

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

MAPEAMENTO SISTEMÁTICO SOBRE O USO
DE ONTOLOGIAS EM INFORMÁTICA MÉDICA

MOISÉS ROBERTO DE ARAÚJO MOTA

JOÃO PESSOA

Agosto - 2013

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

**MAPEAMENTO SISTEMÁTICO SOBRE O USO DE
ONTOLOGIAS EM INFORMÁTICA MÉDICA**

MOISÉS ROBERTO DE ARAÚJO MOTA

JOÃO PESSOA-PB

Agosto-2013

MOISÉS ROBERTO DE ARAÚJO MOTA

**Mapeamento Sistemático Sobre o Uso de Ontologias em
Informática Médica**

DISSERTAÇÃO APRESENTADA AO CENTRO DE INFORMÁTICA DA
UNIVERSIDADE FEDERAL DA PARAÍBA, COMO REQUISITO PARCIAL
PARA OBTENÇÃO DO TÍTULO DE MESTRE EM INFORMÁTICA (SISTEMAS
DE COMPUTAÇÃO).

Linha de Pesquisa: Computação Distribuída

Orientador: Natasha Correia Queiroz Lino

JOÃO PESSOA-PB

Agosto-2013

III

M917m Mota, Moisés Roberto de Araújo.

*Mapeamento sistemático sobre o uso de ontologias
em informática médica / Moisés Roberto de Araújo Mota. -
- João Pessoa, 2013.*

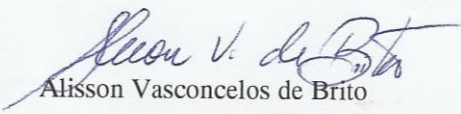
115f. : il.

Orientadora: Natasha Correia Queiroz Lino

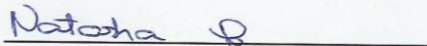
Dissertação (Mestrado) - UFPB/CI

Ata da Sessão Pública de Defesa de Dissertação de
Mestrado de **Moisés Roberto de Araújo Mota**,
candidato ao Título de Mestre em Informática na
Área de Sistemas de Computação, realizada em 06
de Agosto de 2013.

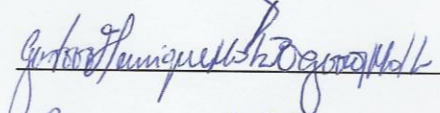
Ao sexto dia do mês de Agosto do ano dois mil e treze, às dez horas, no laboratório 2 da Escola Superior de Redes - Universidade Federal da Paraíba, reuniram-se os membros da Banca Examinadora constituída para examinar o candidato ao grau de Mestre em Informática, na área de "*Sistemas de Computação*", na linha de pesquisa "*Computação Distribuída*", o Sr. **Moisés Roberto de Araújo Mota**. A comissão examinadora foi composta pelos professores doutores: NATASHA CORREIA QUEIROZ LINO (PPGI-UFPB), Orientadora e Presidente da Banca, GUSTAVO HENRIQUE MATOS BEZERRA MOTTA (PPGI-UFPB), examinador interno, LEONARDO VIDAL BATISTA (PPGI-UFPB), examinador interno e ROBSON DO NASCIMENTO FIDALGO (UFPE) como examinador externo. Dando início aos trabalhos, a professora NATASHA CORREIA QUEIROZ LINO cumprimentou os presentes, comunicou aos mesmos a finalidade da reunião e passou a palavra ao candidato para que o mesmo fizesse, oralmente, a exposição do trabalho de dissertação intitulado "*Mapeamento Sistemático Sobre o Uso de Ontologias em Informática Médica*". Concluída a exposição, o candidato foi arguido pela Banca Examinadora que emitiu o seguinte parecer: "*Aprovado*". Assim sendo, deve a Universidade Federal da Paraíba expedir o respectivo diploma de Mestre em Informática na forma da lei e, para constar, eu, Alisson Vasconcelos de Brito, Coordenador deste Programa, servindo de secretário, lavrei a presente ata que vai assinada por mim e pelos membros da Banca Examinadora. João Pessoa, 06 de Agosto de 2013.


Alisson Vasconcelos de Brito

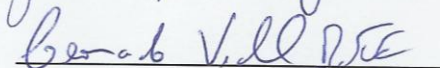
Profa. Dra. Natasha Correia Queiroz Lino
Orientadora (PPGI-UFPB)



Prof. Dr. Gustavo Henrique Matos Bezerra Motta
Examinador Interno (PPGI-UFPB)



Prof. Dr. Leonardo Vidal Batista
Examinador Interno (PPGI-UFPB)



Prof. Dr. Robson do Nascimento Fidalgo
Examinador Externo (UFPE)



Agradecimentos

Agradeço primeiramente à Deus pela saúde e graça proporcionadas para que eu concluísse mais esta etapa da minha vida.

Sou muito grato à UFPB pela oportunidade de cursar esta pós-graduação nesta renomada instituição, à minha orientadora, Natasha Correia Queiroz Lino, por ter aceitado me orientar, pela paciência, por acreditar em mim e me auxiliar em todos os passos até a conclusão deste estudo. Agradeço também à CAPES pelo investimento neste projeto de pesquisa.

Sou muitíssimo grato, especialmente, aos meus pais pelo investimento, incentivo e força para concluir esta etapa tão importante para a vida de um pesquisador, também não posso deixar de agradecer a minha namorada pela ajuda, paciência, e compreensão, em todos os momentos deste mestrado. A presença dela foi essencial em muitos momentos desta pós-graduação. E agradeço aos meus amigos pela paciência.

Agradeço a equipe de pesquisadores que trabalharam comigo durante toda este pesquisa, este mestrado não teria sido concluído sem vocês.

Resumo

A quantidade de estudos no uso de Ontologias em Informática Médica tem crescido através dos anos. Isto demonstra o interesse no desenvolvimento de tecnologias com bases semânticas para esta nova área da ciência.

No entanto, pouco tem sido documentado a respeito de como a área tem se desenvolvido, dificultando, assim, o processo de criação de projetos relevantes na área da Informática Médica, uma vez que não é possível mapear as principais oportunidades de pesquisa, lacunas e necessidades da área.

Logo, o presente projeto tem o objetivo de mapear o desenvolvimento geral desta área, especificamente sobre o uso de Ontologias em Informática Médica, de modo a evidenciar e preencher as necessidades da área.

Para tanto, foi utilizada a metodologia de pesquisa sistematizada conhecida como Mapeamento Sistemático, a qual segue um protocolo de pesquisa bem definido, transparente e rigoroso, permitindo que o estudo seja avaliado, validado e reproduzido pela comunidade científica interessada. Deste modo, foi reunido o conhecimento atual encontrado na literatura sobre o Uso de Ontologias em Informática Médica.

A partir das buscas automáticas e manuais, foram retornados 23788 estudos relacionados ao domínio desta pesquisa. Após a realização de uma seleção rigorosa, este número caiu para 511 artigos relevantes, sobre os quais foram realizadas as análises mais aprofundadas para a construção do Mapeamento Sistemático proposto.

Foram observadas 07 características destes estudos, que permitiram observar as principais tendências de crescimento da área como um todo. A avaliação da qualidade dos trabalhos selecionados forneceu segurança para a construção das conclusões desta pesquisa, tendo em vista que em torno de 95% possuem fortes evidências para apoiar os resultados apresentados.

Apesar dos resultados, foram encontradas algumas lacunas no que diz respeito à profundidade da pesquisa. Foi possível concluir que o Uso de Ontologias em Informática Médica tem crescido de acordo com o esperado, considerando o reuso de ontologias, a integração e interoperabilidade de sistemas e ontologias diferentes. Quanto às oportunidades, identificamos que nesta área há a necessidade de métodos para avaliação, validação, correção, completude e manutenção de ontologias novas ou já estabelecidas, como também o desenvolvimento de aplicações e estudos relacionados a telemedicina, saúde pública, educação, robótica, pesquisas baseadas em evidências, e gestão financeira, voltados para a Informática Médica.

Palavras-Chave: Informática Médica, Ontologias, Mapeamento Sistemático, Representação do Conhecimento.

Abstract

The number of studies on the use of Ontologies in Medical Informatics has grown through the years. It shows the interest in developing technologies with semantic bases for this new area of science.

However, little work has been documented about how the area has developed, thus hampering the creation process of relevant projects in the field of Medical Informatics, since it is not possible to map the main research opportunities, gaps and needs of this area.

Therefore, this project aims to map the overall development of this area, specifically on the use of Ontologies in Medical Informatics, in order to highlight and fill the area needs.

For this, we used the methodology of systematic research known as Systematic Mapping, which follows a research protocol clearly defined, transparent and rigorous, enabling the study be evaluated and validated by the scientific community interested reproduced. In this way, the current knowledge was gathered in the literature on the Use of Ontologies in Medical Informatics.

From automatic and manual searches, it was returned 23788 studies related to this research area. After conducting a rigorous selection, this number dropped to 511 relevant papers, which were performed on the further analysis for the proposed construction of Systematic Mapping.

It was detected 07 characteristics of these studies, which allowed the observation of the main trends of growth of the area as a whole. The quality assessment of the selected works provided security to take conclusions of this research, considering that around 95% have strong evidence to support the presented results.

Despite these results, we found some gaps in relation to the depth of this search. We concluded that the use of Ontologies in Medical Informatics has grown as expected, considering the reuse of ontologies, integration and interoperability of systems and different ontologies. About the opportunities, we identified in this area a need for methods for evaluation, validation, correctness, completeness and maintenance of new or already established ontologies, and also the development of applications and studies related to telemedicine, public health, education, robotics, evidence-based research, and financial management, focused on Medical Informatics.

Keywords: Medical Informatics, Ontologies, Systematic Mapping, Knowledge Representation.

