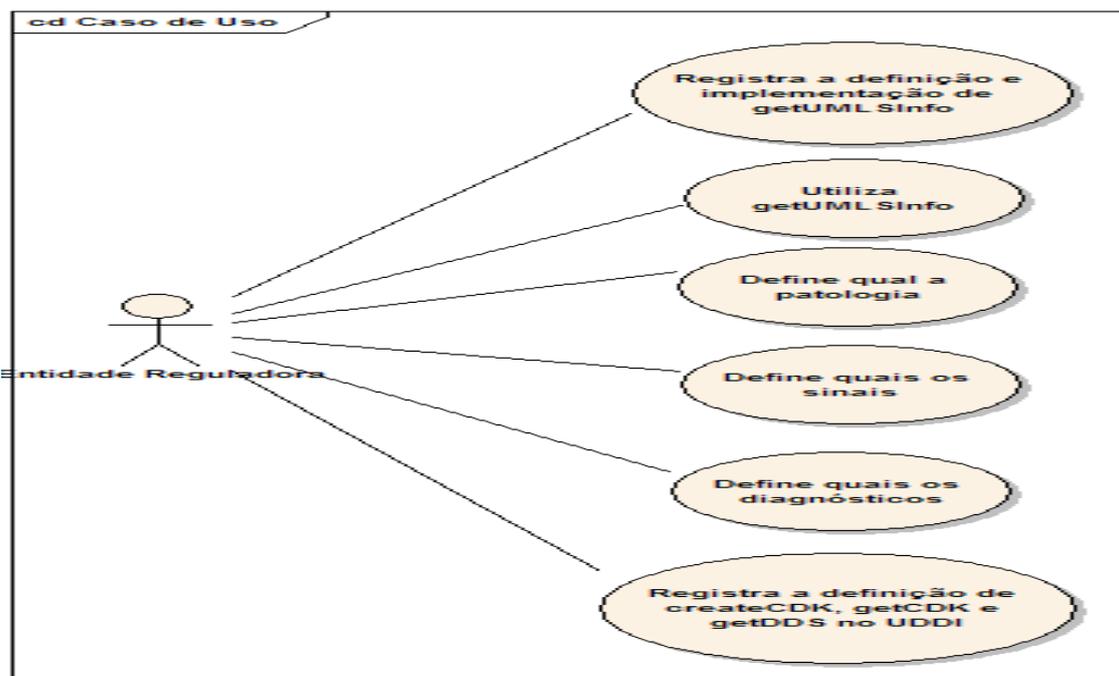


Figura 6.2: Diagrama de Caso de Uso para Módulo de Apoio a Entidade Reguladora.

1. Registrar no servidor UDDI a definição e a implementação do serviço *getUML-Sinfo* da DDSOntWs para que este seja encontrado e utilizado pelas aplicações provedoras e requisitantes de apoio à decisão diagnóstica;
2. Utilizar o serviço *getUMLSInfo* pois os valores de patologia, sinais e diagnósticos que serão definidos devem originar-se da UMLS MetaThesaurus e conter informações da DDSOntADT;
3. Definir qual a patologia utilizada, como dor abdominal aguda ou diabetes;
4. Definir quais os sinais clínicos e laboratoriais da patologia, juntamente com seus respectivos e possíveis valores contínuos ou discretos;
5. Definir quais os possíveis diagnósticos da patologia;
6. Registrar a definição dos serviços *createCDK*, *getCDK* e *getDDS* da DDSOntWs no servidor UDDI na forma de *tModels* de modo que os mesmos possam ser descobertos, e implementados ou utilizados pelas aplicações provedoras e requisitantes de apoio à decisão diagnóstica, respectivamente.

De modo a suportar a realização das atividades de 1 a 6 da entidade reguladora, foi desenvolvido um módulo de software. Este módulo é composto de 4 painéis sobrepostos e selecionáveis por abas intituladas *Get DDSOntADT Info*, *Define Pathology*, *Signals and Diagnosis*, *Registry DDSOntWs Definitions* e *Registry a Service Implementation*. Estes painéis foram utilizados pelo Dr. Halah, com o suporte técnico do autor desta tese, para o preenchimento adequado dos mesmos. Estes são apresentados a seguir.

O acionamento da aba *Define Pathology*, *Signals and Diagnosis* disponibiliza um painel contendo recursos para apoiar as atividades de 3 a 5. Este painel é ilustrado pela Figura 6.3. O painel é dividido em 3 áreas. A área mais acima é utilizada para

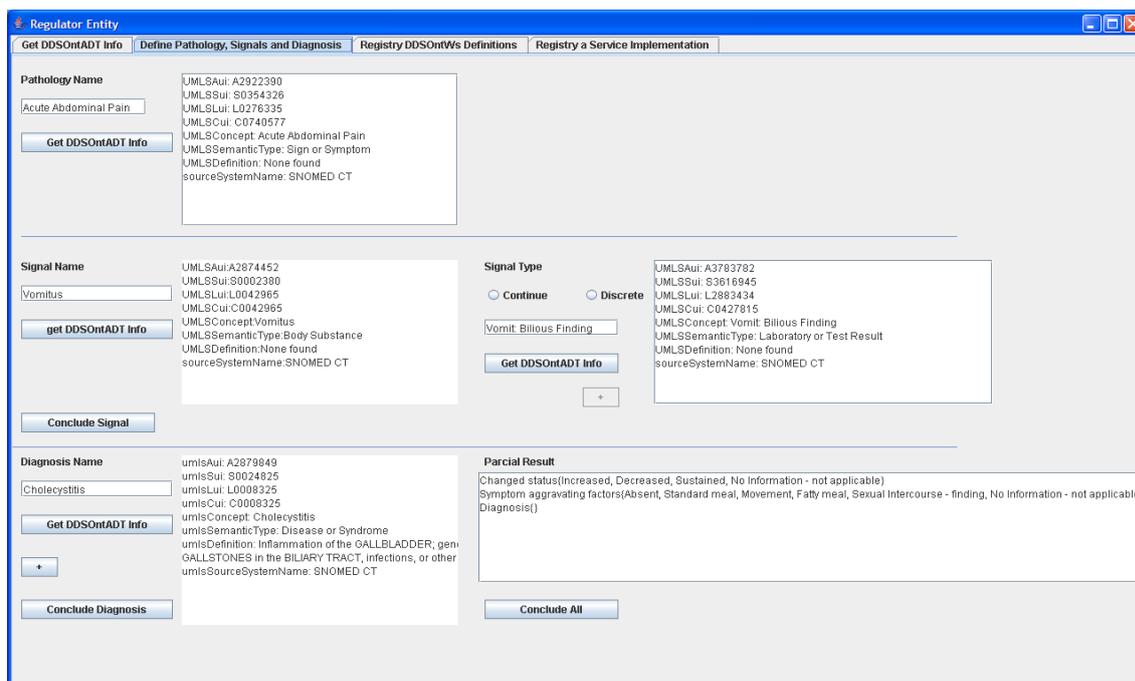


Figura 6.3: Tela para a definição da patologia, dos sinais e diagnósticos.

a entidade reguladora definir o nome da patologia. O nome escolhido deve estar de acordo com a estrutura DDSOntADT da DDSOnt. Para isto, o usuário deve acionar o botão intitulado *Get DDSOntADT Info*, que abre outro painel para auxiliar a obtenção dos valores da UMLS, mostrados como exemplo no campo de texto ao lado do botão. O painel que se abre ao acionar o botão intitulado *Get DDSOntADT Info* é explicado mais adiante. A segunda área do painel relacionado à primeira aba é utilizada para a entidade reguladora definir todos os sinais clínicos e laboratoriais juntamente com seus possíveis valores. Para cada nome de sinal escolhido, deve-se associar informações da DDSOntADT, e para isso deve-se acionar o botão *getUMLSInfo*, que age da mesma forma que o botão anterior. Para cada sinal, define-se os valores possíveis e para cada um destes, define-se se o valor é contínuo ou discreto. Para valores discretos, deve-se associá-lo às informações da DDSOntADT clicando no botão *Get DDSOntADT Info* de mesmo efeito dos outros. O botão intitulado '+' finaliza a definição de um valor possível de sinal e permite que se crie um novo valor possível deste. Ao finalizar os valores possíveis de um sinal, pode-se acionar o botão intitulado *Conclude Signal*, para então se definir um novo sinal. Finalmente, a terceira área permite a entidade reguladora definir os possíveis valores diagnósticos. Estes valores também precisam conter informações DDSOntADT, e para isso o usuário precisar acionar o botão intitulado *Get DDSOntADT Info*. O lado direito desta terceira área contém uma caixa de texto que permite o usuário visualizar um resultado parcial dos valores de sinais e diagnósticos sendo criados para uma patologia. Para finalizar todo o processo, aciona-se no botão com título *Conclude All* para arquivamento dos dados em XML. Pela Figura 6.3, percebe-se que já foram definidos os sinais *Changed status* e *Symptom aggravating factors*, seus possíveis valores, definindo-se assim o diagnóstico *Cholecystitis* para a patologia *Acute Abdominal Pain*. Como descrito no início deste Capítulo, foram criados neste estudo de caso 40 sinais e 13 diagnósticos.

O acionamento do botão com título *Get DDSOntADT Info* do painel explicado anteriormente aciona a aba *Get DDSOntADT Info*, que por sua vez disponibiliza um painel contendo recursos para apoiar a atividade 2 da entidade reguladora. Este painel é

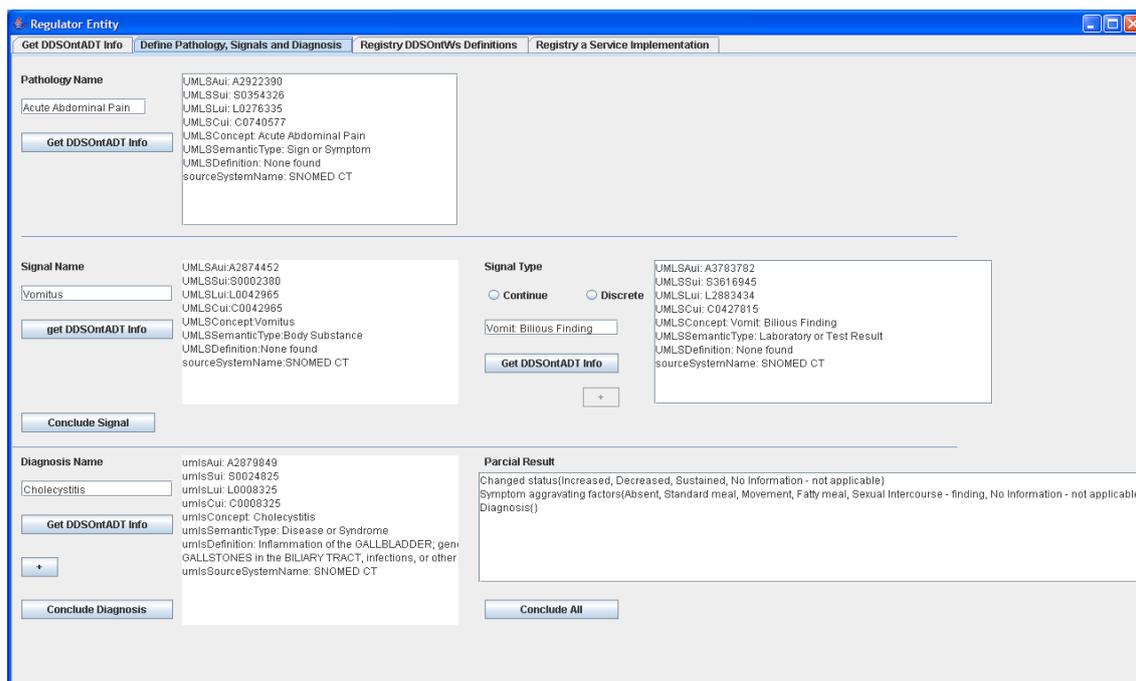


Figura 6.4: Tela para a definição da patologia, dos sinais e diagnósticos suportados pelo serviço getUMLSInfo

ilustrado pela Figura 6.4 e esta dividido em 3 áreas. A área superior é utilizada para a entidade reguladora consultar informações sobre um termo clínico. Para isto, o usuário pode acionar o botão contendo o título *MetaThesaurus Concept Term* e visualizar o resultado no campo de texto ao lado direito. O resultado contém as informações do conceito, do identificador do conceito, do tipo semântico, da definição e dos sinônimos do termo clínico pesquisado. A partir deste resultado, o usuário utiliza-se da área abaixo à esquerda para obter o restante das informações da DDSOntADT. Para isso, ele precisa preencher o nome do termo escolhido (um dos sinônimos resultantes na busca feita anteriormente), o identificador do conceito escolhido, e o nome do seu sistema de terminologia médico, para então clicar no botão com título *Get DDSOntADT*. Este botão finalmente insere na terceira área mais à direita as informações da DDSOntADT associadas ao termo pesquisado.

O acionamento da aba *Registry DDSOntWs Definitions* disponibiliza um painel contendo recursos para apoiar as atividades 1 e 6 da entidade reguladora relacionadas ao registro das definições dos serviços da DDSOntWS em um servidor UDDI. Este painel é ilustrado pela Figura 6.5 que registra a definição do serviço getCDK da DDSOntWs. Este painel também serviu para a definição dos serviços *getDDS*, *createCDK* e *getUMLSInfo*. Todos os serviços foram registrados utilizando-se do pacote JAXR (*Java API for XML Registries*)⁶.

Após o registro ter sido realizado, o mesmo pode ser encontrado para ser implementado por uma entidade provedora ou então utilizada por um entidade requisitante. A partir deste painel, o usuário deve entrar com as seguintes informações relacionadas à estrutura de dados de um *tModel* do serviço Web sendo registrado:

- o nome do serviço Web;

⁶<http://java.sun.com/webservices/jaxr/index.jsp>

Figura 6.5: Tela para registrar a definição de serviços da DDSOntWs.