LISTA DE FIGURAS

Figura 1 - Árvore de Porfírio	6
Figura 2 - Hierarquia de uma Taxonomia	6
Figura 3 - Classificação de Ontologias de acordo com sua generalidade	8
Figura 4 - Arquitetura Bolo de Noiva proposta por Tim Benners Lee	10
Figura 5 - Os principais uso das Tecnologias da Informação em Saúde	16
Figura 6 - Distribuição dos trabalhos retornados pelos engenhos de busca	34
Figura 7 - Total de artigos retornados pelos engenhos acrescido o PubMed e PMC	35
Figura 8 - Representatividade dos Engenhos após a seleção de artigos relevantes	38
Figura 9 - Formulário para a extração de dados desenvolvido para esta pesquisa	50
Figura 10 - Quantidade de Publicações relevantes para a área através dos anos	55
Figura 11- Principais usos de Ontologias em Informática Médica.....	57
Figura 12 - Principais Ferramentas utilizadas para criação e edição de ontologias	68
Figura 13 - Métodos de Avaliação/Validação de Ontologias	70
Figura 14 - Artigos classificados de acordo com a Taxonomia	73
Figura 15 - Qualidade das Trabalhos	76
Figura 16 - Protótipo de Taxonomia Ontológica	83
Figura 17. Taxonomia Ontológica desenvolvida no presente trabalho	108

LISTA DE TABELAS

Tabela 1 - Engenhos utilizados na busca automática	29
Tabela 2 - Engenhos de busca excluídos.....	29
Tabela 3 - Periódicos selecionados para a busca manual	30
Tabela 4 - Classificação desta pesquisa segundo a taxonomia de Cooper	32
Tabela 5 - Classificação da Pesquisa segundo Marconi e Lakatos	33
Tabela 6 - Engenhos adicionados posteriormente	35
Tabela 7 - % sobre o total dos artigos retornados pelos engenhos.....	36
Tabela 8 – Trabalhos selecionados com base análise de título, resumo e introdução	37
Tabela 9 - Aplicação dos Critérios de Inclusão e Exclusão	39
Tabela 10 - Diretrizes para a documentação dos resultados da pesquisa sugerido por Kitchenham	40
Tabela 11 - Definição dos itens a serem preenchidos no formulário para análise quantitativa dos trabalhos selecionados	41
Tabela 12 - Taxonomias pesquisadas.....	44
Tabela 13 - Definição dos Termos da Taxonomia desenvolvida.....	46
Tabela 14 - Conceitos que definem a qualidade dos resultados para a pesquisa.....	51
Tabela 15 - Relação entre valores e qualidade dos itens da lista de avaliação do mapeamento sistemático.....	52
Tabela 16 - Formulário para Avaliação Qualitativa dos estudos selecionados	53
Tabela 17 - Intervalos das qualidades dos trabalhos avaliados	76
Tabela 18 - Formulário para Avaliação do Mapeamento Sistemático.....	77
Tabela 19 - Periódicos selecionados para a divulgação do Mapeamento Sistemático	78

SUMÁRIO

Capítulo 1 - Introdução	1
1.1. Motivação.....	2
1.2. Objetivo Geral	3
1.3. Objetivos Específicos	4
1.4. Estrutura da Dissertação	4
Capítulo 2 - Referencial Teórico.....	5
2.1. Ontologias	5
2.1.1. Classificação de Ontologias	8
2.1.2. Linguagens para Representação de Ontologias	9
2.1.3. Metodologias para a criação de ontologias	12
Capítulo 3 - Informática Médica	15
3.1. Dificuldades.....	18
3.2. Áreas de Atuação da Informática Médica.....	19
Capítulo 4 - Mapeamento Sistemático.....	24
Capítulo 5 - Metodologia	27
5.1. Planejamento do Mapeamento Sistemático	27
5.2. Classificação da Pesquisa	31
5.3. Seleção dos Estudos Relevantes	34
5.4. Seleção dos Estudos Potencialmente Relevantes	37
5.5. Critérios de Inclusão e Exclusão	38
5.6. Extração de Dados	40
5.6.1. Criação da taxonomia utilizada na pesquisa	42
5.6.1.1. Resultado da análise comparativa entre as taxonomias pesquisadas	45
5.7. Extração dos Dados.....	49
5.7.1. Avaliação da qualidade dos trabalhos	50
Capítulo 6 - Síntese dos Dados e Apresentação dos Resultados	54
6.1. Estatísticas Básicas.....	54
6.2. Mapeamento das Evidências	57
6.3. Mapa das evidências investigadas	71
6.4. Mapeamento das Evidências quanto as subáreas da Informática Médica	72
6.5. Avaliação de qualidade dos trabalhos.....	75
6.6. Avaliação do Mapeamento Sistemático	77
Capítulo 7 - Conclusões.....	79

7.1. Principais Contribuições	82
7.2. Trabalhos Futuros	84
7.3. Limitações deste Trabalho	84
Capítulo 8 - Referências	86
APÊNDICE A.....	95
ANEXO I	99
ANEXO II	111

Capítulo 1 - Introdução

Informática Médica surge como uma nova área de pesquisa que combina a Ciência da Computação, Tecnologias da Informação e Medicina. Integra elementos como Ciência da Informação, Tecnologia Informática e Históricos Médicos relacionados à saúde e serviços, (SILACHAN, 2011).

A Informática Médica lida com o armazenamento, recuperação, uso da informação, dados e conhecimentos biomédicos para a resolução de problemas e tomadas de decisão. Dentre as linhas de pesquisas na área de Informática Médica, podemos destacar: digitalização de prontuários médicos, extração de informações em prontuários digitais, reconhecimento de documentos em linguagem natural, ferramentas de apoio ao diagnóstico clínico/médico, mineração de dados em bancos médicos, reconhecimento e diagnósticos baseados em imagens (radiografias, mamografias, etc.), etc. Ou seja, é perceptível que as técnicas de Inteligência Artificial são parte integrante da Informática Médica através das linhas de pesquisa relacionadas, principalmente, à mineração de dados, suporte ao diagnóstico clínico, etc.

A Representação do Conhecimento é a área da Inteligência Artificial que lida como o conhecimento pode ser representado simbolicamente e manipulado de forma automática por programas de raciocínio, (RUSSEL, 2009). Na Ciência da Computação, Ontologia é uma técnica de Representação do Conhecimento que se refere a um artefato de engenharia constituído por um vocabulário específico usado para descrever uma determinada realidade e construir uma série de suposições sobre o significado pretendido. Ontologias descrevem uma especificação formal sobre certo domínio, sendo um entendimento compartilhado sobre o mesmo e um modelo formal e executável por máquina.

Neste contexto, as pesquisas em Informática Médica (IM) agregadas a Inteligência Artificial, tem usufruído de inúmeras oportunidades, mediante a utilização de ontologias, para avançar o estado da arte em vários aspectos, principalmente no tocante a diagnósticos mais eficientes de determinadas doenças, redução de erros médicos, entre outros.

1.1. Motivação

É sabido que o diagnóstico médico depende de uma grande quantidade de dados de natureza heterogênea como imagens, texto, etc., (MARAGOUDAKIS, 2008), além de pesquisadores poderem utilizar diversos tipos de nomenclaturas para se referir a um mesmo termo médico, (SARKAR, 2010). Iniciativas como o SNOMED, (MICHAEL, 2010), tem o objetivo de padronizar a terminologia clínica e conceitos médicos, atuando como uma solução para esse tipo de problema. Porém além das nomenclaturas distintas entre os profissionais da área médica, os dados ainda estão sujeitos a sua heterogeneidade natural.

Ontologias são técnicas de Representação do Conhecimento, uma subárea da Inteligência Artificial, dentre suas principais áreas de aplicação, pode-se citar as seguinte, (SMITH, 2001):

- Sistemas de Gerenciamento de Banco de Dados: ontologias começaram a atuar na resolução de problemas de modelagem conceitual em banco de dados. Entre outros usos, podemos citar a utilização de ontologias em técnicas de Mineração de Dados.
- Engenharia de Software: dentre os principais usos das ontologias nesta área da Ciência da Computação, podemos citar a integração de diversas ferramentas, com o objetivo de prover suporte a todo o processo de desenvolvimento de software. Através do compartilhamento de uma mesma conceituação sobre processos de software.

Após sua utilização bem sucedida para estes fins, as ontologias passaram a ser aplicadas em outras áreas da Ciência da Computação bem como da ciência de um modo geral.

Desde então este recurso tem sido utilizado principalmente para: tornar automático o processamento de certos tipos de dados, criar um vocabulário comum entre homem e máquina, ou entre agentes inteligentes, etc.

A partir da evolução das tecnologias da informação e sua utilização crescente por profissionais da medicina, no que diz respeito desde o uso de softwares para o gerenciamento de dados referentes aos pacientes, até o uso de equipamentos que permitam diagnósticos precisos como o aparelho de ressonância magnética, foi

produzida uma grande quantidade de dados que podem revolucionar o diagnóstico médico, (SARKAR, 2010).

A Informática Médica, portanto, trouxe grandes benefícios para o diagnóstico médico, mas ainda está aquém do seu verdadeiro potencial. Tecnologias como as Ontologias facilitam a integração de dados heterogêneos, bem como tornam possível a comunicação entre sistemas diferentes, além de tornar certos tipos de dados passíveis de processamento automático. Este tipo de recurso aplicado a Informática Médica torna possível certas operações antes impraticáveis, por exemplo: o diagnóstico médico baseado em várias fontes diferentes de dados, a digitalização de prontuários médicos, a mineração de dados em bases de clínicas, etc. Diversas pesquisas têm sido desenvolvidas neste sentido, (SARKAR, 2010), (DANG, 2008).

Portanto, devido a importância que tem adquirido torna-se necessário documentar, de maneira eficiente, o estado da arte do uso de Ontologias na Informática Médica, promovendo uma visão detalhada de como estas aplicações são desenvolvidas, quais seus benefícios, lacunas e oportunidades para deste modo contribuir para o crescimento da mesma guiando o desenvolvimento de novos estudos com o objetivo de promover o crescimento geral da Informática Médica.

1.2. Objetivo Geral

O Objetivo desta pesquisa é investigar o conhecimento existente sobre o uso de Ontologias em Informática Médica, e entender quais são os principais usos de Ontologias neste contexto, bem como os benefícios de sua aplicação nas diversas subáreas da Informática Médica.

Não obstante, este estudo tem como objetivo também apresentar um mapeamento do estado da arte relacionado ao uso de Ontologias em Informática Médica, sendo esta nossa principal contribuição, tendo em vista sua utilidade para pesquisadores, professores, alunos e profissionais da área.

1.3. Objetivos Específicos

A partir do Objetivo Geral desta pesquisa elencamos os seguintes objetivos específicos:

- Pesquisa e seleção de estudos sobre o uso de Ontologias em Informática Médica.
- Avaliação da qualidade dos estudos selecionados.
- Extração de informações relevantes para a pesquisa, bem como a categorização e interpretação dos dados.
- Análise das contribuições deste trabalho e sugestão para trabalhos futuros.
- Apresentação do mapeamento após a combinação dos resultados obtidos.

1.4. Estrutura da Dissertação

Esta dissertação está dividida em 7 Capítulos, de modo que o início da mesma é caracterizado por uma introdução discorrendo sobre os objetivos gerais e específicos deste projeto.

No Capítulo 2, são definidos os conceitos a respeito de Ontologias, como podem ser classificadas, suas formas de representação e linguagens para sua implementação. Por fim as principais metodologias para o desenvolvimento da mesma.

No Capítulo 3 é apresentada o conceito da área de Informática Médica, seus principais benefícios, dificuldades e a especificação de suas sub-áreas.

No Capítulo 4 são definidas as principais metodologias e métodos utilizados para o desenvolvimento deste mapeamento sistemático.

No Capítulo 5, temos a definição de todo o protocolo do Mapeamento Sistemático, classificação desta pesquisa, definição de escopo e extração dos dados.

No Capítulo 6, é apresentada a síntese dos dados, resultados e a discussão dos dados e conclusões.

No Capítulo 7, por fim, são dadas as considerações finais, principais contribuições, juntamente com a perspectiva para trabalhos futuros.

As referências bibliográficas utilizadas nesta pesquisa estão no capítulo homônimo.

Capítulo 2 - Referencial Teórico

Neste Capítulo é apresentada a base teórica necessária para o melhor entendimento e desenvolvimento desta pesquisa através de uma revisão sistemática de literatura. Tem como objetivo apresentar um Mapeamento Sistemático, sobre o tema pesquisado, revelando suas deficiências, necessidades, etc.

2.1. Ontologias

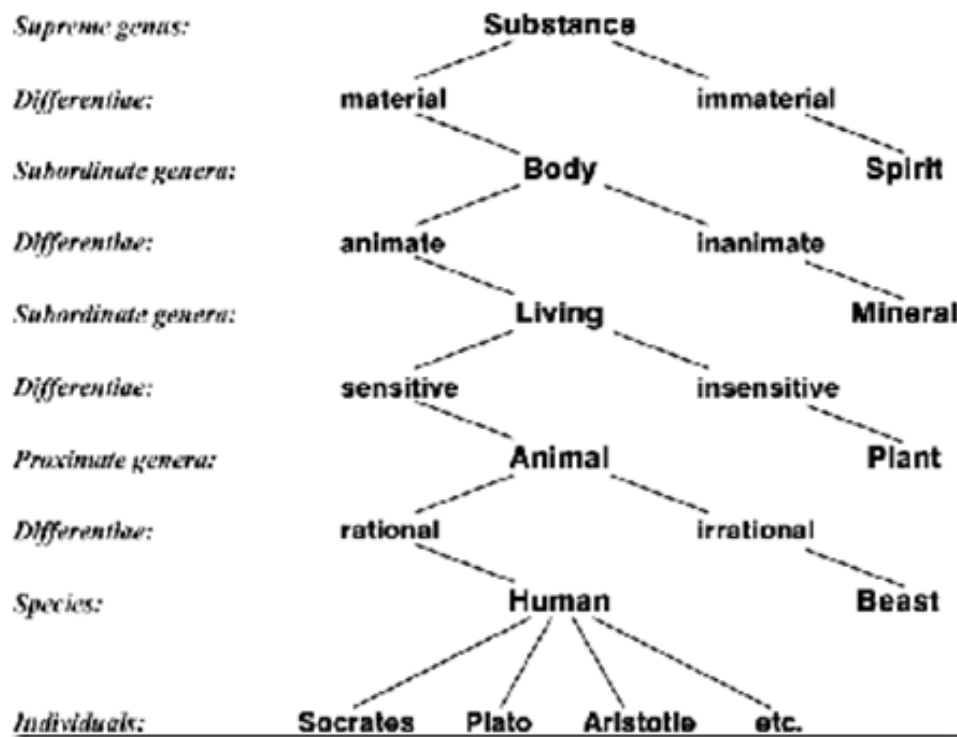
Ontologia, na Filosofia, é a ciência que estuda a natureza e os relacionamentos do ser. Assim como os tipos de estruturas dos objetos, propriedades de eventos, processos e relacionamentos em todas as áreas da realidade. Dito isto, é possível concluir que as ontologias estão, originalmente, dentro da área da metafísica, (BREITMAN, 2001).

O termo ontologia vem do grego *ontos* (ser) + *logos* (palavra), foi cunhado pela primeira vez em 1613, independentemente, por dois pesquisadores, a saber: Rudolf Gückel em sua obra *Lexicon Philosophicum* e por Jacob Lorhard na sua obra *Theatrum philosophicum*, (SMITH, 2001). A primeira ocorrência do termo na língua inglesa foi registrada no Oxford English Dictionary no século XVIII, por volta de 1721, no qual ontologia é definida como “um relato do ser no abstrato”, (SMITH, 2001), (BREITMAN, 2010).

O termo foi introduzido no estudo da Filosofia com dois objetivos: para fazer distinção entre o estudo do ser e o estudo dos diversos tipos de seres vivos existentes, e como disciplina da área de Filosofia como um sistema de categorização para organizar a realidade.

Porém, como é sabido, este não é o primeiro sistema de categorização idealizado na história da ciência, visto que Aristóteles já inventara uma estrutura semelhante. Conhecida como árvore de Porfírio, filósofo grego que tornou conhecida a estrutura criada por Aristóteles, que procura categorizar qualquer objeto como um subgênero do chamado, “gênero supremo” denominado “Substância”. Vale salientar que este é o conceito de uma Taxonomia, vide Figura 1.

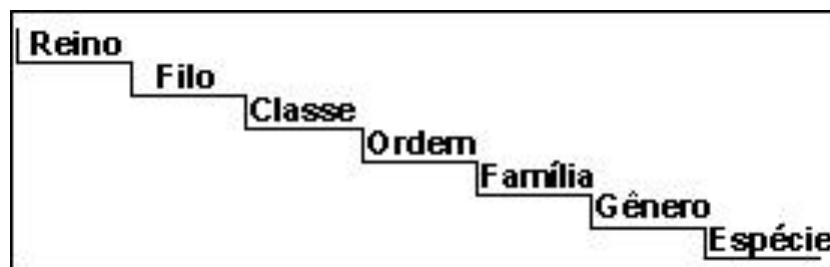
Figura 1 - Árvore de Porfírio



Fonte: Breitman, 2010

A estrutura da Árvore de Porfírio possui semelhanças com a Ontologia quanto às hierarquias dos objetos categorizados. Porém a principal desvantagem do uso da Taxonomia é o tipo de relacionamento entre suas entidades, que segundo Daconta (2003), é uma hierarquia do tipo pai-filho, ou seja, o único tipo de relacionamento existente entre suas entidades é a generalização. Vide Figura 2.

Figura 2 - Hierarquia de uma Taxonomia



Fonte: Autoria própria

Portanto, conclui-se que não há como representar relacionamentos dos tipos: *parte-de*, *causa-efeito*, *localização*, *associação*, etc.

Portanto, de acordo com o que foi dito anteriormente, é possível perceber a diferença entre as ontologias e taxonomias, visto que a primeira possui grande expressividade e deste modo permite a construção de diversos tipos de relacionamento além de um maior nível de detalhamento, (BREITMAN, 2010). Existem diversas definições para o termo Ontologias, entre eles podemos citar seis definições que, segundo Guarino, (1998), estão presentes na literatura. A saber:

- Ontologia é uma disciplina da Filosofia;
- Ontologia é um relato semântico formal;
- Ontologia é uma especificação de uma conceitualização;
- Ontologia é uma representação de um sistema conceitual através de uma teoria da Lógica;
- Ontologia é o vocabulário utilizado por uma teoria lógica;
- Ontologia representa o meta-nível da especificação de uma teoria lógica.

Porém a definição mais aceita, na literatura, sobre ontologias, segundo Breitman, (2010), é a proposta por Gruber:

“Ontologia é uma especificação formal e explícita de uma conceitualização compartilhada.” Gruber, (1995).

Em que, de acordo com, Breitman (2010), conceitualização significa um modelo abstrato de um determinado fenômeno cujos conceitos relevantes para o mesmo são identificados; explícita significa que os elementos e suas restrições estão bem definidos; formal representa que a ontologia é passível de processamento automático; e compartilhada significa que a ontologia é oriunda de um conhecimento consensual, ou seja, aceita por um grupo de indivíduos.

De acordo com o consórcio W3C, (*Ontology for Media Resources 1.0*, 2012), as ontologias devem prover descrições para os seguintes tipos de conceitos:

- Classes: a representação do objeto descrito pela ontologia.
- Relacionamentos: as relações entre as classes.
- Propriedades: atributos das classes.

Ontologias podem ser classificadas de diferentes maneiras, segundo a literatura, (BREITMAN, 2010), podemos citar: Classificação das Ontologias quanto ao seu Espectro Semântico, quanto a sua Generalidade e quanto ao Tipo de Informação de Representam. Dentre as quais foi detalhada o tipo de classificação necessário para o melhor entendimento deste trabalho.

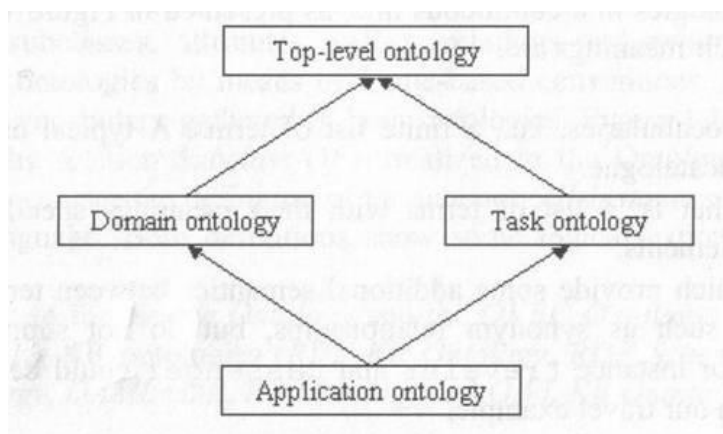
2.1.1. Classificação de Ontologias

Este sistema de classificação utiliza a generalidade da ontologia como principal critério de classificação, (BREITMAN, 2010):

- Ontologias de Nível Superior: descrevem conceitos genéricos, como espaço, tempo e etc. A princípio são independentes de domínio e poderiam ser reutilizadas no desenvolvimento de outras ontologias.
- Ontologias de Domínio: descrevem um domínio específico através a utilização de conceitos especializados presentes na ontologia de nível superior.
- Ontologias de tarefas: descrevem um vocabulário relacionado a uma tarefa genérica através da especialização de ontologias de nível superior.
- Ontologias de Aplicação: são ontologias muito específicas. Os conceitos de aplicação são, de maneira geral, papéis assumidos por entidades de domínio numa determinada tarefa.

A Figura 3 mostra a relação entre as 4 classes propostas por este sistema de categorização de ontologias.

Figura 3 - Classificação de Ontologias de acordo com sua generalidade



Fonte: Gómez-Pérez, Fernandez-López e Corcho, (2004)

Breitman, (2010), afirma que este sistema de categorização foi idealizado por Nicola Guarino, em sua obra, *Formal Ontology and Information Systems*. De acordo com suas conclusões a partir da observação da impossibilidade de relacionar certos tipos de ontologias, Guarino, (1998), sugeriu que ontologias específicas demais poderiam não ter uma relação entre si, mas poderiam estar relacionadas com uma ontologia num nível superior, ou ontologia de alto nível. Deste modo formalizou este sistema de classificação baseando-se no critério da especificidade, tornando possível a relação entre ontologias específicas como as Ontologias de Domínio e Ontologias de Tarefas. Porém, Gómez-Pérez, Fernandez-López e Corcho, (2004), propuseram um sistema de categorização semelhante ao de Guarino, (1998).

Além de existirem modelos diferentes para classificação de ontologias, também existem diversas linguagens para a representação das mesmas. Na última década foram propostas linguagens para a criação de ontologias, linguagens para representação do conhecimento, etc., que apesar de não serem desenvolvidas originalmente para este propósito, também vêm sendo utilizadas para a representação de ontologias, (BREITMAN, 2010).