- o nome da pessoa autorizada a manipular tecnicamente o serviço;
- a descrição do serviço, como exemplo, uma descrição informal e os parâmetros de entrada e saída do mesmo;
- os detalhes do serviço ou de algum tipo de parâmetro de entrada ou saída;
- o identificador e a categoria do serviço a partir de um sistema de classificação;
- o endereço IP do servidor UDDI onde será registrado o serviço;
- o endereço IP da máquina utilizada pela entidade reguladora;
- a quantidade e os valores de sinais, e a quantidade e os valores de diagnósticos. Todos estes são obtidos automaticamente ao acionar o botão entitulado *Load*, como representado na Figura 6.5, que por sua vez recupera os dados do arquivo XML criado no painel com aba *Define Pathology, Signals and Diagnosis*;
- o nome e a filiação do médico; e finalmente,
- a técnica de aprendizado de máquina utilizada;

Todas estas informações são obrigatórias e devem ser informadas pela entidade reguladora no momento do registro da definição. Durante a escolha dos valores do nome da patologia, e dos sinais e diagnósticos utilizados na definição, é necessário que o usuário informe as propriedades do tipo semântico DDSOntADT. Para isso, este deve clicar no botão entitulado *Get DDSOntADT Info* que abre um painel semelhante ao painel aberto pelo botão *Get DDSOntADT Info* da aba *Define Pathology, Signals and Diagnosis*. Ao final do registro, um *tModelID* é devolvido do servidor UDDI para a aplicação da entidade reguladora dentro da aba sendo discutida. Este *tModelID* identifica unicamente os *tModels* registrados no servidor.

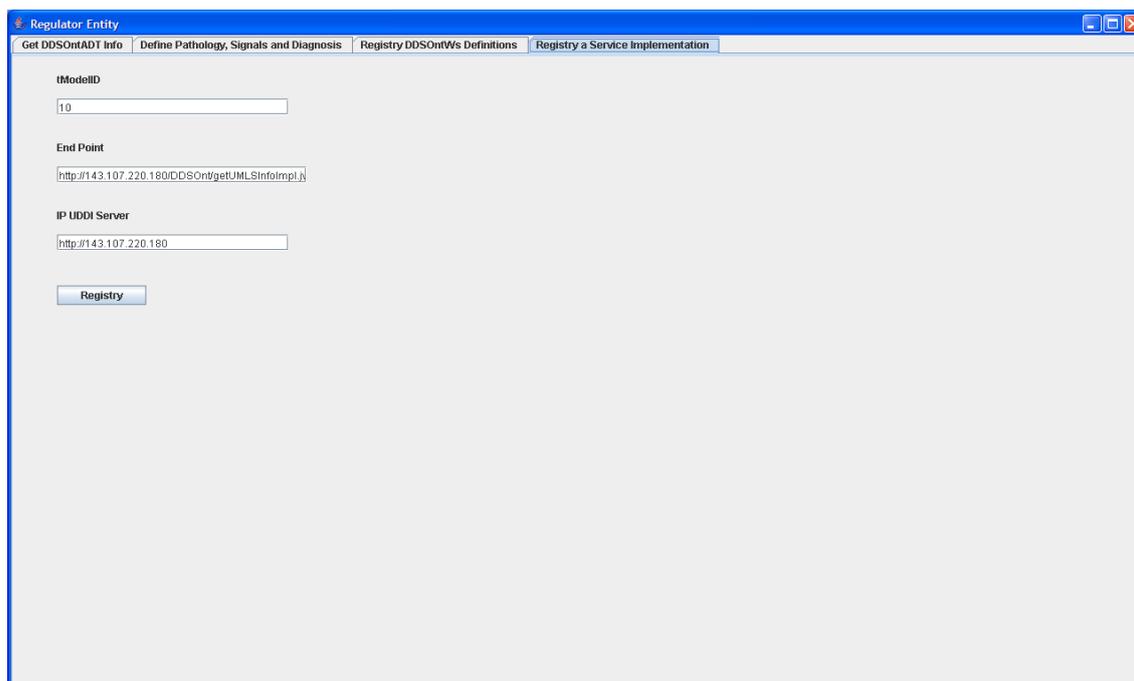


Figura 6.6: Tela para registrar a implementação do serviço getUMLSInfo.

Durante o uso dos serviços da DDSontWS pode ser necessário a atualização dos valores de sinais e diagnósticos, ou seja, sinais ou diagnósticos podem ser adicionados ou removidos pela entidade reguladora, o que provocará a atualização da definição de um registro no servidor UDDI. Eventualmente, estas mudanças podem ocorrer em função de sugestões das entidades requisitantes ou provedoras. Eventualmente ainda, o excesso de atualizações pode também necessitar a definição de um novo registro dos serviços DDSontWs de mesma patologia, porém com diferentes sinais e diagnósticos.

Para finalizar a explicação do módulo da entidade reguladora, o acionamento da aba *Registry a Service Implementation* disponibiliza um painel contendo recursos para apoiar também a atividade 1 da entidade reguladora. Este painel, ilustrado pela Figura 6.6, é utilizado para que a entidade reguladora possa registrar a implementação do único serviço Web de sua responsabilidade, o getUMLSInfo. A implementação dos outros serviços da DDSontWs são registrados pela entidade provedora. Para o registro da implementação de getUMLSInfo, o usuário informa o identificador *tModelID* do *tModel* localizado no servidor UDDI, o local onde se encontra o *endPoint* desta implementação, e o endereço IP do servidor UDDI. A Figura 6.6 ilustra o registra desta implementação. A implementação deste serviço foi realizada utilizando-se o pacote JAXR (*Java API for XML Registries*)

Particularmente em relação aos detalhes de implementação do serviço Web getUMLSInfo realizado pela entidade reguladora, foram utilizadas a classe interface DDSontWS, apresentada na seção refddsws, a API JDDSOnt, apresentada na seção 5.3.2, e também a API UMLSKS, apresentada na seção 3.4.

6.4 Módulo de Apoio para Juntas Médicas ou Médicos Provedores

O segundo módulo criado no estudo de caso tem como objetivo apoiar médicos ou juntas médicas provedoras de decisões diagnósticas, especificamente relacionadas à síndrome de dor abdominal aguda. Estas juntas médicas ou médicos podem ser representados na realidade por qualquer entidade que deseja oferecer serviços de apoio à decisão diagnóstica, e de criação e compartilhamento de bases de conhecimento para apoiar a decisão diagnóstica. Neste estudo de caso, a entidade provedora é representada pelo próprio Dr. Halah, que se baseou no livro de Porto (Porto 1991) para simular este papel.

A Figura 6.7 ilustra um Diagrama de Caso de Uso da UML que apresenta os requisitos e funcionalidades deste módulo. Por esta figura, pode-se perceber que a junta médica ou um médico provedor:

1. Consulta o servidor UDDI a fim de obter definições de *tModels* relacionados aos serviços Web da DDSOntWs. Particularmente, os *tModels* de *createCDK*, *getCDK* e *getDDS* são buscados pois estes serão implementados pela entidade provedora. A busca pelo *tModel* de *emphgetUMLSInfo* é necessário apenas para a utilização do serviço já implementado pela entidade reguladora;
2. Registra no servidor UDDI as implementações dos serviços da DDSOntWs obtidos na consulta anterior, especificamente os serviços *createCDK*, *getCDK*, e *getDDS*;
3. Utiliza o serviço Web *getUMLSInfo* implementado pela entidade reguladora a fim de recuperar os valores referentes às propriedades do tipo semântico *DDSOntADT*;
4. Cria o conjunto de treinamento contendo os casos clínicos com sinais e diagnósticos;
5. Utiliza o serviço Web *createCDK* para criar uma base de conhecimento relacionada ao conjunto de treinamento citado no item anterior;
6. Eventualmente, a junta médica pode querer testar o resultado de sua base de conhecimento gerada no item anterior, o que exigirá o uso do serviço *getDDS*.

De modo a suportar a realização das atividades de 1 a 6 de uma entidade provedora, foi desenvolvido um módulo de software. Este módulo é composto de 9 painéis sobrepostos e selecionáveis por abas intituladas *Search for DDSOntWs Definitions*, *Get DDSOntADT Info*, *Registry DDSOntWs Implementation*, *Patient Identification*, *Patient Complaint/Period - Part 1*, *Patient Complaint/Period - Part 2*, *Patient Physical Exam*, *Patient Laboratory Exam* e *Patient Diagnostic*. Estes painéis foram utilizados pelo Dr. Halah, com o suporte técnico do autor desta tese. Estes são apresentados a seguir.

O acionamento da aba *Search for DDSOntWs Definitions* disponibiliza um painel contendo recursos para apoiar a atividade 1 da entidade provedora. O painel é dividido em 2 áreas. A área mais à esquerda é utilizada para a entidade provedora procurar por um *tModel* no servidor, e a área mais a direita mostra o resultado desta procura. Um *tModel* é definido e registrado por uma entidade reguladora explicada na seção 6.3. Para buscar um *tModel* associado a um serviço Web da DDSOntWs, é

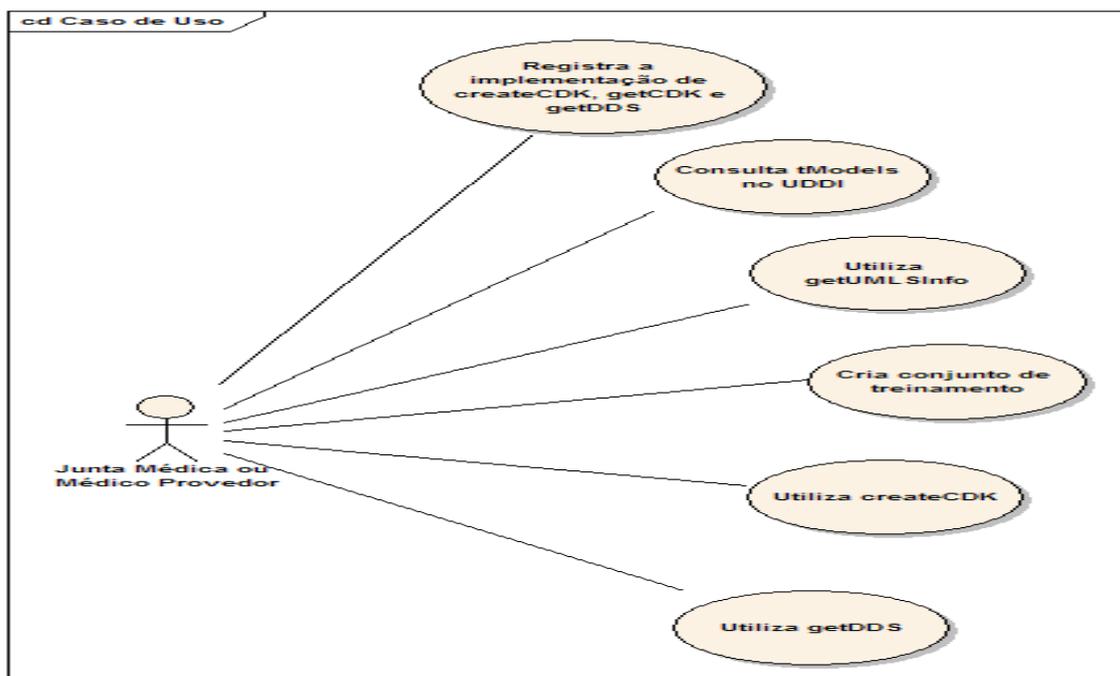


Figura 6.7: Diagrama de Caso de Uso para Módulo de Apoio à Junta Médica.

preciso que o usuário preencha as seguintes informações a respeito deste, como mostra a Figura 6.8, que ilustra a busca por um tModel relacionado ao serviço getCDK: uma descrição, o nome da patologia, a quantidade de sinais, os valores dos sinais, podendo ser algum dos descritos ou todos eles, a quantidade de diagnósticos, os valores de diagnósticos, podendo ser algum deles ou então todos eles, o nome e a filiação do médico, um identificador e um categoria relacionado a um sistema de classificação, a técnica de aprendizado de máquina, e por fim o endereço IP de onde fica o servidor UDDI. Este painel também foi utilizado para buscar a definição dos serviços getDDS, getUMLSInfo e createCDK. A busca por definição de serviços foram realizadas utilizando-se o pacote JAXR (*Java API for XML Registries*)

Todas as informações de busca apresentadas são opcionais e podem deixar de ser informadas pelo médico no momento da consulta. Durante a escolha dos valores do nome da patologia, e dos sinais e diagnósticos utilizados na busca, é necessário que o usuário informe valores das propriedades do tipo semântico DDSOntADT. Para isso, este deve clicar no botão entitulado *DDSOntADT*, que abre um painel representado pela aba *Get DDSOntADT Info*, idêntico ao painel aberto pelo botão *Get DDSOntADT Info* do módulo da entidade reguladora explicado na seção *sec:reguladora*. Assim, o painel *Search for DDSOntWs Definition* da entidade provedora suporta também a atividade 3 apresentada no diagrama da Figura 6.7.

Ainda dissertando a respeito do painel que apóia a atividade 1 da entidade provedora, a área mais à direita mostra os resultados da consulta por um tModel relacionado a um serviço Web da DDSOntWs. O resultado da consulta retorna informações como o nome do serviço Web, uma descrição deste contendo os parâmetros de entrada e de saída, um resumo específico do serviço descrevendo os tipos de dados utilizados, o identificador do TModel, e por fim o endereço de acesso ao documento WSDL definido pela entidade reguladora. Todas estas informações relativas a um tModel são necessárias para que um serviço da DDSOntWs seja efetivamente implementado pela entidade provedora, ou então utilizado, como é o caso do getUMLSInfo,

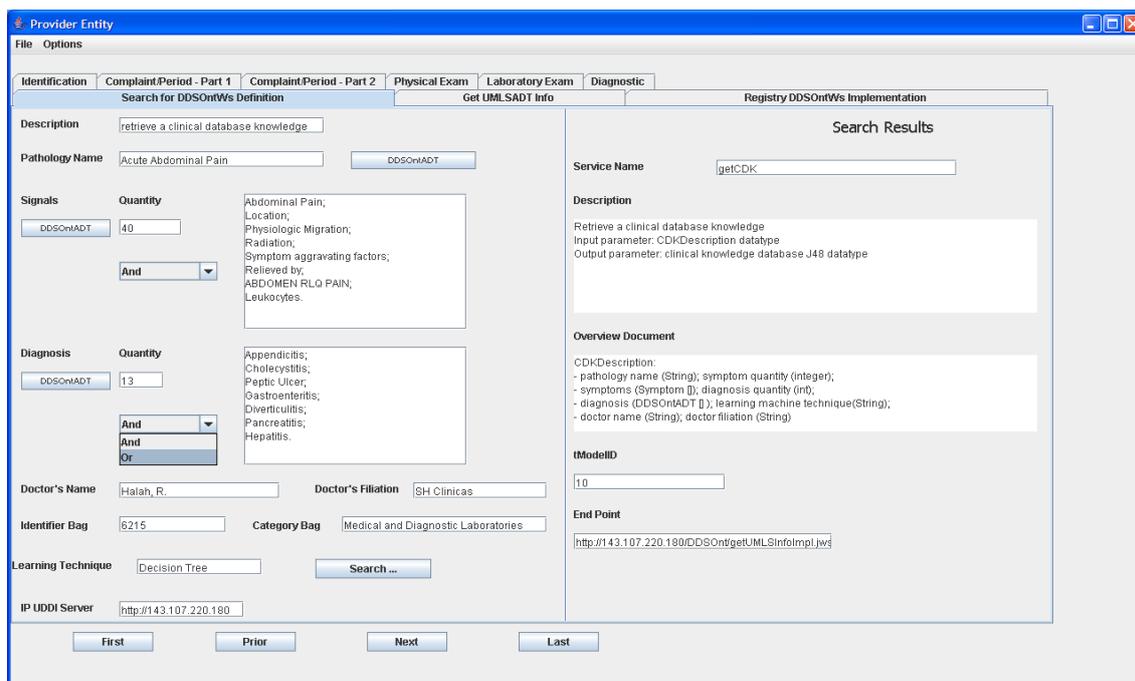


Figura 6.8: Tela para apoiar a busca por TModels no servidor UDDI

no painel representado pela aba *Get DDSOntADT Info*. Caso exista mais de um tModel que se encaixe na busca, estes podem ser navegados pelos botões na parte de baixo do painel.

Tendo encontrado a definição os tModels de seu interesse no servidor UDDI, a entidade provedora pode então implementar os serviços da DDSOnt, especificamente o createCDK, o getDDS e o getCDK. Particularmente em relação aos detalhes de implementação destes serviços, foram utilizadas a classe interface DDSOntWS, apresentada na seção 5.3.3, e a API JDDSOnt, apresentada na seção 5.3.2. Com relação específica à implementação dos serviços createCDK e getDDS, foi utilizada a API WEKA, apresentada na seção 5.3.3 para a geração e inferência das base de conhecimento.

O acionamento da aba *Registry DDSOntWs Implementation* da entidade provedora disponibiliza um painel idêntico ao mesmo painel da entidade reguladora. Este painel tem como objetivo apoiar a atividade 3 da entidade provedora, que é o registro no servidor UDDI das implementações dos serviços createCDK, getCDK e getDDS. Assim, estes podem ser utilizados pela entidade requisitante, explicada na próxima seção. O registro das implementações foram realizados utilizando-se o pacote JAXR (*Java API for XML Registries*).

A atividade 4 de competência de uma entidade provedora envolve a criação de um conjunto de treinamento composto por casos clínicos. O nome conjunto de treinamento é utilizado neste contexto pois os casos clínicos criados gerarão futuramente uma base de conhecimento. Sendo assim, a máquina utiliza-se destes casos clínicos para poder treinar e aprender, e assim gerar uma base de conhecimento. Cada um dos casos clínicos é formado por um conjunto de sinais e um diagnóstico. Para que o usuário possa preencher os sinais, 5 abas foram criadas no módulo de apoio à entidade provedora. A primeira aba, *Patient Identification*, ilustrada pela Figura 6.9, permite o usuário entrar com dados relativos à identificação do paciente. A segunda aba, *Complaint/Period - Part 1*, ilustrada pela Figura 6.10, permite o usuário entrar com uma primeira parte dos sinais clínicos relacionados a queixa/duração do paci-

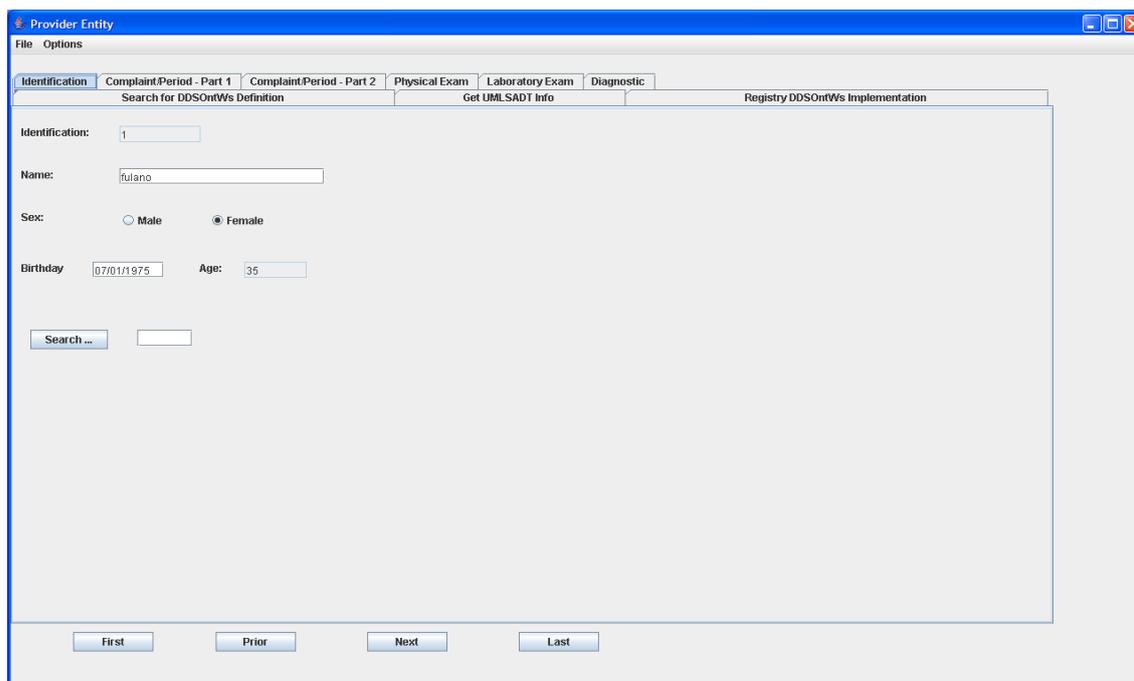


Figura 6.9: Tela para apoiar a entrada de dados relativos a identificação do paciente

ente. A terceira aba, *Complaint/Period - Part 2*, ilustrada pela Figura 6.11, permite o usuário entrar com a segunda parte dos sinais clínicos relacionados a queixa/duração do paciente. A quarta aba, *Physical Exam*, ilustrada pela Figura 6.12, permite o usuário entrar com dados relativos aos exames físicos do paciente. Finalmente, na quinta aba, *Laboratory Exam*, ilustrada pela Figura 6.13, permite o usuário entrar com dados relacionados aos exames laboratoriais do paciente.

É importante observar que os termos utilizados no módulo da entidade provedora são provenientes do sistema de terminologia SNOMED CT. Para a obtenção dos valores da UMLS MetaThesaurus referentes ao tipo semântico *DDSOntADT*, foi utilizado o serviço *getUMLSInfo*. Ainda, que a entrada dos dados clínicos é estruturada, como foi observado no início deste capítulo.

Para finalizar o apoio à atividade de criar um conjunto de treinamento, foi criado a aba *Diagnostic*. Esta aba ativa um painel que permite o usuário escolher um diagnóstico, como mostra a Figura 6.14, e assim finalizar seu preenchimento de um caso clínico. Com o dados do caso clínico preenchidos totalmente, o usuário pode clicar no botão com título *Save Clinical Case*, que salva os dados em banco de dados local. Assim, pode-se repetir o processo de preenchimento novamente e quantas vezes forem necessárias. Neste estudo de caso, foram criados 211 casos clínicos referente a dor abdominal aguda.

Este último painel em discussão dá apoio ainda às atividades 5 e 6 da entidade provedora. Para que o usuário possa criar uma base de conhecimento a partir dos casos clínicos salvos, é preciso clicar no botão entitulado *Call createCDK*, que vai executar o serviço Web *createCDK*, implementado pela própria entidade provedora.

Eventualmente, a entidade provedora pode querer solicitar apoio à decisão diagnóstica de um caso clínico atual e duvidoso, ou mesmo testar a sua própria base de conhecimento criada. Para isto, é preciso clicar no botão com título *Call getDDS*,

The screenshot shows a software window titled "Provider Entity" with a menu bar (File, Options) and several tabs: Identification, Complaint/Period - Part 1 (selected), Complaint/Period - Part 2, Physical Exam, Laboratory Exam, and Diagnostic. Below the tabs are three sub-sections: "Search for DDSOntWs Definition", "Get UMLSADT Info", and "Registry DDSOntWs Implementation". The main form area contains the following fields:

- Abdominal pain: Yes No
- Location:
- Migration: yes No Initial Location:
- Character Initial:
- Character Changes: Yes No
- Character Final:
- Severity:
- Severity Changes:
- Radiation: Yes No
- Radiation Location:
- Exacerbating Factors:

At the bottom of the form are four navigation buttons: First, Prior, Next, and Last.

Figura 6.10: Tela para apoiar a entrada de dados relativos a primeira parte da queixa/duração do paciente

The screenshot shows the same "Provider Entity" window, but with the "Complaint/Period - Part 2" tab selected. The form area contains the following fields:

- Alleviating Factors:
- Bowel Movements:
- jaundice: Yes No
- Alcoholism: Yes No
- Anorexia: Yes No
- Vomiting: Yes No
- Vomiting Character:
- Last Menstruation:
- Dyspareunia: Yes No
- Vaginal discharge: Yes No

At the bottom of the form are four navigation buttons: First, Prior, Next, and Last.

Figura 6.11: Tela para apoiar a entrada de dados relativos a segunda parte da queixa/duração do paciente

The screenshot shows the 'Provider Entity' application window with the 'Physical Exam' tab selected. The form contains the following fields and options:

- Tachycardia: Yes No
- abdominal Auscultation:
- Hydration:
- distention: Yes No
- temperature:
- murphy: Yes No
- bloomberg: Yes No
- Tenderness in RLQ: yes no
- Rebound Pain: yes no

Navigation buttons at the bottom: First, Prior, Next, Last.

Figura 6.12: Tela para apoiar a entrada de dados relativos aos exames físicos do paciente

The screenshot shows the 'Provider Entity' application window with the 'Laboratory Exam' tab selected. The form contains the following fields and options:

- WBC:
- Shift to left: yes no
- Hematuria: yes no
- Anytase:
- Alkaline phosphatase:
- SGOT:
- SGPT:

Navigation buttons at the bottom: First, Prior, Next, Last.

Figura 6.13: Tela para apoiar a entrada de dados relativos aos exames laboratoriais do paciente

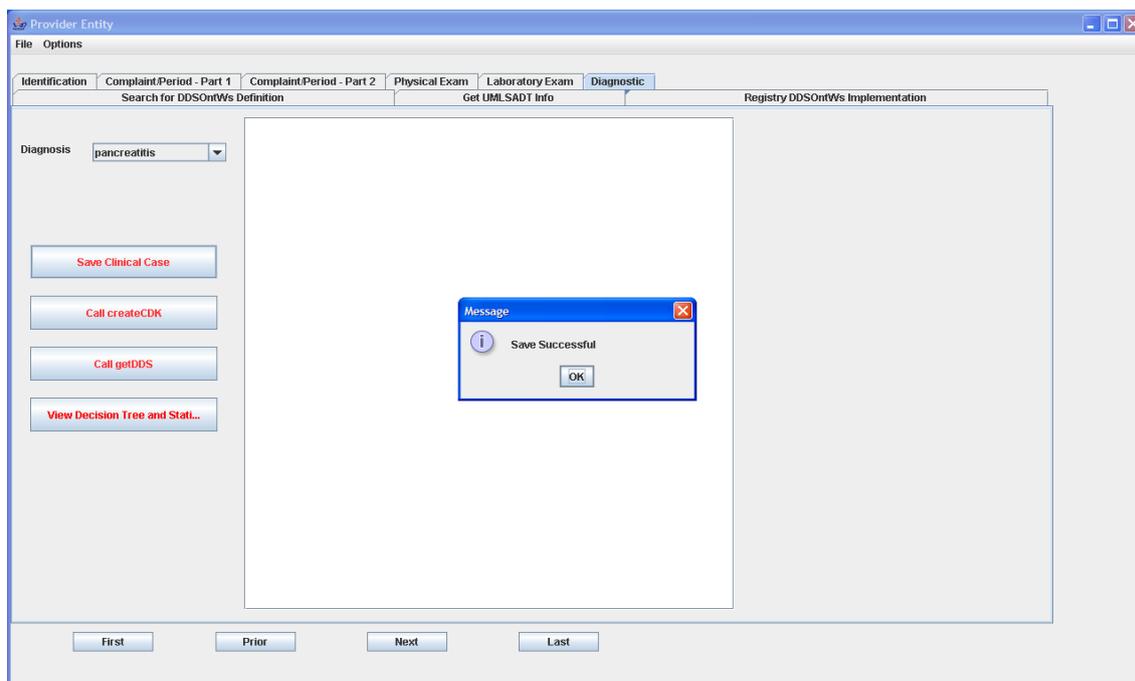


Figura 6.14: Tela para apoiar a entrada de dados relativos ao diagnóstico do paciente