2.1.2. Linguagens para Representação de Ontologias

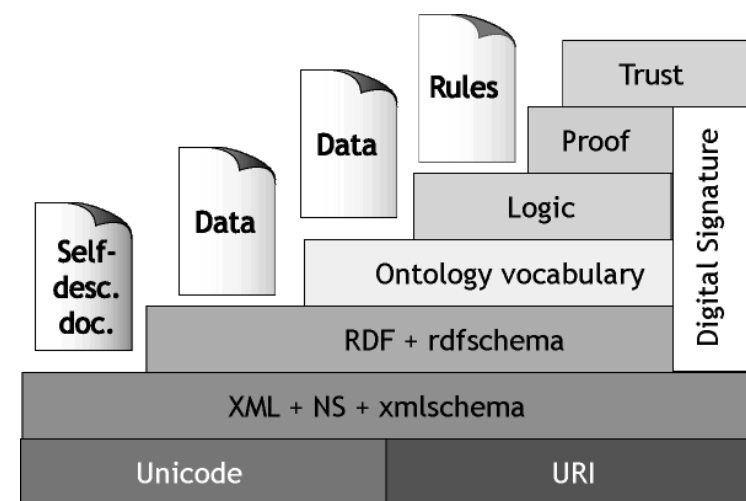
Segundo Breitman, (2010), Tim Berners Lee, na XML 2000 Conference, propôs um modelo em camadas para a arquitetura web do futuro. Esta ficou conhecida como “bolo de noiva”, vide Figura 4. Tem como principal objetivo a reestruturação da Internet, construindo em cima do que já existe. Esse desenvolvimento progressivo e gradual não descartaria as páginas existentes, mas utilizaria estes recursos como base para alcançar a *web* baseada em semântica, ou *web* semântica.

Inicialmente, o termo *web* semântica, no presente trabalho, se mostra deslocado ou fora do escopo do mesmo, contudo se considerarmos que técnicas para representação de conhecimento como ontologias foram desenvolvidas com esta finalidade, bem como as linguagens que as representam, compreendemos que a *web* semântica encontra-se inserida neste contexto.

O *W3 Consortium* (*World Wide Web Consortium*), uma comunidade internacional em que as organizações-membros trabalham juntas para desenvolver padrões para *web*, é liderada por Tim Berners Lee, e seu objetivo é desenvolver o

pleno potencial da *web*, (BREITMAN, 2010). Para tanto definiram as bases para a criação, padronização, utilização e evolução das tecnologias semânticas para *web*.

Figura 4 - Arquitetura Bolo de Noiva proposta por Tim Benners Lee



Fonte: BAILER, 2009.

O W3C (*W3 Consortium*) recomenda a utilização das seguintes linguagens para a representação de ontologias, a saber: RDF e OWL.

- **RDF (*Resource Description Framework*)**: A Linguagem RDF foi desenvolvida pelo W3C com o objetivo de codificar metadados para tornar possível o processamento automático por máquinas, (BREITMAN, 2010), ou seja, fornece interoperabilidade entre aplicativos que trocam informações numa rede. Esta linguagem se baseia nos princípios das redes semânticas para ser utilizada na descrição de recursos disponíveis na *Web*. O RDF fornece primitivas básicas para a criação de ontologias simples, incluindo relacionamentos de generalização para classes e propriedades. No entanto não possui a expressividade necessária para ser propriamente uma linguagem para a representação de ontologias, suportando apenas ontologias simples, ou leves (ou *lightweight*), principalmente as que não necessitem de representações para conectivos lógicos como negação, disjunção e conjunção. Sendo, portanto, bastante limitada quanto ao seu poder de comunicação. O RDF está projetado para fornecer a interoperabilidade e semântica aos metadados de modo a facilitar a busca por recursos na *Web*. Esta linguagem possui três princípios fundamentais: recursos, propriedades e frases, (BREITMAN, 2010). A proposta para qual foi desenvolvida no que diz respeito a representação de metadados, mas não possui recursos

necessários para uma linguagem de ontologias. Para tanto foi desenvolvida, o RDF-Schema (RDFS), que consiste em uma extensão do RDF convencional oferecendo primitivas de modelagem que permitem a construção de hierarquias, classes, subclasses, propriedades e subpropriedades. Foi proposto em 2000, passando por revisões em 2002 e 2003, (BREITMAN, 2010).

O RDF-Schema é utilizado junto com o RDF, atuando como uma espécie de dicionário que possibilita o processamento automático por máquinas. O conjunto é denominado apenas pela sigla RDFS.

- *OWL (Web Ontology Language)*: O consórcio W3C desenvolveu a *Web Ontology Language* (OWL) como uma revisão da linguagem DAML+OIL. Foi projetada para atender as necessidades das aplicações para *Web Semântica*. Dentre seus objetivos principais, podemos citar, segundo Breitman (2010): Construir Ontologias: criar ontologias, bem como explicitar seus conceitos e propriedades e fornecer informações sobre os mesmos; explicitar fatos sobre determinado domínio: Fornecedor informações sobre os objetos que fazem parte do domínio em evidência; racionalizar sobre ontologias e fatos: Ou seja, determinar as conseqüências do que foi construído e explicitado.

Com OWL é possível representar conceitos e seus relacionamentos na forma de uma ontologia. Existem três versões para esta linguagem: OWL Lite – indicada para a criação de hierarquias simples, como as de classificação, e suas restrições, que não possuem estruturas de relacionamento sofisticadas como axiomas. Além de servir como estágio de transição para a próxima versão da OWL; OWL-DL – (DL é o acrônimo para lógica descritiva) essa linguagem pode ser mapeada neste tipo de lógica e, obviamente, possui um nível de expressividade maior que a versão anterior; OWL Full, segundo Breitman, (2010), e McGuinness, (2003), esta versão da linguagem suporta o máximo de expressividade enquanto mantém completude computacional, garantindo tempo finito para todas as computações.

As linguagens RDF e OWL, bem como outras que não foram citadas, são essenciais para o compartilhamento de informação e o reuso de conhecimento sobre domínios específicos em diversas aplicações, como a *Web Semântica*. Porém, linguagens para o desenvolvimento de ontologias não são o único recurso

necessário para o crescimento da área, segundo Breitman, (2010), é notória a necessidade da criação de uma metodologia que formalize o processo de criação de uma ontologia.

2.1.3. Metodologias para a criação de ontologias

Segundo Breitman, (2010), diversos pesquisadores estão buscando a metodologia mais adequada para o desenvolvimento de ontologias para a *Web Semântica*.

De acordo com Guarino e Welty, (2002), a disciplina de ontologia está evoluindo e nesse processo está surgindo claramente a necessidade de uma metodologia.

Porém, de acordo com Breitman, (2010), exista um grande número de variáveis para serem consideradas na criação de uma metodologia para ontologias, deste modo talvez não seja possível desenvolver um único método adequado para todas as situações. E, portanto, a melhor alternativa para a resolução desse problema seja a composição de diversas metodologias.

Existe outra vertente que possui um objetivo mais modesto com relação ao desenvolvimento de metodologias para a criação de ontologias. Através da engenharia, como ciência aplicada, (Breitman, 2010), John Sowa propõe a seguinte alternativa para a resolução deste problema:

“Engenheiros têm um objetivo mais modesto. Ao contrário de procurar a melhor teoria possível, que resolva todos os problemas, ficam satisfeitos com a teoria que é boa o suficiente para o problema específico que têm em mãos.[...]” John Sowa, (2000).

John Sowa, (2000), sugere que não é necessário desenvolver uma solução definitiva para aplicá-la na criação de uma metodologia para o desenvolvimento de ontologias, as soluções que possuímos atualmente são boas o suficiente para o presente problema. Portanto, as diversas metodologias com este propósito existentes são igualmente válidas, bem como podem ser consideradas soluções pontuais para o problema.

São exemplos das principais metodologias para a construção de ontologias existentes:

- Metodologia Cyc: De acordo com Breitman, (2010), o primeiro registro de método ou conjunto de heurísticas para a criação de ontologias foi proposto e publicado como resultados do projeto Cyc (redução de *enCYCclopaedia*, ou enciclopédia). Esse projeto teve como objetivo a construção de uma grande base de dados para armazenar os termos mais gerais da realidade consensual dos humanos, foi desenvolvido pela *Microelectronics and Computer Corporation*, MCC. A base de dados foi alimentada por, cerca de, um milhão de sentenças que foram incluídas manualmente na base de conhecimentos do Cyc.
- Metodologia proposta por Uschold: Também conhecida como *skeletal methodology*, Uschold propôs uma metodologia baseada na técnica de cenários, proposta por John Carrol para design de interfaces. Este processo é composto por quatro estágios distintos, (BREITMAN, 2010). Identificação de propósito da ontologia; Construção; Avaliação; Documentação.
- Metodologia do Projeto TOVE: Esta metodologia, segundo Breitman (2010), é derivada do desenvolvimento de ontologias para os domínios de processos de negócios e corporativos. Outra base para esta metodologia foi a utilização de cenários propostos por Uschold. As etapas para a construção de ontologias com essa metodologia são as seguintes: Descrição de cenários de motivação; Formulação informal das questões de competência; Especificação dos termos da ontologia através de uma linguagem formal; Descrição formal das questões de Competência; Verificação da completude da ontologia.
- Methontology: é um *framework* baseado no processo-padrão IEEE para o desenvolvimento de software, (BREITMAN, 2010), as atividades definidas podem ser separadas em três grupos: Atividades de gerenciamento de ontologias, Atividades ligadas ao desenvolvimento de ontologias e Atividades de manutenção de ontologias.
- Método do Projeto KACTUS: segundo Breitman, (2010), este processo de desenvolvimento de ontologias é baseado no reuso e adaptação de conceitos existentes em outras ontologias. Seu desenvolvimento abrange desde a

especificação da aplicação até o refinamento para promover a evolução da ontologia.

Existem também métodos simplificados para a criação de ontologias como o método 101, proposto por Natalya Noy e Deborah McGuinness, e o método de Ian Horrocks. Porém não faz parte do escopo deste trabalho apresentar e detalhar todos os métodos existentes para o desenvolvimento de ontologias, visto que nenhuma das questões de pesquisa permite um aprofundamento neste tópico.

Capítulo 3 - Informática Médica

Informática médica tem evoluído como uma nova área de pesquisa em um período, relativamente, curto de tempo. Seu surgimento é, em parte, devido aos múltiplos desafios que enfrentam a prática da medicina de hoje. Como exemplo, os médicos precisam: ser mais eficientes, migrar de base de dados em papel para bases digitais, reduzir os erros de diagnóstico e gerenciar e investigar todas as informações relacionadas ao paciente que esteja a seu alcance. Tecnologia tem o potencial para ajudar com cada uma dessas áreas. Com o advento da Internet, computadores de alta velocidade, tecnologias de reconhecimento de voz, tecnologia sem fio e móvel, os profissionais de saúde de hoje têm muitas ferramentas disponíveis à sua disposição (HOYT, 2009).

Informática médica por definição incorpora um conjunto de metodologias que são relacionadas, principalmente, ao gerenciamento de dados, informação e conhecimento aplicado, de um modo geral, na medicina, (SARKAR, 2010). Apesar do grande desenvolvimento recente de aplicações dentro deste domínio, existem registros de pesquisas nesta área desde Shortliffe, (1979) Morris e Casimir, (1987), em que em seus trabalhos eram propostas a utilização de tecnologias computacionais na medicina para o auxílio de diagnóstico, gerenciamento de dados, etc.