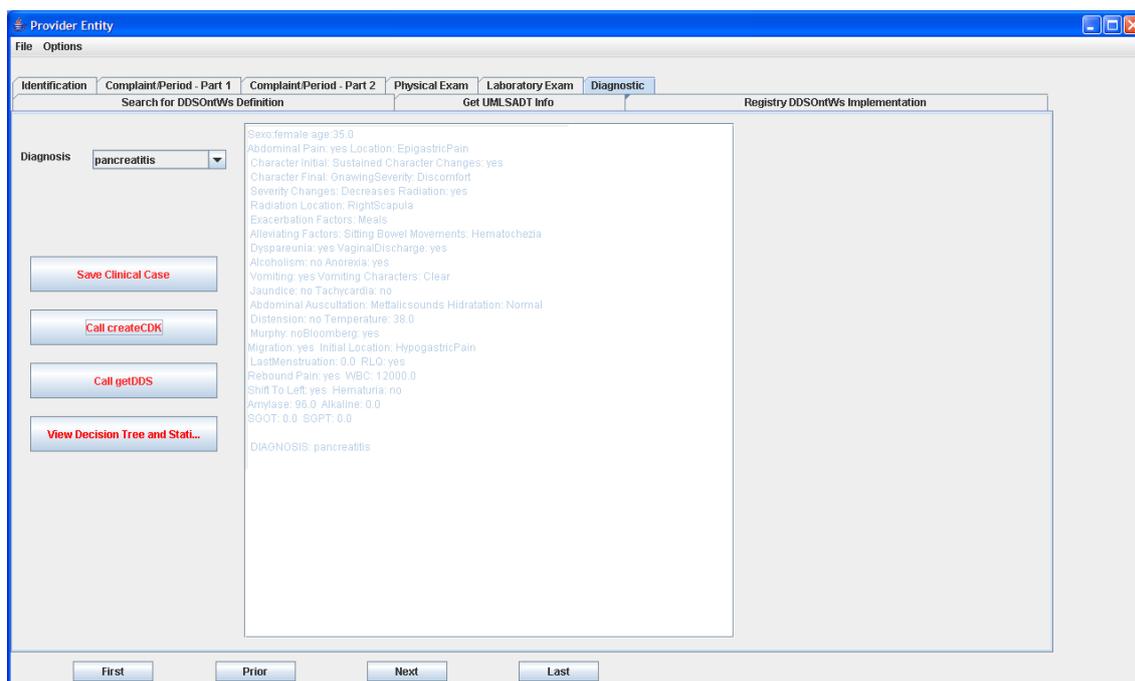


Figura 6.15: Tela para apoiar a consulta por uma resposta diagnóstica a partir de base de conhecimento local

que executa o serviço Web getDDS, também implementado pela própria entidade. A ação deste botão sempre solicita um apoio à decisão diagnóstica a partir da base de conhecimento desta própria entidade. A Figura 6.15 ilustra esta situação, mostrando que a resposta diagnóstica do caso duvidoso é *pancreatitis*.

Finalmente, caso o médico deseja visualizar a árvore de decisão criada a partir

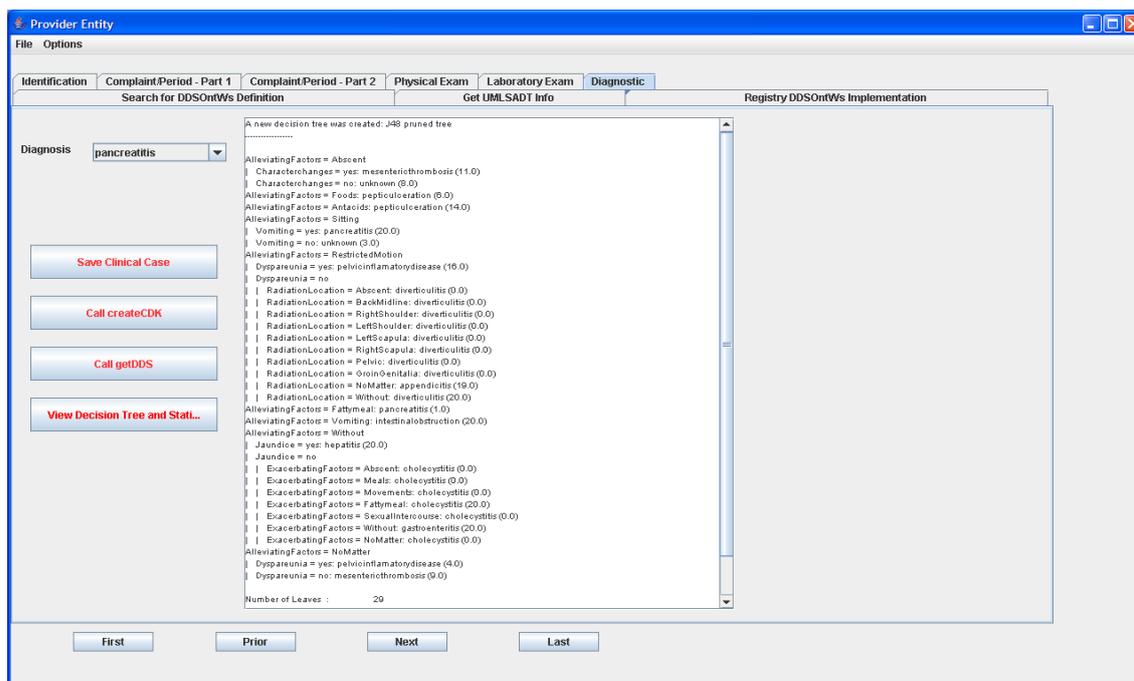


Figura 6.16: Tela para apoiar a consulta das estatísticas e da árvore de decisão da base de conhecimento local

de sua base de conhecimento para se chegar ao diagnóstico sugerido pela sua base de conhecimento, juntamente com estatísticas de percentual de acertos, é preciso que ele pressione o botão *View Decision Tree and Statistics*. A técnica *10-fold cross-validation* foi utilizada para validar a base de conhecimento, e o resultado foi que 98.75% dos casos de teste foram classificados corretamente. Estas características podem ser obtidas da própria base de conhecimento e a implementação disto é de responsabilidade da própria entidade, não sendo isto disponível em forma de serviço Web pela DDSOntWs. A Figura 6.16 ilustra o usuário visualizando uma árvore de decisão com estatísticas da base de conhecimento.

6.5 Módulo de Apoio para Médicos Requisitantes

O terceiro e último módulo criado no estudo de caso tem como objetivo apoiar médicos requisitantes de atividades de apoio diagnóstico, especificamente relacionadas à patologia de dor abdominal aguda. Esta entidade requisitante pode ser representada na realidade por qualquer médico iniciante ou não, que deseja utilizar serviços relacionados a sugestão diagnóstica, e recuperação de bases de conhecimento para apoio à decisão diagnóstica. Neste estudo de caso, a entidade requisitante é representada pelo próprio Dr. Halah, que se baseou em seus casos clínicos reais de sua clínica. A Figura 6.17 ilustra os requisitos e funcionalidades deste módulo. Pela figura, pode-se perceber que o médico requisitante pode:

1. Consultar o servidor UDDI a fim de obter a definição dos tModels relacionados aos serviços da DDSOntWs. Com estas respostas, a entidade requisitante pode fazer uso dos serviços *getCDK*, *getDDS*, e *createCDK* implementados pelas entidades provedoras, e especificamente do *getUMLSInfo*, implementado pela própria entidade reguladora;

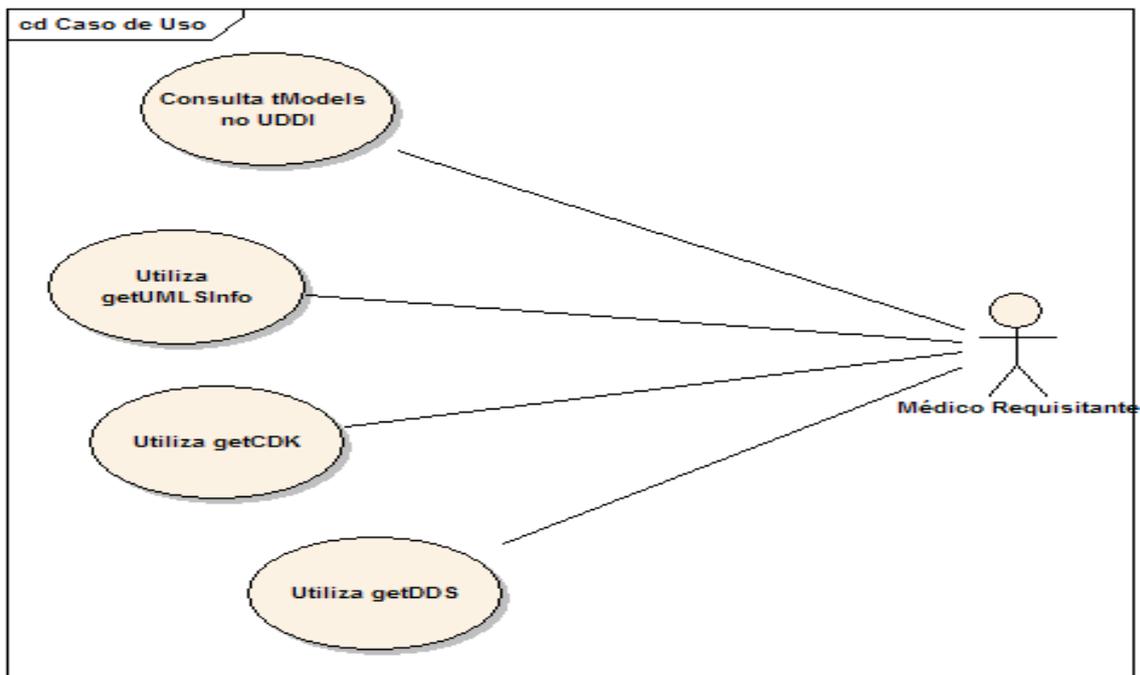


Figura 6.17: Diagrama de Caso de Uso para Módulo de Apoio ao Médico

2. Utilizar o serviço Web `getUMLSInfo` da entidade reguladora a fim de recuperar os valores da UMLS MetaThesaurus e compor a estrutura de dados `DDSOntADT`;
3. Utilizar o serviço Web `getCDK` para obter uma base de conhecimento relacionada à sua patologia ou valores de sinais e diagnósticos;
4. Utilizar o serviço Web `getDDS` para ser apoiado durante a decisão diagnóstica de caso clínico duvidoso a partir da base de conhecimento obtida no item anterior.

De modo a suportar a realização das atividades de 1 a 4 de uma entidade requisitante, foi desenvolvido um módulo de software. Este módulo é composto de 9 painéis sobrepostos e selecionáveis por abas intituladas *Search for DDSOntWs Definitions*, *Get DDSOntADT Info*, *Identification*, *Complaint/Period – Part 1*, *Complaint/Period – Part 2*, *Physical Exam*, *Laboratory Exam*, *Diagnostic* e *Search for DDSOntWs Implementation*. Estes painéis foram utilizados pelo Dr. Halah, com o suporte técnico do autor desta tese. Estes são apresentados a seguir.

O acionamento da aba *Search for DDSOntWs Definitions* disponibiliza um painel contendo recursos para apoiar a atividade 1 da entidade requisitante. Este painel é idêntico ao mesmo painel da entidade provedora, e permite procurar pela definição de um `tModel` em um servidor UDDI. Como resposta, este painel recebe o nome do serviço Web, uma descrição deste contendo os parâmetros de entrada e de saída, um resumo específico do serviço descrevendo os tipos de dados utilizados, o identificador do `TModel`, e por fim o endereço de acesso ao documento WSDL definido pela entidade reguladora. Estas informações são suficientes para que a entidade requisitante possa utilizar os serviços da `DDSOntWs`.

Portanto, pode-se dizer que este painel apóia a atividade 1 da entidade requisitante. E como este painel possibilita a ativação do painel *Get DDSOntADT Info*, como mostrado na descrição do módulo da entidade provedora, também apóia a atividade 2 da entidade requisitante.