A informática médica, também conhecida como informática clínica, informática em saúde, possui três definições clássicas segundo Hoyt, (2009), a saber:

- Campo científico que lida com recursos, dispositivos e métodos formais para otimizar o armazenamento, recuperação e gerenciamento de informações biomédicas para resolução de problemas e tomada de decisão;
- Aplicação de tecnologias da informação em todos os campos da medicina - assistência, educação e pesquisa médica;
- Compreensão e utilização de habilidades e ferramentas que permitem o compartilhamento e uso da informação para fornecer cuidados médicos e promover a saúde.

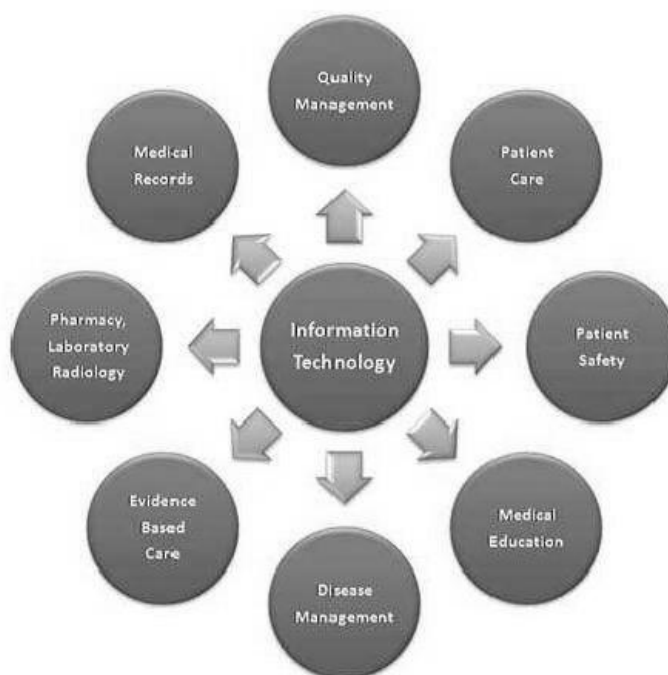
Historicamente a informática médica tem sido dividida entre dois focos distintos: Foco no paciente e o foco no conhecimento. Respectivamente, as aplicações são desenvolvidas tem como objetivo o tratamento do paciente, ou o

gerenciamento do conhecimento existente em determinado escopo, (SARKAR, 2010). Deste modo podem ser identificados os objetivos das aplicações desenvolvidas a partir dos requisitos predefinidos de acordo com o foco para o qual a pesquisa é direcionada.

Segundo Hyot, (2009), a informática médica enfatiza o compartilhamento de grandes quantidades de informações entre pessoas e entidades de saúde, bem como utiliza a tecnologia para tornar a transferência e análise de dados eficiente, confiável e rápida, podemos citar os seguintes exemplos de informações que precisam ser compartilhadas: resultados de exames laboratoriais, resultados de exames de raio-x, estado de vacinação do paciente, medicamentos aos quais o paciente é alérgico, entre outros. Esta área de estudo está frequentemente atuando em conjunto com os campos das ciências clínicas, saúde pública, biblioteconomia, entre outros.

Tecnologias da informação estão sendo utilizadas para realizar muitas funções importantes na área de saúde, atuando como grandes facilitadores para diversas atividades, vide a Figura 5.

Figura 5 - Os principais uso das Tecnologias da Informação em Saúde



Fonte: Hoyt, (2009), p. 3.

Existem muitas forças que impulsionam a adoção das tecnologias da informação em saúde, mas as mais importantes são as necessidades de: diminuir custos, melhorar a segurança do paciente e padronizar e melhorar a qualidade dos cuidados médicos, (HOYT, 2009).

Entre os principais objetivos da informática médica podemos citar, de acordo com Hoyt, (2009), e Siddiqi (2009):

- Reduzir os erros de diagnóstico médico e litígios resultantes;
- Melhorar a comunicação entre agentes, instituições de saúde e pacientes;
- Melhorar a qualidade dos cuidados médicos;
- Reduzir a duplicação de testes ou prescrições ordenadas;
- Melhorar os índices de mortalidade de uma determinada região;
- Padronizar as informações geradas por diferentes médicos e instituições de saúde;
- Melhorar a produtividade do clínico;
- Acelerar o processo de cuidados médicos e operações administrativas por parte das instituições de saúde;
- Proteger a privacidade e garantir a segurança dos pacientes, médicos e instituições;

Além destes objetivos é possível identificar também os interesses das partes que constituem a área de informática médica e listar seus interesses, ou objetivos. Os citados anteriormente podem ser considerados como objetivos comuns a todos os profissionais e instituições que compõem esta nova área de pesquisa. Os grupos de profissionais e instituições identificados por Hoyt (2009) são: Pacientes, Clínicos, Enfermeiros e profissionais de suporte (anestesiologistas, etc.), Saúde Pública, Governo, Educadores, Empresas de Seguro, Hospitais, Pesquisadores e Fornecedores de Tecnologia. No entanto, em oposição aos principais benefícios oferecidos através da concretização dos objetivos citados existem certas dificuldades para a implantação das tecnologias desenvolvidas nas pesquisas dentro do domínio da informática médica.

3.1. Dificuldades

Hoyt, (2009), também cita as principais dificuldades para implantação das tecnologias promissoras desenvolvidas nas pesquisas em informática médica, dentre os quais se destacam:

- Tempo: dentre as principais dificuldades impostas pelos profissionais da medicina quanto a implantação de tecnologias da informação na área de saúde, destaca-se a afirmação de que estes profissionais não tem tempo para aprender a utilizar novas tecnologias em seus ambientes de trabalho, bem como não recebem investimentos para aderir a novas práticas como esta. Quando necessário recorrem a terceirização deste tipo de serviço.
- Custo: tecnologias para análise, armazenamento e transmissão de imagens médicas, por exemplo, ainda possuem um custo elevado. Apesar da diminuição gradual do custo deste tipo de tecnologia gradualmente ser evidente de acordo com o tempo, não são todas as instituições de saúde que possuem recursos para implantar e manter as mesmas.
- Falta de Interoperabilidade: com a existência de diversos modelos para o armazenamento de dados médicos, a interoperabilidade entre sistemas de diferentes instituições apenas com a implantação de padrões. Portanto, padrões precisam ser desenvolvidos e adotados nacional e internacionalmente.
- Mudanças no ambiente de trabalho: mudanças significativas serão necessárias não só no ambiente de trabalho, de modo geral, mas principalmente no modo de trabalho dos profissionais de medicina. Mesmo que o desenvolvimento da informática médica economize tempo e dinheiro, certos profissionais oferecem resistência a mudança de hábitos em seu ambiente de trabalho.
- Privacidade e Legalidade: leis regulamentares quanto ao uso, compartilhamento, transmissão e segurança de dados médicos referentes à pacientes e instituições de saúde precisam ser criadas com o objetivo de proteger os usuários.

No mais, Hyot, (2009), enfatiza o comportamento dos profissionais de medicina quanto à implantação de novas tecnologias em seu modo de trabalho, qualificando-os como céticos no que tange aos benefícios que tais recursos proporcionarão de fato.

É possível também identificar que existem dificuldades específicas de determinadas sub-áreas da informática médica, Hyot, (2009), porém as mesmas não fazem parte do escopo deste trabalho.

3.2. Áreas de Atuação da Informática Médica

Nesta sessão, serão citadas as principais áreas de atuação da informática médica identificadas por Hyot, (2009), a saber:

- Registros Eletrônicos de Saúde: também conhecidos como prontuários eletrônicos, pode ser considerado o principal foco do desenvolvimento da informática médica visto que através dele que são produzidas melhorias na qualidade do tratamento do paciente, análise e recuperação de informação médica. As Ontologias são utilizadas para definição semântica dos termos existentes nos registros eletrônicos, como também promovem uma estruturação semântica de uma base de dados médica.
- Sistemas Integrados de gestão de práticas: estes sistemas funcionam com base nos registros eletrônicos de saúde e tem como um dos principais objetivos é atuar na cobrança de serviços prestados automaticamente pela instituição de saúde. São sistemas administrativos direcionados para a gerência de hospitais, centros de saúde pública, etc.
- Interoperabilidade de sistemas de informação em saúde: pode ser definido como a capacidade de aplicações eletrônicas da área de saúde tem de compartilharem informações. A interoperabilidade neste tipo de sistema é uma característica crítica para o desenvolvimento pleno desta área. O compartilhamento de informações de saúde precisa ser desenvolvido nos níveis local, regional, nacional e internacional. Para tanto é necessária a construção de ontologias que sirvam como ponte entre diferentes sistemas e bases de dados.

- **Sistemas de Auto-Atendimento para Pacientes:** o desenvolvimento de sistemas de auto-atendimento em saúde é um dos objetivos recentes da Informática Médica. Consiste em permitir que o paciente tenha acesso a serviços simples, como certos serviços clínicos, de maneira automática, eficiente e segura, independente de sua localização geográfica, deste modo é necessário um banco de dados semântico bem definido que possa ser utilizados para consultas seguras e eficientes sem o auxílio de um médico.
- **Recursos Médicos Online:** consiste no desenvolvimento de bases de dados de informações médicas semânticas confiáveis, de modo que possam ser consultados para finalidades como educação, recuperação de informação por médicos e pacientes, entre outros. O uso de ontologias é essencial para o desenvolvimento pleno desta subárea.
- **Tecnologias Móveis Aplicadas a Informática Médica:** a evolução e popularização deste tipo de tecnologia tornaram possível o desenvolvimento de aplicações médicas para dispositivos portáteis com a finalidade de melhorar a qualidade de vida de determinadas pessoas. Portanto fornecer informações de maneira rápida e confiável é essencial para que os usuários de tecnologias móveis possam utilizar este tipo de recurso de maneira eficiente, a construção de bases de dados semânticas através de ontologias e aplicativos que possam extrair conhecimento dos mesmos é importante para esta área de aplicação.
- **Medicina Baseada em Evidências:** consiste no auxílio tecnológica a avaliação das subáreas da medicina baseada na busca por evidências. Ou seja, a avaliação da qualidade dos estudos nas diversas áreas da medicina baseada na recuperação e verificação da importância dos mesmos para a evolução da medicina como um todo. O auxílio das tecnologias informáticas é importante visto que boa parte destes estudos tem como seu principal modo de acesso a *web*, através de periódicos eletrônicos, anais de congressos, etc. Portanto o desenvolvimento destas tecnologias especificamente com a finalidade de auxiliar no processo de recuperação de evidências no campo da medicina é importante para esta subárea da informática médica. Dispor destas informações organizadas através de ontologias faz com que esta avaliação seja eficaz e completa.