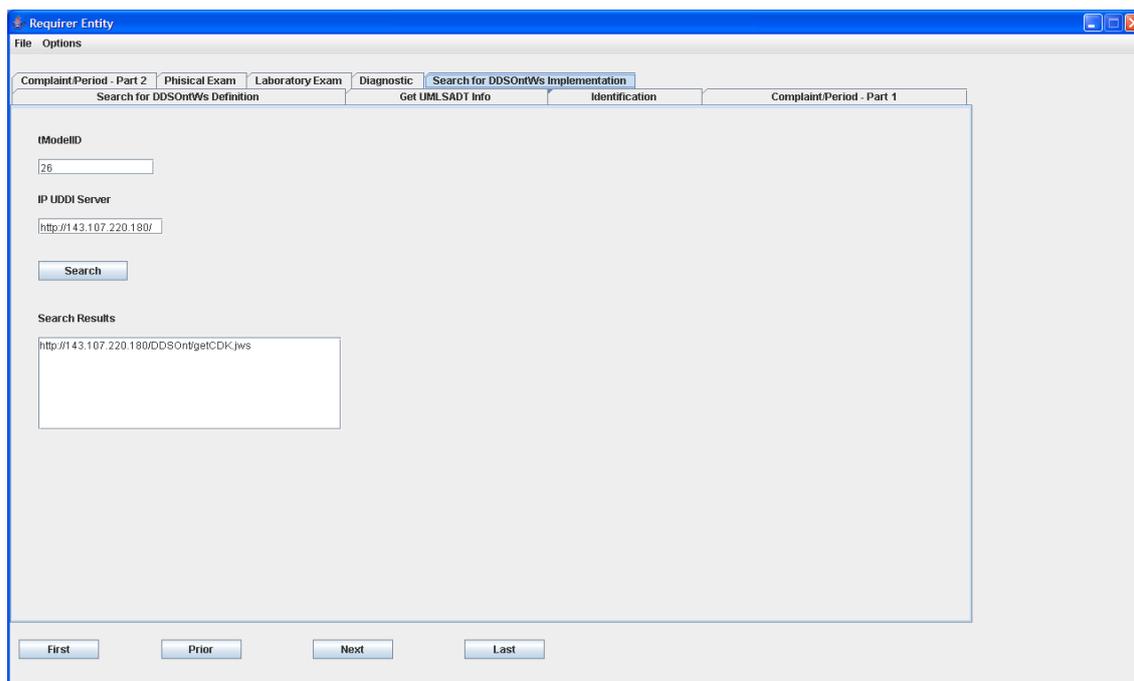


Figura 6.18: Tela para apoiar a busca por implementações de um *tModel*

Tendo encontrado a definição dos *tModels* de seu interesse no servidor UDDI, a entidade requisitante precisa agora de encontrar as implementações destes *tModels*, para então fazer uso dos serviços da DDSOntWs. Para isto, é necessário que o usuário se utilize do painel ativado pela aba *Search for DDSOntWs Implementation*, ilustrado pela Figura 6.18. Este painel permite o usuário entrar com um identificador de *tModel* encontrado no painel de busca por definição de *tModels*, e um endereço IP do servidor UDDI, e obter como resposta os endereços IPs de todas as entidades provedoras que implementaram o *tModel* especificado na busca. Na Figura 6.18, esta-se buscando por implementações do ID igual a 26, que neste estudo de caso é o identificador do serviço Web *getDDS*. Da mesma forma que este painel foi utilizado para procurar as implementações de *getDDS*, este também foi utilizado para a busca das implementações de *getCDK* e *getUMLSInfo*.

A busca por definições de *tModels* e por implementações destes foi realizada utilizando-se o pacote JAXR (*Java API for XML Registries*).

Estando apto a utilizar os serviços da DDSOntWs, a entidade requisitante pode utilizar-se das atividades de apoio à decisão diagnóstica. Assim, o usuário pode fazer uso das abas *Identification*, *Complaint/Period – Part 1*, *Complaint/Period – Part 2*, *Physical Exam* e *Laboratory Exam* para preencher as informações de um caso clínico duvidoso, e navegar para a aba *Diagnostic*. Esta última aba ativa um painel ilustrado na Figura 6.19. Por este painel, o usuário pode clicar no botão entitulado *Call GetCDK* para obter uma base de conhecimento de uma entidade provedora, e então finalmente ser apoiado para uma decisão diagnóstica clicando no botão com título *Call GetDDS*. A Figura 6.20 ilustra uma resposta diagnóstica obtida a partir da base de conhecimento importada de um outro médico, especificamente neste caso, da entidade provedora.

A entidade requisitante pode ainda, a partir de seu módulo, criar seus próprios casos clínicos e laboratoriais de acordo com sua possível experiência clínica. Neste estudo de caso, a entidade requisitante criou 26 casos clínicos a partir das experiências clínicas reais do Dr. Halah em sua clínica. De posse dos casos clínicos, a entidade

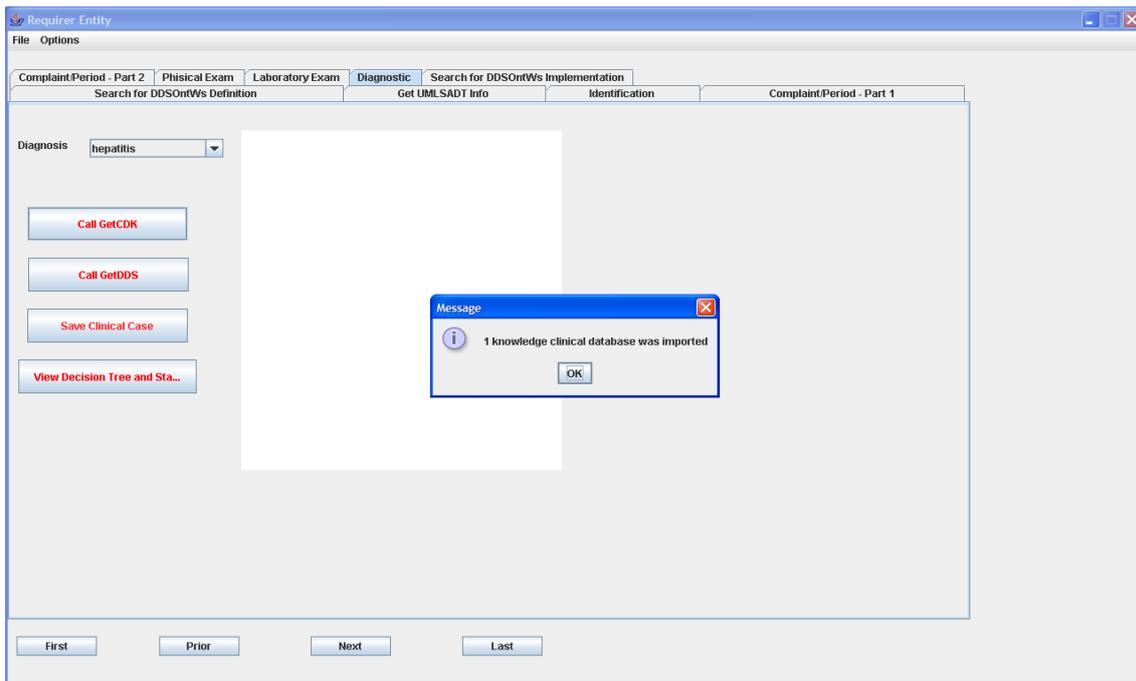


Figura 6.19: Tela para apoiar a busca por um base de conhecimento

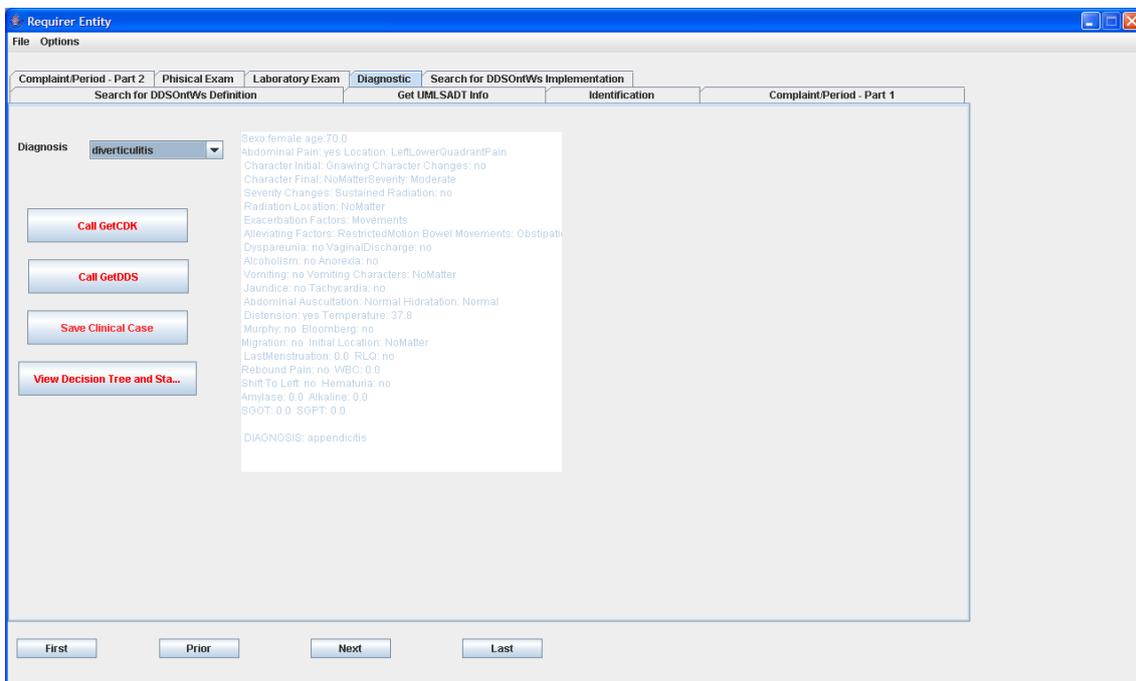


Figura 6.20: Tela para apoiar a consulta por uma resposta diagnóstica

requisitante fez uso do serviço Web *createCDK* para criar sua própria base de conhecimento local. A técnica *10-fold cross-validation* foi utilizada para validar esta base, e o resultado foi que 97.55% dos casos de teste foram classificados corretamente. De posse desta sua base de conhecimento, o médico pode utilizar o serviço *getDDS* para comparar o resultado da resposta diagnóstica apresentada anteriormente (a partir da base de conhecimento da entidade provedora) com o resultado de uma resposta diagnóstica a partir da sua própria base. Ainda, com sucessivas utilizações dos serviços *getCDK* e *getDDS*, é possível o médico fazer comparações entre respostas diagnósticas obtidas de bases de conhecimento de várias entidades provedoras.

Eventualmente, a entidade requisitante pode tornar-se entidade provedora. Para isto, é necessário que a primeira implemente o serviço Web *getCDK* e registre esta implementação no servidor UDDI. Desta maneira, outros médicos poderão encontrar e utilizar sua base de conhecimento compartilhada com a comunidade médica.

É importante observar que os termos utilizados no módulo da entidade requisitante são provenientes do sistema de terminologia ICD-9. Para a obtenção dos valores da UMLS MetaThesaurus referentes ao tipo semântico *DDSOntADT*, foi utilizado o serviço *getUMLSInfo*. Ainda, que a entrada dos dados clínicos é estruturada, como foi observado no início deste capítulo.

6.6 Considerações Finais

Este Capítulo apresentou um estudo de caso que faz uso da *DDSOnt*, *JDDSOnt* e *DDSOntWs* para realizar atividades relacionadas a apoiar a decisão clínica da síndrome de dor abdominal aguda em ambiente colaborativo. Mostrou-se o funcionamento, e a finalidade de 3 módulos que são utilizados por 3 atores diferentes: uma entidade reguladora dos serviços da *DDSOntWs*, uma entidade provedora de base de conhecimento e de apoio à decisão diagnóstica, e uma entidade requisitante, que procura bases de conhecimento e solicita por respostas diagnósticas à entidade requisitante.

As entidades reguladoras, provedoras e requisitantes neste estudo de caso, como visto, foram representadas por apenas um médico, o Dr. Halah. O médico utilizou os módulos de software, que funcionaram adequadamente de acordo com os propósitos definidos. Entretanto, novos testes podem e precisam ser realizados com outros médicos e com novas síndromes diferentes da dor abdominal aguda. São potenciais usuários dos módulos de software desenvolvidos: médicos experientes ou juntas médicas especializadas, médicos iniciantes, médicos distantes de centros médicos de referência, graduandos em medicina, e médicos residentes. Estes novos testes são melhor discutidos na seção de trabalhos futuros presente no próximo capítulo.

A versão da UMLS utilizada no estudo de caso é essencialmente escrita na língua inglesa, e não possui uma tradução para o português dos termos presentes nos sistemas de terminologia. Assim, foi necessário a aplicação no estudo de caso da língua inglesa. Entretanto, existe um vocabulário em português na UMLS, como parte de um conjunto maior composto por vários vocabulários. Em função disto, os números identificadores de conceitos, de átomos, e de termos léxicos que são referências a conceitos, átomos e termos léxicos, respectivamente, podem ser traduzidos para o português.

O próximo e último Capítulo tem como objetivo tecer as considerações finais deste trabalho, e apresentar alguns trabalhos relacionados e os futuros trabalhos.

Discussão dos Resultados

7.1 Considerações Finais

Este trabalho apresentou os resultados da investigação dos problemas de interoperabilidade semântica e técnica durante a troca de informações para apoiar a decisão diagnóstica em ambiente colaborativo na Internet. Foi proposto uma ontologia original denominada DDSOnt desenvolvida para a definição de um conjunto de tipos semânticos necessários para a criação de uma estrutura que suporte o compartilhamento de bases de conhecimento de modo que possam ser utilizadas por médicos no apoio à decisão diagnóstica. Esta ontologia estende a ontologia UMLS SN, reutilizando os tipos semânticos Entidade e Evento, com objetivo de promover a compatibilidade ontológica.

A escolha de se estender a UMLS SN em detrimento de outras ontologias associadas é justificada pela forma em que esta foi construída. A UMLS SN criou seus tipos semânticos a partir de conceitos provenientes de mais de 100 sistemas de terminologias, enquanto que as demais ontologias criaram seus tipos a partir da experiência dos seus autores. Espera-se portanto, que a compatibilidade ontológica seja mais frequentemente conseguida quando se utiliza a UMLS SN, principalmente em um cenário que promove a compatibilidade terminológica com a *UMLS MetaThesaurus*.

Neste contexto, a proposta deste trabalho de criar um ambiente colaborativo e interoperável para apoiar a decisão diagnóstica torna-se dependente e relacionada ao projeto da UMLS. Um aspecto positivo deste relacionamento é que a proposta se aproveita da experiência e trabalho realizado pela NLM no desenvolvimento da base de dados UMLSKS desde 2002, procurando sempre unificar conceitos e tipos semânticos da área médica. Por outro lado, a solução proposta exige das aplicações o uso de um sistema de terminologia que seja contemplado pela tesouro da UMLS. Outro aspecto negativo são as diferenças de granularidade presentes na UMLS SN. Ainda, a UMLS, além de apresentar alguns erros nas versões mais atuais ¹, faz com frequência atualizações de suas versões, o que dificulta a sua adaptação pelos sistemas presentes

¹http://www.nlm.nih.gov/research/umls/release_notes.html

nas instituições em saúde.