- Diretrizes para Prática Clínica: consiste em uma das diversas aplicações da Medicina Baseada em Evidência. Baseados nestes resultados gerados são formulados guias de práticas clínicas para tratamento de determinadas doenças ou pacientes em certas condições de saúde. Este processo pode ser realizado de maneira automática com a utilização de sistemas de suporte a decisão, por exemplo. Neste tipo de aplicação as ontologias têm o objetivo de permitir uma busca automática eficaz por guias de práticas clínicas.
- Registro e Gestão de Doenças: o principal objetivo desta subárea é auxiliar na gestão de pacientes crônicos através da utilização dos recursos aqui apresentados, visto que boa parte das aplicações de informática médica são direcionadas para casos críticos. Além de permitir o compartilhamento de conhecimento entre sistemas e bases de dados clínicas promovendo um crescimento no entendimento de diversas doenças, para tanto se faz necessário o uso de ontologias para que seja possível a integração destas informações.
- Pagamento por Desempenho: esta subárea consiste na definição da tributação e o retorno financeiro aos clínicos por aplicações na área clínica, ou médica. Ou seja, consiste no desenvolvimento de modelos de negócio para estes novos tipos de aplicativos. Estes modelos serão aplicados nas diversas plataformas existentes, como dispositivos móveis, terminais de auto-atendimento, entre outros. A definição dos termos de maneira formal e consensual permite o desenvolvimento conjunto de diversos tipos de aplicações com este objetivo.
- Segurança do Paciente e a Tecnologia: o objetivo principal desta subárea é melhorar o tratamento de pacientes em hospitais, inclusive evitando erros médicos. A tecnologia auxilia na segurança do paciente, tornando o tratamento mais eficiente, preciso, etc. Deste modo a utilização de bases de dados semânticas e sistemas que possam processar automaticamente estas informações é essencial.
- Prescrição Médica Eletrônica: além da mortalidade causada por erros médicos, existe a mortalidade provocada por erros farmacêuticos através das prescrições médicas escritas de maneira ilegível pelos profissionais da medicina. As tentativas para eliminar a prescrição médica escrita não

obtiveram sucesso, portanto a alternativa viável é o desenvolvimento de sistemas de prescrição eletrônica automática, principalmente através de ontologias, ou sistemas que reconheçam as prescrições escritas manualmente.

- Telessaúde e Telemedicina: consiste na transferência de dados administrativos relacionado a instituições de saúde e a monitoração remota de pacientes respectivamente. Quanto a Telemedicina, esta tem sido desenvolvida principalmente com o objetivo de permitir o acesso a saúde em comunidades de difícil acesso aos profissionais da medicina. A definição semântica dos dados utilizados se faz necessário para que não existam diagnósticos errados, imprecisos e ineficientes.
- Armazenamento de Imagens e Sistemas de Informação: estes tipos de sistemas facilitam o diagnóstico, elaboração de relatórios, consultas remotas através de imagens médicas. Bem como a transmissão, comunicação, armazenamento destes tipos de dados. No que diz respeito a comunicação, armazenamento e a realização de consultas e diagnósticos o uso de ontologias se faz necessário para tornar estas atividades mais eficientes.
- Bioinformática: consiste na fusão das áreas de biologia, ciências da computação e tecnologia da informação numa única disciplina. Uma das principais aplicações deste domínio é no gerenciamento, manipulação e descobertas de informações genéticas. Inclusive utilizando técnicas de inteligência artificial, armazenamento de dados, etc. O compartilhamento de conhecimento é essencial para o crescimento desta recente subárea, uma vez que estes estudos precisam reunir grandes quantidades de informações que se definidas semanticamente permite uma utilização melhor das mesmas.
- Saúde Pública: subárea direcionada para as populações e não indivíduos. Saúde pública acompanha as tendências na saúde das populações, com o objetivo de prevenir as doenças ou detectá-las precocemente para iniciar o tratamento adequado nas populações estudadas. Construir sistemas de saúde adequados que tenham acesso a diversos tipos de informações clínicas diferentes e sejam capazes de gerenciar e extrair conhecimento destes, é o grande desafio desta subárea. Para tanto as ontologias auxiliam

no gerenciamento, armazenamento, processamento e extração do conhecimento necessário para este tipo de aplicação.

Tendo em vista as principais áreas de atuação apresentadas nesta sessão, é perceptível que boa parte destas possui a necessidade de compartilhar informações, definir padrões para comunicação, etc. Baseado na definição de ontologia e seu detalhamento apresentados no Capítulo 1, podemos identificar diversos usos para esta tecnologia semântica na informática médica.

Portanto, com o objetivo de verificar os possíveis usos de ontologias em informática médica, seus benefícios, o estado da arte desta nova área de pesquisa e sua evolução, foi proposto a realização de um mapeamento sistemático, (KITCHENHAM, 2010), visto que não foi encontrado estudos semelhantes nesta temática.

Capítulo 4 - Mapeamento Sistemático

Estudos sobre engenharia de software baseados em evidências existem desde o final da década de 80, início da década de 90. Este paradigma surgiu nos estudos da área médica visando integrar as melhores evidências de pesquisa com experiências clínicas e avaliação de pacientes, (KITCHENHAM, 2004).

Embora tenha surgido na área da medicina, o paradigma baseado em evidência possui grande aplicabilidade, de modo que é possível utilizar seus princípios nos domínios em que a construção do conhecimento através de revisões sistemáticas estão presentes.

No domínio da Ciência da Computação, a área em que primeiro surgiu à aplicabilidade deste paradigma, foi na Engenharia de Software. Segundo Kitchenham, (2004), a Engenharia de Software Baseada em Evidência consiste na integração das melhores evidências provenientes da pesquisa a experiência prática e valores humanos no processo de tomada de decisão tendo em consideração o desenvolvimento e a manutenção de software.

A Engenharia de Software Baseada em Evidência, tem como um dos seus principais objetivos auxiliar o profissional a adotar tecnologias adequadas para sua pesquisa, evitando as inadequadas, sempre em busca das melhores práticas e procedimentos, (KITCHENHAM, 2004), para que assim seja possível a construção de uma avaliação qualitativa da evolução da área pesquisada.

De acordo com Dyba, (2005), a Engenharia de Software Baseada em Evidência reúne e avalia as evidências existentes através de cinco etapas de uma metodologia, a saber:

1. Transformar o problema ou necessidade em questão de pesquisa;
2. Pesquisar na literatura por melhores evidências disponíveis para responder às perguntas;
3. Avaliar criticamente as evidências, quanto a sua validade, impacto e aplicabilidade;
4. Integrar as evidências avaliadas com experiências práticas, valores e circunstâncias para tomar decisões;

5. Avaliar o desempenho das etapas anteriores com o objetivo de tentar melhorá-las.

Segundo Dyba, (2005), e Kitchenham, (2004), as etapas 2 e 4 são realizadas através de uma revisão sistemática de literatura ou um mapeamento sistemático, dois dos principais métodos utilizados na Engenharia de Software Baseada em Evidência, doravante citada como ESBE. Estes podem ser classificados como tipos de estudos secundários visto que depende de estudos anteriores para construir o conhecimento e agregar evidências. Kitchenham, (2004) define estes métodos como:

- Revisões Sistemáticas de Literatura: identificam, avaliam e interpretam, de forma não enviesada, todas as evidências disponíveis relevantes para servir de apoio para uma determinada questão de investigação científica específica, área temática, ou fenômeno, da seguinte forma: numa população P, a intervenção I, num contexto C, é mais eficaz na obtenção de resultados O do que uma intervenção C?
- Mapeamento Sistemático: tem como objetivo identificar todas as pesquisas relacionadas a um tópico específico, ou seja, responder a questões mais amplas relacionadas à evolução da investigação. Tais questões podem ser classificadas como exploratórias. Por exemplo: O que sabemos sobre o tema X?

Conforme a definição optou-se pelo Mapeamento Sistemático tendo em vista que o seu principal objetivo é identificar todas as pesquisas relacionadas a um tópico específico. Deste modo, tem-se uma visão geral da evolução do que pode ser uma determinada área de pesquisa, apresentando sua evolução até o momento da construção do mapeamento.

Não obstante, Arksey e O'Malley, (2005), apresentam quatro razões que justificam a realização de um Mapeamento Sistemático, a saber:

- Estudar a extensão, natureza, alcance e natureza do fenômeno a ser investigado. O mapeamento é considerado uma maneira útil verificar a evolução de certas áreas de estudo, em que é difícil visualizar a variedade de materiais disponíveis.

- Determinar a necessidade da elaboração de uma completa revisão sistemática de literatura. Podendo ser considerado, portanto, uma etapa preliminar para a realização de uma revisão sistemática completa de literatura, identificando a relevância de sua elaboração.
- Resumir e divulgar o resultado de pesquisas. Descrevendo de maneira detalhada os resultados e abrangência da pesquisa em determinadas áreas de estudo.
- Identificar lacunas nas pesquisas da literatura existente. Este tipo de mapeamento sistemático leva à disseminação de conclusões sobre o estado global da área investigada.

Portanto, fundamentado sobre os objetivos e motivação deste trabalho, como também na definição de Mapeamento Sistemático (KITCHENHAM, 2004) e nos motivos apresentados por Arksey e O'Malley, (2005), decidimos utilizar deste método para que fosse possível estudar a evolução da área e mapear o estado da arte em Ontologias na Informática Médica.

Capítulo 5 - Metodologia

A execução desta pesquisa foi baseada nas diretrizes para o desenvolvimento de um Mapeamento Sistemático recomendadas por Kitchenham (2004), outros trabalhos fundamentais foram os de Jacinto (2010), Almeida (2011), Suassuna (2011), Cardozo (2012), bem como os trabalhos de Condori-Fernandez *et. al* (2009), Brereton e Budgen (2006), todos contribuíram para a realização desta pesquisa principalmente no que diz respeito a execução das diferentes etapas do Mapeamento Sistemático e no reporte dos dados e análise dos resultados.

5.1. Planejamento do Mapeamento Sistemático

Segundo Kitchenham, (2007), o mapeamento sistemático decorre da necessidade de resumir toda informação existente sobre determinada área de estudo de maneira completa e imparcial. Para tanto, é necessária a definição de um protocolo para guiar o desenvolvimento do mapeamento.

O protocolo de um Mapeamento Sistemático deve ser seguido rigorosamente, também é preciso que seja imparcial e passível de auditoria, Travassos, (2007), além de permitir que esta pesquisa seja replicada por outros autores.

As atividades do protocolo podem ser divididas em três etapas, Kitchenham, (2007), a saber: Planejamento, Condução e Reporte do Mapeamento. Detalhamos cada uma destas etapas segundo Kitchenham, (2007):

Etapa de Planejamento:

- Identificação da necessidade de uma revisão;
- Definição do tema sobre o qual será executado o mapeamento;
- Especificação das Questões de Pesquisa;
- Desenvolvimento do protocolo de pesquisa;
- Avaliação do protocolo de pesquisa.

Etapa da Condução:

- Identificação do tipo de pesquisa;
- Seleção dos estudos primários;
- Avaliação da qualidade da revisão;
- Extração de dados e monitoramento;

- Síntese dos dados.

Etapa de Reporte:

- Especificação dos mecanismos de reporte;
- Formatação do Mapeamento Sistemático;
- Avaliação final do Mapeamento Sistemático.

Dentro da etapa de Planejamento do Mapeamento, a atividade mais importante, de acordo com Kitchenham, (2007), é a criação das questões de pesquisa. Elas conduzem toda a metodologia do estudo, e é através delas que: são identificados os estudos primários relacionados à pesquisa; durante a extração de dados as questões de pesquisa atuam como fatores de inclusão e exclusão, visto que apenas os dados que estão relacionados com elas serão avaliados; por fim, a avaliação dos dados selecionados também é realizada tendo como base as questões de pesquisa.

As questões de pesquisa que orientaram este trabalho são:

- QP1 - Qual(is) o(s) principal(is) uso(s) de Ontologias em Informática Médica?
- QP2 - Qual(is) o(s) benefícios do uso de Ontologias em Informática Médica?
- QP3 - Quais as ontologias utilizadas em Informática Médica?
- QP4 - É considerado o reuso de ontologias nos estudos da área de Informática Médica?
- QP5 - Quais as ferramentas e linguagens utilizadas para a criação e edição de Ontologias?
- QP6 - Quais repositórios de Ontologias são utilizados?
- QP7 - Como os trabalhos de Ontologias em Informática Médica estão sendo validados/avaliados?

A definição do protocolo, de acordo com Kitchenham, (2007), é realizada após o desenvolvimento das questões de pesquisa.

O próximo passo consiste em especificar os engenhos de busca para a execução da busca automática. Os engenhos de busca apresentados na Tabela 1 foram selecionados por possuírem, segundo a literatura pesquisada, Dyba (2005), Kitchenham (2007), Travassos (2007), Brereton (2007), Hoyt (2009), boa parte das pesquisas desenvolvidas no domínio deste trabalho.

Tabela 1 - Engenhos utilizados na busca automática

Engenhos de Busca	Endereços Web
IEEE Xplorer	ieeexplore.ieee.org/
ACM Digital Library	dl.acm.org/dl.cfm
Science Direct	sciencedirect.com/
Google Scholar	scholar.google.com.br/
CiteSeerX	citeseer.uark.edu:8080/citeseerx/index
IET	http://digital-library.theiet.org/

Fonte: Autoria própria

Foi excluído o engenho *IET* pela indisponibilidade de seus artigos, o *CiteSeerX* foi retirado desta pesquisa por não abordar temas que não são de interesse desta pesquisa. A Tabela 2 mostra os citados engenhos e seus respectivos endereços para consulta.

Tabela 2 - Engenhos de busca excluídos

Engenhos Excluídos
CiteSeerX citeseer.uark.edu:8080/citeseerx/index
IET http://digital-library.theiet.org/

A Tabela 3 mostra que também foram definidos periódicos para compor a busca manual que, de acordo com Jacinto (2010), Almeida (2011), Suassuna (2011), Cardozo (2012), são úteis para que o pesquisador faça uma revisão ampla que aborde de fato tudo o que foi desenvolvido relacionado a área investigada. Tem o objetivo de cobrir toda a literatura existente a respeito do tema escolhido, portanto é preciso selecionar os periódicos cujos temas estejam dentro do escopo da pesquisa, (KITCHENHAM, 2007).

Uma vez definidos os engenhos de busca manual e automática é preciso construir a String de Busca, (KITCHENHAM E TRAVASSOS, 2007), de modo que represente o que será abordado na pesquisa baseado nas questões de pesquisa.

Tabela 3 - Periódicos selecionados para a busca manual

Periódicos	Qualis
Applied Ontology	B2
International Journal of Metadata, Semantics and Ontologies	B2
International Journal of Medical Informatics	A2
Journal of Biomedical Informatics	A1
Journal of the American Medical Informatics Association	A1
Medical informatics and the internet in Medicine	B2

Fonte: Autoria própria

A String de Busca é construída a partir da concatenação dos termos encontrados nas questões de pesquisa através de operadores booleanos do tipo “AND” e “OR”, conforme o Quadro 1. Uma vez desenvolvida a String de Busca é possível iniciar a pesquisa nos engenhos selecionados, tendo em vista que cada engenho possui uma sintática própria foi necessário adequar a String de Busca de acordo suas especificações. Porém, segundo com Jacinto (2010), Almeida (2011), Suassuna (2011), Cardozo (2012) não é necessário definir estas adequações.

Quadro 1 - String de Busca Desenvolvida

String de Busca
("Knowledge Engineering" OR "Knowledge Representation" OR "Ontology Engineering") AND ("Health" OR "E-Health" OR "Electronic Health" OR "Medical Informatics" OR "Medical Systems" OR "Medical" OR "Clinical Decision Support" OR "Medical Decision Support") AND ("Ontology" OR "Ontologies")

Para auxiliar a seleção dos estudos relevantes para um Mapeamento Sistemático é preciso que existam critérios de inclusão e exclusão, (Jacinto, 2010; Almeida, 2011; Suassuna, 2011; Cardozo, 2012).

Os seguintes tipos de estudos foram incluídos nesta avaliação sistemática:

- Estudos completos publicados em revistas ou conferências, que tenham sido revisados, sobre o uso de Ontologias em Informática Médica e que respondam uma das questões de pesquisa;
- Estudos secundários, ou seja, que dependam de estudos primários;
- Estudos teóricos com o objetivo de apresentar conceitos para o entendimento da área;
- Estudos experimentais relacionados à área;
- Apenas trabalhos escritos em língua inglesa serão considerados.

Os seguintes estudos serão excluídos:

- Estudos que não estejam claramente relacionados à área de Ontologias e Informática Médica;
- Estudos que não respondem a nenhuma das questões de pesquisa;
- Artigos duplicados, ou seja, aqueles encontrados em mais de uma fonte da busca automática e/ou manual;
- Artigos convidados, tutoriais, relatórios técnicos, teses, dissertações e relatórios de workshops que sejam incompletos;
- Estudos não disponíveis para download nos engines de busca definidos.

Após a conclusão dos passos da etapa de planejamento, cuja última atividade é a definição dos critérios de inclusão e exclusão dos trabalhos, (KITCHENHAM, 2007), foi iniciada a etapa de condução do mapeamento. Esta é caracterizada pela classificação da pesquisa, seleção e extração dos estudos relevantes para este trabalho.

5.2. Classificação da Pesquisa

De acordo com Cooper, (1988), com a finalidade de complementar a descrição metodológica desta pesquisa, a mesma foi classificada segundo a taxonomia proposta para revisões sistemáticas. Esta usa seis características para categorizar este tipo de pesquisa, a saber: foco, objetivo, perspectiva, cobertura, organização e audiência.

Portanto, conforme a taxonomia desenvolvida por Cooper, (1988), esta pesquisa é classificada como mostra a Tabela 4.

Tabela 4 - Classificação desta pesquisa segundo a taxonomia de Cooper

Classificação Metodológica	
Foco	Baseado em Resultados
Objetivo	Integração
Perspectiva	Neutra
Cobertura	Revisão Exaustiva com Seleção
Organização	Conceitual
Audiência	Pesquisadores em Geral

Fonte: Autoria própria

Cooper, (1988) define os elementos de sua taxonomia da seguinte maneira:

- Foco: conforme a análise da taxonomia adotada se concentra nos resultados obtidos dos estudos primários selecionados. Este tipo de revisão é caracterizada pela síntese de trabalhos que é desenvolvida, permitindo chegar a conclusões, bem como identificar lacunas nas pesquisas selecionadas como também da área em geral.
- Objetivo: pode ser definido como a integração dos trabalhos selecionados, tendo em vista o mapeamento da área do Uso de Ontologias em Informática Médica.
- Perspectiva: não faz parte do escopo desta pesquisa validar ou avaliar os trabalhos selecionados. Portanto, é assumida a neutralidade no que diz respeito ao posicionamento dos pesquisadores envolvidos. Pretende-se revelar o estado da arte da área de pesquisa estudada, salientando as principais tendências e lacunas existentes na mesma.
- Cobertura: quanto a este aspecto da pesquisa, tendo em vista seu objetivo de apresentar um mapeamento do estado da arte da área de estudo, porém selecionando os tipos de trabalhos que serão avaliados, esta é classificada como Revisão Exaustiva com Seleção.
- Organização: esta pesquisa é classificada como sendo organizada conceitualmente. Ou seja, uma revisão teórica focada nas teorias presentes na literatura clássica relacionada.
- Audiência: a principal contribuição desta pesquisa é fornecer uma visão geral sobre o estado da arte da área pesquisada. Portanto é destinada a

acadêmicos, de um modo geral, que tenham interesse em desenvolver novos trabalhos dentro deste domínio.

Segundo Randolph, (2009), devido à importância que uma revisão da literatura possui para o desenvolvimento de novos trabalhos de pesquisa, é necessário que o desenvolvimento da mesma esteja alinhado com determinados objetivos para que obtenha resultados coerentes com o proposto. Portanto, é importante que seja seguida uma categorização deste tipo de pesquisa, de modo a produzir conhecimento relevante relacionado à área do estudo.

Esta pesquisa também optou pelo método de abordagem indutiva baseada em dados de natureza qualitativa, conforme a qualificação proposta por Marconi e Lakatos, (2007), vide Tabela 5, em que é possível verificar outras características da pesquisa. A técnica utilizada neste trabalho foi o Estudo de Mapeamento Sistemático, definido por Kitchenham, (2007). Esta técnica foi selecionada tendo em vista a necessidade de resumir toda informação existente sobre determinada área de estudo de maneira completa e imparcial, Kitchenham e Dyba (2004), Dyba, (2005), Kitchenham, (2007), Jacinto, (2010), Almeida, (2011), Suassuna, (2011) e Cardozo, (2012).

Tabela 5 - Classificação da Pesquisa segundo Marconi e Lakatos

Quadro Metodológico	
Método de abordagem	Indutivo
Natureza das variáveis	Qualitativa
Procedimento	Mapeamento Sistemático
Variável Independente	Ontologias
Variável Dependente	Os usos de Ontologias em Informática Médica

Fonte: Autoria própria

Quanto ao método indutivo, segundo Marconi e Lakatos, (2007), classifica-se esta pesquisa deste modo tendo em vista que este método pode ser definido como um processo mental, por intermédio do qual partindo de dados particulares, suficientemente constatados, infere-se uma verdade geral ou universal, não contida nas partes examinadas.