As atividades de confecção da ontologia DDSOnt com a metodologia 101 apresentadas no Capítulo 5, bem como o uso desta no estudo de caso para construção e compartilhamento de bases de conhecimento, e consultas a respostas diagnósticas na patologia de dor abdominal aguda, foram realizadas com êxito. Isto se justifica uma vez que estas atividades mostraram-se satisfatórias para as situações de colaboração médica propostas no estudo de caso. Assim, o objetivo de se criar um ambiente colaborativo com a DDSOnt, a JDDSOnt e a DDSOntWs que apóie a decisão diagnóstica e promova a compatibilidade ontológica foi alcançado.

De modo a suportar que softwares clínicos pudessem fazer uso da DDSOnt para a criação e disponibilização de bases de conhecimento e consulta a respostas diagnósticas diversas, a API JDDSOnt foi projetada para a definição e construção de documentos clínicos estruturados da Web Semântica em conformidade com a DDSOnt. Estes documentos são utilizados pelos serviços da DDSOntWs. Este último permite que a construção e o compartilhamento de bases de conhecimento, assim como a consulta por respostas diagnósticas, possam ocorrer em ambiente compartilhado e distribuído, orientado a serviços, e independente de plataforma, promovendo assim a compatibilidade técnica. De modo a promover a compatibilidade terminológica entre as aplicações que façam uso da DDSOnt, o serviço getUMLSInfo faz uso da UMLSKS API para relacionar termos de diversos vocabulários médicos com os conceitos da *UMLS MetaThesaurus*.

As atividades de construção dos softwares JDDSOnt e DDSOnWs, assim como o uso destes no estudo de caso apresentado foram realizadas com êxito, uma vez que estes puderam ser utilizados pelos módulos de apoio às entidades reguladora, provedora e requisitante. Assim, o objetivo de se criar um ambiente que promova a compatibilidade terminológica e tecnológica especificamente para apoiar a decisão diagnóstica foi alcançado.

7.2 Trabalhos Relacionados

Este projeto investigou e estudou alguns trabalhos na literatura que propõem soluções que permitam que CDSSs recebam informações provenientes de sistemas clínicos e hospitalares heterogêneos. Estes são apresentados a seguir.

Bilykh e colaboradores (Bilykh et al. 2006) apresentam um CDSS nomeado EGADSS, de padrão aberto, orientado a componentes, e possível de ser utilizado por qualquer EMR. A justificativa de se criar o EGADSS, segundo os autores, é que os CDSSs são, em sua maioria, proprietários, fortemente acoplados com EMR específicos, e caros de serem desenvolvidos. Assim, os autores propõem dois modelos de documentos baseados no CDA HL7 como formato para troca de informações clínicas entre o EGADSS e os EMRs: um para sumários do paciente e outro para recomendações clínicas ao paciente. As regras médicas das bases de conhecimento do EGADSS são representadas utilizando-se da linguagem Arden Syntax e consultadas com a linguagem XML XPath².

O trabalho de Bilykh e co-autores relaciona-se com este trabalho pois ambos buscam criar uma proposta que permite alimentar um CDSS com dados provenientes de EMRs heterogêneos. Enquanto este trabalho propõe uma solução baseada nos sis-

²<http://www.w3.org/TR/xpath20/>.

temas da UMLS para promover interoperabilidade, o trabalho estudado escolheu as tecnologias relacionadas ao projeto HL7. Especificamente, este trabalho promove a compatibilidade ontológica com a ontologia DDSOnt extendida da UMLS SN, que permite a criação de uma estrutura de documentos que suporte a manipulação de bases de conhecimento de modo que possam ser utilizadas por médicos no apoio à decisão clínica. Já o trabalho de Bilykh citeBil2006 usa a própria ontologia RIM do HL7 para relacionar as informações semânticas aos dados contidos nos documentos clínicos propostos a partir do CDA HL7. Pode-se então comparar a estrutura suportada pela DDSOnt com os modelos de documentos relacionados ao EGADSS.

Ainda comparando os dois trabalhos, este projeto promove a compatibilidade terminológica com a *UMLS MetaThesaurus* e o trabalho de Bilykh citeBil2006 não aborda e não propõe uma solução aos problemas terminológicos. Finalmente, ambos os trabalhos propõe promover a compatibilidade técnica com as especificações relacionadas a serviços Web, apesar do trabalho de Bilykh citeBil2006 não discutir uma solução para disponibilização e busca dos serviços, e conseqüentemente, da possibilidade de compartilhamento e recuperação de bases de conhecimento. Isto efetivamente permite criar um ambiente colaborativo de apoio a decisão diagnóstica, como demonstrado neste trabalho.

Achour e colaboradores (Achour et al. 2001) apresentam uma ferramenta de aquisição de conhecimento baseada em regras e na UMLS que tem como objetivo facilitar a criação e a manutenção de uma base de conhecimento por médicos, e o seu compartilhamento e reuso por profissionais de outras instituições. Para isso, a ferramenta oferece ao seu usuário a possibilidade de construir uma ontologia de domínio a partir de um mecanismo navegacional sobre a rede semântica da UMLS SN. Com a ontologia construída, a base de conhecimento pode ser criada utilizando termos da *UMLS MetaThesaurus*. Posteriormente, as regras de decisão médica da base de conhecimento são representadas por módulos lógicos médicos (MLMs) utilizando a linguagem Arden Syntax, que segundo os autores, facilita o reuso e o compartilhamento desta. Um estudo de caso é apresentado com informações associadas a transfusão de sangue.

O trabalho de Achour e co-autores relaciona-se com este trabalho pois ambos procuram criar um sistema de apoio à decisão clínica preocupados com o reuso e o compartilhamento da base de conhecimento deste. Entretanto, enquanto o trabalho estudado explora especificamente as atividades de criação e representação da base de conhecimento de um CDSS, este trabalho verifica os problemas de interoperabilidade dos dados durante o apoio à decisão diagnóstica em ambiente colaborativo na Web, analisa as possibilidades de decisão diagnóstica com bases de conhecimento compartilhadas, e faz uma proposta que promove a compatibilidade técnica, terminológica e ontológica. Em função disto, Achour e co-autores permitem a criação de ontologias de domínios diferentes para cada patologia sendo estudada a partir da UMLS SN, enquanto que este trabalho propõe a DDSOnt que estende a UMLS SN especificamente para suportar o apoio à decisão diagnóstica em ambiente colaborativo. Outro aspecto comparável entre os dois projetos é a técnica para representação da base de conhecimento. O primeiro utiliza-se de MLMs e a linguagem Arden Syntax, enquanto que este utiliza-se da técnica de árvore de decisão.

Kazemzadeh e Sartipi (Kazemzadeh 2005) produzem um *framework* que promove a interoperabilidade de dados clínicos e de conhecimento médico provenientes de EMRs heterogêneos. De acordo com os autores, a interoperabilidade dos dados é conseguida com o uso das tecnologias RIM e CDA relacionadas ao projeto HL7. E a interoperabilidade do conhecimento clínico é promovida com a criação de uma linguagem XML nomeada PMML, a utilização das técnicas Arden Syntax e MLM, e também do uso da técnica de guia de conduta GLIF. Especificamente, o trabalho estudado busca promo-

ver a compatibilidade ontológica com a própria ontologia RIM do HL7 para relacionar as informações semânticas aos dados contidos nos documentos clínicos propostos a partir do CDA HL7. Entretanto, os autores não propõem uma solução aos problemas de compatibilidade terminológica e técnica, assim como não exploram as diversas possibilidades decorrentes de um ambiente que permite apoiar a decisão diagnóstica em ambiente colaborativo na Web.

Outro trabalho relacionado com este é o apresentado por Pisa e colaboradores em (Pisa et al. 2004). Neste trabalho é apresentado o projeto Lepidus R3, que tem por objetivo desenvolver, implementar e avaliar uma arquitetura distribuída que permite oferecer interoperabilidade técnica para o sistema Lepidus através do uso de tecnologias de Serviços Web baseados em SOAP/XML. O sistema Lepidus é um sistema de apoio à decisão diagnóstica baseado em redes neurais e que utiliza uma forma de representação dos sinais clínicos através de funções senoidais amortecidas (Silva and Roque 2000). Segundo os autores, com esta arquitetura, sistemas de instituições em saúde legados podem fazer uso do Lepidus na Internet para auxiliar seus usuários no apoio diagnóstico.

O trabalho de Pisa e co-autores relaciona-se com este trabalho pois ambos permitem o uso e acesso de um CDSS através da tecnologia de serviços Web. Entretanto, enquanto o primeiro oferece o apoio à decisão diagnóstica a partir de casos clínicos individuais, este trabalho propõe o apoio a decisão diagnóstica colaborativa, oferecendo aos médicos serviços de criação, compartilhamento e consulta de bases de conhecimento em ambiente que promove ainda a compatibilidade terminológica e ontológica.

Foram feitas ainda comparações da proposta deste trabalho com as tecnologias relacionadas ao projeto HL7 e Galen. Os resultados são apresentados a seguir.

A solução proposta neste trabalho proporciona às aplicações clínicas que a utilizam a geração de documentos clínicos com conteúdo que possui uma associação semântica à ontologia DDSOnt. Por exemplo, um documento gerado pela JDDSOnt possui conteúdo formado por termos que são semanticamente relacionados ao tipo semântico *DDSOntADT*, *Signal* ou *ClinicalDiagnosisEvaluation*. Ainda, os termos clínicos presentes no documento estão relacionados aos conceitos do sistema de terminologia da *UMLS MetaThesaurus* para promover a compatibilidade terminológica. Por último, estes documentos gerados devem estar de acordo com uma estrutura estabelecida por um XML Schema da DDSOnt.

Os recursos da DDSOnt apresentados anteriormente estão relacionados com parte do funcionamento das tecnologias disponibilizadas pela Open Galen e pelo HL7. Na tecnologia Galen, a linguagem de representação GRAIL é utilizada para modelar e formalizar a ontologia CRM, e ainda criar documentos clínicos com conteúdo relacionado semanticamente à CRM. Entretanto, não está previsto no projeto Galen que um termo clínico presente em um documento esteja relacionado com algum sistema de terminologia que promova a compatibilidade terminológica, como a *UMLS MetaThesaurus*. Estes documentos criados possuem uma estrutura e organização estabelecida conforme as regras léxicas e sintáticas formais da linguagem GRAIL ³.

Assim como a DDSOnt possui uma representação semântica através de documentos OWL, a ontologia CRM também detém documentos OWL em conjunto aos projetos CO-ODE e HyOntUse (Rector et al. 2007). Ao contrário da GRAIL ⁴, a JDDSOnt possibilita que aplicações consigam criar documentos clínicos em XML e RDF, suportando e promovendo assim a interoperabilidade sintática entre as aplicações médicas. Por

³<http://www.opengalen.org/faq/faq3.html>

⁴<http://www.opengalen.org/faq/faq5.html>

fim, de acordo com o estudo realizado na GRAIL, não existe uma regra formal específica para transferir documentos que suportem o apoio à decisão diagnóstica em ambiente colaborativo como propõe a DDSOnt.

Em relação a tecnologia HL7, documentos clínicos e administrativos podem ser construídos a partir de um esquema originado da arquitetura CDA. O conteúdo destes documentos possuem associação semântica com a ontologia RIM. Os termos clínicos presentes nos documentos estão relacionados com os sistemas de terminologia contemplados no HL7 Vocabulary, apesar de não ser citado em ⁵ que este promova a compatibilidade terminológica.

A ontologia HL7 possui documentos OWL apoiados pelo projeto Protegé (Tu 2007) para representar sua ontologia RIM. Apesar da arquitetura CDA possuir um documento XML Schema que formaliza e valida se os documentos clínicos estão em conformidade com este, HL7 não possui um mecanismo que apóie as aplicações na geração automática destes. Por fim, o HL7 possui um padrão para construção de modelos de documentos clínicos, chamados de *Templates*. Este padrão define um formato e uma estrutura de documentos clínicos para diversas situações clínicas, como relatos ou notas de histórico de uma condição física, de uma operação, de um procedimento, ou da conta de um paciente. Apesar destes formatos, em primeiro momento, mostrarem-se contribuir significativamente para o compartilhamento de dados médicos, a versão final destes documentos ainda encontram-se em fase de votação e aceitação (*ballot*) pela comunidade HL7, como explicado na seção 4.3.3, situação esta que impossibilita uma melhor avaliação e conclusão. Ainda, dentre estes tipos de documentos não é proposto um documento clínico que apóie especificamente a decisão diagnóstica em ambiente colaborativo, como propõe a DDSOnt.

7.3 *Contribui original para a área de Informática em Saúde*

Durante a comparação das atividades desenvolvidas e os resultados obtidos destes trabalhos estudados com esta proposta, verificou-se que:

- Os trabalhos encontrados, em comparação a este trabalho, exploram de maneira superficial os problemas de incompatibilidade técnica, terminológica e ontológica durante o fornecimento de informações heterogêneas por parte de aplicações legadas aos CDSSs;
- Os trabalhos estudados procuram enfatizar propostas de representação interoperável especificamente do conteúdo da base de conhecimento de um CDSS, ou seja, buscam representar as regras de decisão médica presentes na base de conhecimento através de linguagens baseadas em regras e independentes de plataforma. Para isto, são utilizadas técnicas baseada nas linguagens MLM (Hripsak G et al. 1994), Arden Syntax (Jenders et al. 2003) e GLIF (Boxwala et al. 2004). Já este trabalho procura abstrair a forma com que as regras de decisão médica são representadas em uma base de conhecimento, mas sabe-se que este recurso é realizado em nossa proposta por técnicas de aprendizado de máquina disponível na literatura, como árvore de decisão, redes neurais, redes bayesianas, e outras mais. Estes aspectos específicos de representação de uma base

⁵<http://www.hl7.org/about/>

de conhecimento são também abstraídos pelo usuário quando este faz uso do serviço Web getDDS() que compõe a solução DDSOntWs;

- O fato de abstrair as regras de decisão médica presentes em uma base de conhecimento de um CDSS é justificado em função da proposta deste trabalho. Procurou-se explorar e criar uma solução que promova interoperabilidade em termos dos dados clínicos em nível terminológico, técnico e ontológico, e não em nível de regras de decisão médica presentes em um base de conhecimento;
- Viabilizado o cenário de um CDSS poder formar bases de conhecimento a partir de informações fornecidas por diferentes médicos em aplicações heterogêneas, os artigos encontrados não exploram, como faz este trabalho, as diversas possibilidades decorrentes de um ambiente que permite apoiar a decisão diagnóstica em ambiente colaborativo na Web. São possibilidades descritas por este trabalho a criação, o compartilhamento e a recuperação de bases de conhecimento, e a obtenção de uma resposta diagnóstica a partir de bases de conhecimento diversas;
- Finalmente, os tipos semânticos *DDSOntADT*, *Signal* e *ClinicalDiagnosisEvaluation* definidos na DDSOnt visam facilitar o entendimento semântico, por parte de aplicações computacionais, de informações sendo transferidos durante o apoio de uma decisão diagnóstica em ambiente colaborativo. As ontologias investigadas e estudadas na literatura não fornecem tipos semânticos específicos, como faz a DDSOnt, para suportar uma estrutura que auxilie no compartilhamento e na recuperação de bases de conhecimento utilizadas por médicos no apoio à decisão diagnóstica em ambiente colaborativo;

É neste sentido portanto, que este trabalho efetivamente contribui de maneira original para a área de Informática em Saúde.

7.4 Trabalhos Futuros

O processo diagnóstico pode ser melhorado quando suportado por um CDSS, como foi apresentado neste trabalho, a partir da construção de um cenário onde entidades reguladoras, provedoras e requisitantes de bases de conhecimento podem trocar experiências de apoio à decisão diagnóstica em ambiente colaborativo através de recursos computacionais como a ontologia DDSOnt, e os softwares JDDSOnt e DDSOntWs.

Um outro cenário interessante e não explorado neste trabalho é quando o médico pode ser apoiado por ferramentas de comunicação como Vídeo-Conferência, *Chat*, Comunicador Instantâneo (CI) e *Whiteboard*⁶ para trocar experiências clínicas com outros médicos. Estas ferramentas permitem que médicos possam trocar textos, manter contatos áudio-visuais, transferir arquivos digitais de exames (imagens, gráficos, etc.), fazer desenhos ilustrativos, etc., objetivando melhor concluir e apurar uma segunda opinião diagnóstica.

Estes dois cenários, de apoio a decisão diagnóstica e de suporte a segunda opinião diagnóstica, quando integrados, podem criar um terceiro e novo cenário inovador caracterizado por discussões diagnósticas suportadas por ferramentas de comunicação multimídia e fomentadas por uma ontologia e bases de conhecimento. As bases de

⁶Lousa eletrônica utilizada para escrita e desenho

conhecimento que produzem melhores percentuais de classificação correta de diagnósticos são as formadas pelos casos clínicos contendo sinais que melhor expressam o diagnóstico associado. Sendo assim, a atividade de criar, discutir e atualizar o conjunto de casos clínicos precisa ocorrer em um ambiente que possibilite que médicos possam discutir e trocar suas experiências clínicas para a formação de uma base de conhecimento mais precisa e refinada, traduzindo-se em diagnósticos mais apurados, rápidos, e efetivos (Pires et al. 2006b).

A possibilidade de discussão pelos melhores casos clínicos e de análise de opiniões diagnósticas diversas entre alunos e professores em ambiente como o cenário inovador apresentado pode enriquecer especificamente o processo educacional na área médica. Neste caso, alunos e professores podem construir conhecimento clínico em ambiente suportado por ferramentas de comunicação multimídia, e ainda por um LMS (*Learning Management System*). O TIDIA-Ae⁷ possui essas características e apresenta-se como um bom ambiente de estudo de caso para este projeto.

O projeto TIDIA-Ae, financiado e apoiado pela FAPESP, tem como objetivo criar um conjunto de esforços para fomentar pesquisa e desenvolvimento de tecnologias computacionais que suportem o gerenciamento de atividades de ensino-aprendizagem, podendo estas serem executadas na modalidade presencial ou a distância. A ferramenta LMS é responsável pelo gerenciamento de usuários, como a autorização, a autenticação, o papel e a permissão dos participantes. Ainda, a partir do ambiente TIDIA-Ae pode-se instanciar ferramentas de comunicação como comunicador instantâneo (CI) (Lobato et al. 2006), *chat*, fórum de discussão, *whiteboard*, editor colaborativo, e ferramenta de integração de experimentos remotos. Especificamente, o autor deste trabalho participa no desenvolvimento do CI.

A ontologia DDSOnt e os softwares relacionados serão adicionados ao ambiente TIDIA-Ae de modo que este ambiente possa suportar a troca de informações diagnósticas objetivando:

1. oferecer alternativas de atividades didáticas no ensino médico e,
2. refinar bases de conhecimento que auxiliam no apoio à decisão diagnóstica.

O projeto de pesquisa do aluno Pereira (Pereira et al. 2007) está relacionado ao primeiro objetivo descrito, enquanto que o projeto de pesquisa de Martins (Martins et al. 2007) relaciona-se com o segundo objetivo. Ambos são orientados pelo autor desta tese. Estes dois projetos de pesquisa estão em andamento e tem previsão de término para o final de 2008.

Com as adições dos recursos da DDSOnt no ambiente TIDIA-Ae, pretende-se implementar os seguintes cenários, considerando a existência de salas remotas com alunos e professores de medicina participando de uma reunião em diferentes locais, e as ferramentas de comunicação disponíveis são o CI com os recursos de áudio e

vídeo (Pires et al. 2006b):

1. Configuração do contexto onde será realizada a reunião: definição dos participantes, agendamento prévio da patologia discutida, planejamento das ferramentas de comunicação que poderão ser utilizadas. Pode-se definir ainda se as informações criadas serão compartilhadas apenas para o grupo envolvido, ou para todos os participantes do ambiente;

⁷Tecnologia da Informação para o Desenvolvimento da Internet Avançada - Aprendizado Eletrônico – <http://tidia-ae.incubadora.fapesp.br/portal>.

2. Análise dos diagnósticos que serão considerados para o estudo da patologia com a ferramenta CI. Por exemplo, para a patologia de dor abdominal aguda serão considerados os diagnósticos de gastrite, de diverticulite, e de inflamação pélvica;
3. Discussão dos sinais clínicos e laboratoriais que serão considerados importantes para a patologia, bem como seus possíveis valores. Por exemplo, para a patologia citada no item anterior, os sinais podem ser a dilatação (sim ou não), a intensidade da dor (desconforto, moderado e severo), e fatores de aumento da dor (movimento, alimento, e exercícios);
4. Discussão dos casos clínicos que irão originar o conjunto de treinamento para a criação de uma base de conhecimento;
5. Durante a definição dos diagnósticos, dos sinais e de seus valores, e dos casos clínicos, pode acontecer de eventualmente alguns professores criarem ou definirem valores alternativos com objetivos de se criar bases de conhecimento diferentes para comparações futuras de diferentes respostas diagnósticas. Neste caso, pode-se redefinir os tipos de acesso para aquelas informações em particular;
6. Considerando que existam diversas bases de conhecimento de diferentes patologias, deve-se buscar por uma base de conhecimento antes de consultar por uma resposta diagnóstica;
7. Efetivamente realizar consultas por respostas diagnósticas e comparar o resultado com participantes que eventualmente utilizaram outras bases de conhecimento.

Para a implementação dos cenários apresentados e que serão realizados no contexto dos trabalhos de Pereira e Martins, pretende-se estabelecer contato e parceria com grupos de médicos e residentes da Faculdade de Medicina da USP/Ribeirão Preto. Sendo assim, os valores de sinais e diagnósticos clínicos, assim como os casos clínicos criados e avaliados no ambiente do TIDIA-Ae e suportados pela DDSOnt podem efetivamente possuir valor médico e científico. Ainda, estes estudos de caso servirão também como um teste de usabilidade para a DDSOnt e dos softwares relacionados.

Outros trabalhos futuros podem ser realizados com objetivo de adicionar funcionalidades ou estender a abrangência da proposta deste trabalho:

- Foram considerados e testados no estudo de caso apenas os termos clínicos no idioma inglês da *UMLS MetaThesaurus*, visto que as entidades reguladora, provedora e requisitante utilizam-se apenas deste idioma. Sendo assim, propõe-se um estudo de caso onde as entidades utilizam-se de idiomas diferentes, e assim mais de um servidor da *UMLS MetaThesaurus* tenha que ser utilizado. Segundo a documentação da *UMLS MetaThesaurus* (Thesaurus 2007), sua estrutura acomoda traduções dos sistemas de terminologias presentes em idiomas diferentes do inglês, e que existem muitas traduções em diferentes idiomas na última versão presente na UMLSKS 2007AA;
- Estudar e utilizar as linguagens MLM e Arden Syntax de modo a propor uma representação interoperável das regras de decisão médica presentes nas bases de conhecimento suportada pela DDSOnt;

- A API JDDSOnt foi utilizada pelas entidades reguladora, provedora e requisitante no estudo de caso para a construção de documentos XML/RDF em conformidade com a DDSOnt. Estes documentos são utilizados na implementação e uso dos serviços Web da DDSOntWs. Como atividade futura, pode-se estender as funcionalidades da JDDSOnt com a criação de programas que permitam responder as perguntas que a DDSOnt foi projetada para responder, como apresentado no Capítulo 5. Estas funcionalidades podem ser atualmente implementadas com o uso da API JENA.

De posse das idéias apresentadas neste trabalho, da ontologia DDSOnt definida, e dos artefatos de softwares JDDSOnt e DDSOntWs confeccionados, o autor deste trabalho, juntamente com seu grupo de pesquisa, deve aplicá-los e colocá-los em prática por meio dos trabalhos futuros apresentados, e outros que podem ser vislumbrados posteriormente. E através dos resultados destas aplicações, mostrar de que forma pode-se contribuir ainda mais para a área de Informática em Saúde.

Referências

- Achour, S., Dojat, M., Rieux, C., Bierling, P., and Lepage, E. (2001). A umls-based knowledge acquisition tool for rule-based clinical decision support system development. *Journal of American Medical Informatics Association*, 8(4):351–360.
- Ashburner, M., Ball, C., Blake, J., Botstein, D., Butler, H., Cherry, J., Davis, A., Dolinski, K., Dwight, S., Eppig, J., Harris, M., Hill, D., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J., Richardson, J., Ringwald, M., Rubin, G., and Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. the gene ontology consortium. *Nature Genetics*, 25(1):25–29.
- Barnett, G., Justice, N., Somand, M., Adams, J., Waxman, B., Beaman, P., Parent, M., Van Deusen, F., and Greenlie, J. (1979). Costar: A computer-based medical information system for ambulatory care. *Proceedings of the IEEE*, 67(9):1226–1237.
- Barstow, A., Hendler, J., Skall, M., Pollock, J., Martin, D., Marcatte, V., McGuinness, D. L., Yoshida, H., and Roure, D. (2004). Owl web ontology language for services (owl-s), <http://www.w3.org/submission/2004/07/>. Última visita:2007.
- Beeler, G., Case, J., Curry, J., Hueber, A., Mckenzie, L., Schadow, G., and Shakir, A. (2003). Hl7 reference information model, <http://www.hl7.org/v3ballot/html/infrastructure/rim/rim.htm>. Última visita:2007.
- Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific American*, 5:28–37.
- Bilykh, I., Jahnke, J., McCallum, G., and Price, M. (2006). Using the clinical document architecture as open data exchange format for interfacing emrs with clinical decision support systems. In *Proceedings of the 19th IEEE International Symposium on Computer-Based Medical Systems*, volume 1, pages 855–860, Salt Lake City, USA. IEEE.
- Booth, D., Haas, H., McCabe, F., Newcomer, E., Champion, M., Ferris, C., and Orchard, D. (2006). Web service architecture w3c working group, <http://www.w3.org/tr/ws-arch/>. Última visita:2007.
- Boxwala, A., Peleg, M., and Tu, S. e. a. (2004). Glif3: a representation format for sharable computer-interpretable clinical practice guidelines. *Journal of Biomedical Informatics*, 37(3):147–161.
- Boyer, S., Chute, C., Cimino, J., Hammond, W., and Huff, W. (2007). Hl7 vocabulary domains, <http://www.hl7.org/v3ballot/html/infrastructure/vocabulary/vocabulary.htm>. Última visita:2007.

- Bradley, C. P. (2005). Can we avoid bias ? *British Medical Journal*, 330:784.
- Breitman, K. (2005). *Web Semântica: A Internet do Futuro*. Editora LTC, Rio de Janeiro.
- Chinnici, R., Gudgin, M., Moreau, J., Schlimmer, J., and Weerawarana, S. (2006). Web services description language (wsdl) version 2.0 part 1: Core language, <http://www.w3.org/tr/2004/wd-wsdl20-20040803/>. Última visita:2007.
- CID-10 (1993). Cid-10 - classificação estatística internacional de doenças e problemas relacionados à saúde, faculdade de saúde pública da universidade de são paulo and organização mundial de saúde and organização pan-americana de saúde, <http://www.datasus.gov.br/cid10/webhelp/cid10.htm>. Última visita:2007.
- Coskun, N. and Sessions, R. (1992). Class objects in som, ibm personal systems developer. (summer 1992): 67-77.
- Cote, R., Rothwell, D., Palotay, J., Beckett, R., and Brochu, L. (1993). The systematized nomenclature of human and veterinary medicine: Snomed international. northfield, il: College of american pathologists.
- Dolin, R., Alschuler, L., Beebe, C., Biron, P., Boyer, S., Behlen, F., and Shabo, A. (2005). Clinical document architecture, release 2, <http://www.hl7.org/v3ballot/html/infrastructure/cda/cda.htm>. Última visita:2007.
- Dolin, R., Alschuler, L., Beebe, C., Biron, P., Boyer, S., Essin, D., Kimber, E., and et al (2001). The hl7 clinical document architecture. *Journal of American Medical Informatics Association*, 8(1):552-569.
- Fernández-López, M., Gómez-Pérez, A., and Juristo, N. (1997). Methontology: From ontological art towards ontological engineering. In *Proceedings of the AAAI Symposium on Ontological Engineering*, volume 1, Stanford - EN. ACM.
- Greive, G., Hamm, R., Esler, B., and et al. (2007). Templates specification, <http://www.hl7.org/v3ballot/html/infrastructure/templates/templates.htm>. Última visita:2007.
- Gruninger, M. and Fox, M. (2002). Tove ontology project, <http://www.eil.utoronto.ca/enterprise-modelling/tove/>. Última visita:2007.
- Guarino, N. and Welty, C. (2002). Evaluating ontological decisions with ontoclean. *Communications of the ACM*, 45(2):61-65.
- Haas, H. (2006). Web service activity statement, <http://www.w3.org/2002/ws/activity>. Última visita:2007.
- Hassan, E. (2007). Hyperglycemia management in the hospital setting. *American Journal of Health-System Pharmacy*, 64(20):9-14.
- Haug, P., Rocha, B., and Evans, R. (2003). Decision support in medicine: lessons from the help system. *International Journal of Medical Informatics*, 69(2-3):273-284.
- Horstmann, M. and Kirtland, M. (1997). Dcom architecture, <http://msdn2.microsoft.com/en-us/library/ms809311.aspx>. Última visita: 2007.
- Hripcsak G, P., Ludemann, T., Pruor, O., and Wigertz, P. (1994). Rationale for the arden syntax. *Computers in Biomedical Research*, 27(4):291-324.
- Jenders, R., Corman, R., and Dasgupta, B. (2003). Making the standard more standard: a data and query model for knowledge representation in the arden

- syntax. In *Proceedings of the AMIA Annual Symposium*, volume 1, pages 323–330, Washington, USA. AMIA.
- Kazemzadeh, R.S. and Sartipi, K. (2005). Interoperability of data and knowledge in distributed health care systems. In *Proceedings of the 13th IEEE International Workshop on Software Technology and Engineering Practice*, volume 1, pages 230–240, Budapest, Hungary. IEEE.
- Krause, A., Hartl, D., Theis, F., Stangl, M., Gerauer, K., and Mehlhorn, A. (2004). Mobile decision support for transplantation patient data. *International Journal of Medical Informatics*, 73(5):461–464.
- Lipscomb, C. E. (2000). Medical subject headings (mesh). *Journal of the Medical Library Association*, 88(3):265–266.
- Liu, S., Ma, W., Moore, R., Ganesan, V., and Nelson, S. (2005). Rxnorm: Prescription for electronic drug information exchange. *IEEE Computer Society*, 5(1):17–23.
- Lobato, D. C., Baldochi, L. A., Pires, D. F., Pessoa, T. R. M., Gasparin, R., Montoro, F., and Teixeira, C. (2006). A multimedia instant messenger for an e-learning environment. In *Proceedings of II Workshop do TIDIA*, volume 1, pages 1–8, São Paulo, Brasil. FAPESP.
- Martin, D., Burstein, M., Hobbs, J., Lassila, O., McDermott, D., McIlraith, S., Narayanan, S., Paolucci, M., Parsia, B., Payne, T., Sirin, E., Srinivasan, N., and Sycara, K. (2004). Owl-s: Semantic markup for web services, <http://www.w3.org/submission/2004/subm-owl-s-20041122/>. Última visita:2007.
- Martins, J. A. C. M., Pires, D. F., and Teixeira, C. A. C. (2007). Construção de bases de conhecimento clínicas para segunda opinião diagnóstica em ambiente suportado por ontologias e pela comunicação multimídia. projeto de pesquisa tt2 adjacente ao processo 05/60653-1. tidia-ae fase ii. fapesp.
- McCray, A. T. (2003). An upper-level ontology for the biomedical domain. *Comparative and Functional Genomics*, 4(1):80–84.
- McGuinness, D. and Harmelen, F. (2004). Owl web ontology language overview, <http://www.w3.org/tr/owl-features/>. Última visita:2007.
- Miller, E. and Hendler, J. (2006). Web ontology language (owl), <http://www.w3.org/2004/owl/>. Última visita:2007.
- Miller, E., Hongsermeier, T., Neumann, E., and Gilman, B. (2006a). W3c semantic web health care and life sciences interest group, <http://www.w3.org/2001/sw/hcls/>. Última visita:2007.
- Miller, E., Swick, R., and Brickley, D. (2006b). Resource description framework (rdf), <http://www.w3.org/rdf/>. Última visita:2007.
- Miller, E., Swick, R., Brickley, D., McBride, B., Hendler, J., Schreiber, G., Wood, D., and Connolly, D. (2006c). Semantic web activity, <http://www.w3.org/2001/sw>. Última visita:2007.
- Mitra, N. (2006). Soap version 1.2 part 0: Primer, <http://www.w3.org/tr/2003/rec-soap12-part0-20030624/>. Última visita:2007.
- Model, R. I. (1994). Health level seven incorporation. hl7 reference information model, <http://www.hl7.org/library>. Última visita:2007.
- Murff, H. J., Gandhi, T. K., Karson, A. K., Mort, E. A., Poon, E. G., Wang, S. J., Fairchild, D. G., and Bates, D. W. (2003). Primary care physician attitudes concerning follow-up of abnormal test results and ambulatory decision support systems. *International Journal of Medical Informatics*, 71(2-3):137–149.

- Noy, N. and McGuinness, D. L. (2001). *Ontology Development 101 - A guide to creating your first ontology*. KSL Technical Report. Stanford University.
- Osterwalder, A. and Pigneur, Y. (2002).
An e-Business Model Ontology for Modeling e-Business, <http://inforge.unil.ch/aosterwa/Documents/eBusinessModels/Publications/Bled02.htm>. Última visita:2007.
- Pereira, C. M. M., Pires, D. F., and Teixeira, C. A. C. (2007). Projeto de pesquisa ttl adjacente ao processo 05/60653-1. tidia-ae fase ii. fapesp.
- Pires, D., Halah, R., and Ruiz, E. (2006a). Sharing web clinical knowledge between medical community to support a second opinion diagnosis. In *Proceedings of XX-XIII Seminário Integrado de Software e Hardware. XXVI Congresso da Sociedade Brasileiro de Computação*, volume 1, pages 361–375.
- Pires, D., Halah, R., Ruiz, E., and Teixeira, C. (2006b). Compartilhamento do conhecimento clínico suportado por ontologias e pela comunicação multimídia. In *Proceedings of XII Simpósio Brasileiro de Sistemas Multimídia e Web*, volume 2, pages 23–25, Natal. WebMedia.
- Pires, D., Halah, R., Tinos, R., and Ruiz, E. (2004a). Uma arquitetura para suportar o compartilhamento de conhecimento clínico em sistemas pep integrados a sistemas de auxílio ao diagnóstico. In *Proceedings of the IX Congresso Brasileiro de Informática em Saúde*, volume 1, pages 593–598.
- Pires, D., Halah, R., Tinos, R., and Ruiz, E. (2004b). Uma arquitetura para suporte à atividade de telemedicina através do compartilhamento do conhecimento clínico extraído de sistemas de prontuário eletrônico. In *Proceedings of the 4th WebMedia PhD and MSc Proposals Workshop*, volume 2, pages 226–229.
- Pisa, I., Galina, A., Lopes, P., Barsottini, C., and Silva, A. (2004). Lepidus r3: Implementação de sistema de apoio à decisão médica em arquitetura distribuída usando serviços web. In *Proceedings of the IX Congresso Brasileiro de Informática em Saúde*, volume 1, pages 779–785, Ribeirão Preto, Brasil. SBIS.
- Porto, C. (1991). *Semiologia Médica*. Editora Guanabara Koogan, Goiânia.
- Read, J. and Benson, T. (1986). Comprehensive coding. *British Journal of Health Care Computing*, 1(5):22–25.
- Rector, A. and Glowinski, A. (1995). Medical-concept models and medical records: An approach based on galen and penpad. *Journal of American Medical Informatics Association*, 2(1):19–35.
- Rector, A., Goble, C., and Horrocks, I. and Stevens, R. I. v. (2007). Co-ode / hyontuse, <http://www.co-ode.org/>.
- Rector, A., Rogers, J., and Pole, P. (1996). The galen high level ontology. In *Proceedings of the Fourteenth International Congress of the European Federation for Medical Informatics*, volume 1, pages 1–6, Copenhagen - Denmark. MIE.
- Rezende, S. (2002). *Sistemas Inteligentes: Fundamentos e Aplicações*. Manole.
- Samore, M., Bateman, K., Alder, S. C., Hannah, E., Donnelly, S., Stoddard, G. J., Haddadin, B., Rubin, M., Williamson, J., Stults, B., Rupper, R., and Stevenson, K. (2005). Clinical decision support and appropriateness of antimicrobial prescribing. *Journal of the American Medical Association*, 294(18):2305–2314.
- Schuyler, P. L., Hole, W. T., Tuttle, M. S., and Sherertz Liu, D. D. (1993). The umls metathesaurus: representing different views of biomedical concepts. *Journal of the Medical Library Association*, 2(81):217–222.
- SemanticNetwork, U. (2006). Umls knowledge sources documentation, <http://www.nlm.nih.gov/research/umls/umlsdoc.html>. Última visita:2007.

- Seven, H. L. (2006). Health level 7, <http://www.hl7.org>. Última visita:2007.
- Short, D., Frisher, M., and Bashford, J. (2004). Barriers to the adoption of computerized decision support system in general practice consultations: a qualitative study of GP's perspectives. *International Journal of Medical Informatics*, 73(4):357–362.
- Shortliffe, E., Perreault, L., Wiederhold, G., and Fagan, L. (2000). *Medical Informatics: Computer Applications in Health Care and Biomedicine*. Springer.
- Silva, R. and Roque, A. (2000). Clinical medical diagnosis using a signal-processing approach. In *Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences*, volume 1, pages 13–18, Las Vegas, USA. CSREA.
- Sjöborg, B., Bäckström, T., Arvidsson, L., Andersén-Karlsson, E., Blomberg, L., Eiermann, B., Eliasson, M., Henriksson, K., Jacobsson, L., and Jacobsson, U., e. a. (2007). Design and implementation of a point-of-care computerized system for drug therapy in stockholm metropolitan health regionÜbridging the gap between knowledge and practice. *International Journal of Medical Informatics*, 76(7):497–506.
- Smith-Akin, K., Bearden, C., Pittenger, S. T., and E.V., B. (2007). Toward a veterinary informatics research agenda: An analysis of the pubmed-indexed literature. *International Journal of Medical Informatics*, 76(4):306–312.
- SNOMED (2007). What is snomed ct. snomed clinical terms core content, http://www.snomed.org/snomedct/what_is.html. Última visita:2007.
- Spackman, K., Campbell, K., and Côté, R. (1997). Snomed rt: A reference terminology for health care. In *Proceedings of the American Medical Informatics Association Annual Symposium*, volume 1, pages 1–5.
- Stearns, M., MPhil, C., Spackman, K., and Wang, A. (2001). Snomed clinical terms: Overview of the development process and project status. In *Proceedings of the American Medical Informatics Association Annual Symposium*, volume 1, pages 1–5, Washington - DC. AMIA.
- Sun, M. (2007). Java remote method invocation - distributed computing for java, <http://java.sun.com/javase/technologies/core/basic/rmi/whitepaper/index.jsp>. Última visita:2007.
- Sycara, K., Klusch, M., Widoff, S., and Lu, J. (1999). Dynamic service matchmaking among agents in open information environments. *ACM SIGMOD Record (Special Issue on Semantic Interoperability in Global Information Systems)*, 28(1):47–53.
- System, U. M. L. (2006). Unified medical language system - national library of medicine, <http://www.nlm.nih.gov/research/umls/>. Última visita:2007.
- Tan, T., Quek, C., and Ng, G. (2005). Ovarian cancer diagnosis using complementary learning fuzzy neural network. In *Proceedings of the 2005 IEEE International Joint Conference on Neural Networks*, volume 5, pages 3034–3039, Montreal, Canada. IEEE.
- Thesaurus, U. (2007). Umls knowledge sources documentation, <http://www.nlm.nih.gov/research/umls/umlsdoc.html>. Última visita: 2007.
- Tu, S. (2007). Health level seven (hl7) data types and top-level reference information model (rim) classes.
- UDDI, O. (2007). Oasis uddi advancing web services discovery standard, <http://www.uddi.org/>. Última visita:2007.

- Wyatt, J. and Spiegelhalter, D. (1991). Field trials of medical decision-aids: potential problems and solutions. In *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, volume 1, pages 3–7, New York. McGraw-Hill.
- Zheng, K., Padman, R., Johnson, M., and Diamond, H. (2005). Understanding technology adoption in clinical care: Clinician adoption behavior of a point-of-care reminder system. *International Journal of Medical Informatics*, 74(7-8):535–543.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)