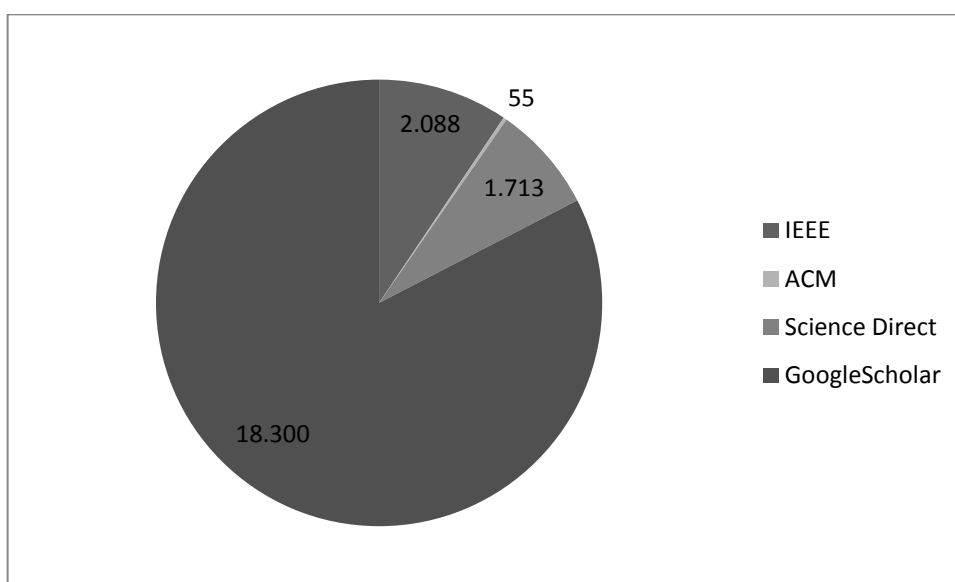
Os métodos de procedimentos consistem na etapa concreta da investigação, o mapeamento sistemático, utilizado neste trabalho. Foi bem definido nas sessões anteriores, bem como as principais motivações responsáveis por esta escolha, de acordo com a literatura consultada, a saber, Kitchenham e Dyba (2004), Dyba, (2005) e Kitchenham, (2007).

A natureza das variáveis, apesar de existir neste trabalho uma análise quantitativa dos dados, é qualitativa. Isso se deve ao fato que o mapeamento será avaliado qualitativamente quanto às reais contribuições, (MARCONI e LAKATOS, 2007), para a área de Ontologias e Informática Médica.

5.3. Seleção dos Estudos Relevantes

Utilizando a *String* de busca do Quadro 1 nos engenhos de busca automática da Tabela 1, seguindo a sintaxe exigida por cada engenho, foram retornados o total de 22.156 trabalhos. Dos quais, 2.088 pertencem ao engenho *IEEE Xplorer*, 55 ao *ACM Digital Library*, 1.713 ao *Science Direct* e 18.300 ao *Google Scholar*. A Figura 6 mostra a distribuição dos trabalhos retornados.

Figura 6 - Distribuição dos trabalhos retornados pelos engenhos de busca



Fonte: Autoria própria

Observando a Figura 6 os trabalhos retornados pelo engenho *ACM Digital Library*, são, aparentemente, irrelevantes no que diz respeito a quantidade diante do volume de trabalhos do *IEEE Xplorer*, *Science Direct* e *Google Scholar*, porém não implica que possuam uma menor relevância para este trabalho como um todo, visto que este gráfico está organizado quanto a quantidade de trabalho, e não quanto a importância dos mesmos para a pesquisa.

Em detrimento a exclusão dos engenhos citados na sessão 5.1., bem como falta de engenhos direcionados para área de Informática Médica, foram adicionados dois novos engenhos de busca durante a execução do Mapeamento Sistemático, Tabela 6.

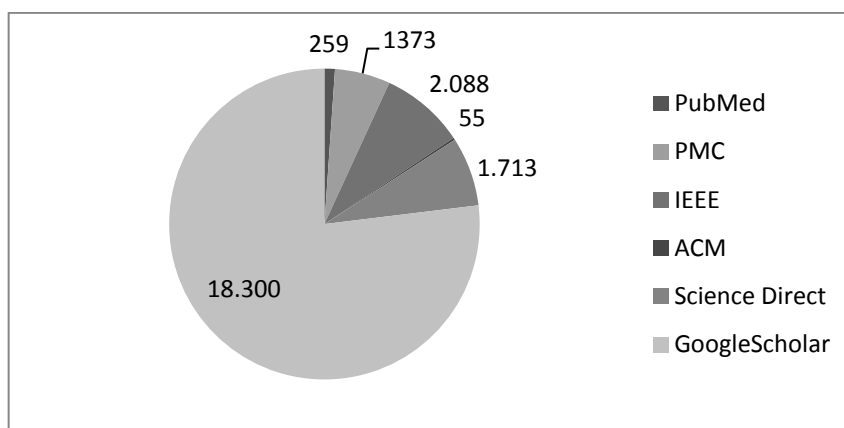
Tabela 6 - Engenhos adicionados posteriormente

Engenhos Adicionados Posteriormente	
PubMed	www.ncbi.nlm.nih.gov/pubmed
PMC	www.ncbi.nlm.nih.gov/pmc/

Apesar da exigência de seguir rigorosamente o protocolo especificado para este Mapeamento Sistemático, de acordo com a literatura (KITCHENHAM, 2007; JACINTO, 2010; ALMEIDA, 2011; SUASSUNA, 2011; CARDOZO, 2012), é possível que o protocolo sofra alterações durante a pesquisa devido as necessidades que possam ser identificadas após a execução do protocolo.

Deste modo a quantidade de artigos retornados no total passou a ser conforme a Figura 7, ou seja, 23.788 retornados.

Figura 7 - Total de artigos retornados pelos engenhos acrescido o PubMed e PMC



Para a seleção destes estudos potencialmente relevantes durante a pesquisa realizada nos engenhos de busca, foram consideradas os seguintes campos: título, resumo, conteúdo do documento e *metadata* (que consiste numa definição semântica dependente do engenho). Considera-se que esta busca poderia ter retornado uma quantidade menor de estudos caso não fossem considerados todos esses campos de busca, trabalhos como os de Jacinto, (2010), Almeida, (2011), Suassuna, (2011), Cardozo, (2012), apenas foram considerados os títulos e resumo. Porém, durante a análise dos artigos, detalhada posteriormente constatou-se que se apenas os títulos e resumos fossem considerados para esta pesquisa certos artigos importantes não teriam sido selecionados.

Quanto aos artigos referentes a busca manual foram retornados 536 artigos no total de todos os periódicos selecionados, porém os mesmos foram classificados como duplicados, ou seja, já haviam sido indexados pelos engenhos de busca automática. De acordo com (KITCHENHAM, 2007) este resultado é plausível para este tipo de pesquisa sistematizada.

A Tabela 7 mostra o percentual de contribuição de cada engenho nesta primeira etapa, a pré-seleção de trabalhos potencialmente relevantes.

Tabela 7 - % sobre o total dos artigos retornados pelos engenhos

Engenhos	Resultados Engenhos	% sobre o total dos engenhos
IEEE Xplorer	2088	8.8%
ACM Digital Library	55	0.23%
Science Direct	1713	7.2%
Google Scholar	18300	77%
PubMed	259	1%
PMC	1373	6%
Total	23.788	≈100%

Fonte: Autoria própria

5.4. Seleção dos Estudos Potencialmente Relevantes

A seleção dos trabalhos relevantes para este Mapeamento foi realizada em duas etapas, a primeira foi através da análise do título, resumo, e introdução, nesta ordem de importância, estes deveriam estar relacionadas ao escopo da pesquisa, bem como responder a pelo menos uma das questões de pesquisa, (KITCHENHAM, 2007; JACINTO, 2010; ALMEIDA, 2011; SUASSUNA, 2011; CARDOZO, 2012), definidas na Sessão 1.2. Os resultados desta etapa podem ser vistos na Tabela 8.

Para tanto uma equipe composta por 3 pesquisadores foi criada, a saber: um aluno de mestrado e dois alunos de iniciação científica, o total dos artigos foi dividido igualmente entre os mesmos. Com o objetivo de eliminar possíveis vieses durante esta seleção, cada trabalho foi avaliado duas vezes por dois pesquisadores diferentes. Se necessário, quando as avaliações se tornavam conflitantes diante de algum critério estabelecido, um terceiro pesquisador era consultado, caso a dúvida permanecesse a mesma era apresentada a orientadora, conforme Beretron, (2007).

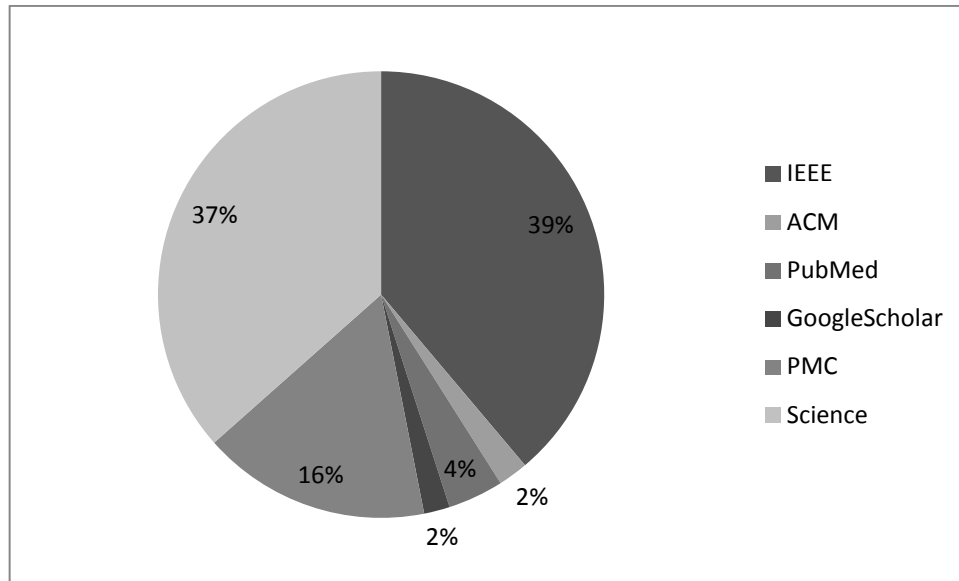
Tabela 8 – Trabalhos selecionados com base análise de título, resumo e introdução

Engenhos	Resultados retornados	Trabalhos Relevantes	% de Aprovação por Engenho
IEEE Xplorer	2088	270	13%
ACM Digital Libray	55	15	27%
Science Direct	1713	28	17%
Google Scholar	18300	13	0.07%
PubMed	259	115	44%
PMC	1373	254	20%
Total	23.788	695	2.9%

Fonte: Autoria própria

A Figura 8 mostra a representatividade de cada engenho após a seleção de artigos relevantes.

Figura 8 - Representatividade dos Engenhos após a seleção de artigos relevantes



Com 28% dos trabalhos considerados relevantes, o *Science Direct* possui uma grande representatividade, porém baseado nos percentuais os engenhos *ACM Digital Library* e *PubMed* foram os que mais retornaram artigos relevantes uma vez que obtiveram os maiores índices de aprovação.

Este é um resultado comum na seleção de artigos potencialmente relevantes em estudos secundários sistemáticos, a saber, revisões e mapeamentos, que utilizam processos automáticos de busca em função das características e funcionalidades dos engenhos disponíveis, (KITCHENHAM, 2007).

5.5. Critérios de Inclusão e Exclusão

A etapa seguinte foi a aplicação dos critérios de inclusão e exclusão nos trabalhos potencialmente relevantes selecionados, (KITCHENHAM E DYBA, 2004), (DYBA, 2005) e (KITCHENHAM, 2007), estes foram previamente definidos na Sessão 1.2., a saber:

- Estudos completos publicados em revistas ou conferências, que tenham sido revisados, sobre o uso de Ontologias em Informática Médica e que respondam uma das questões de pesquisa;
- Estudos secundários, ou seja, que dependam de estudos primários;

- Estudos teóricos com o objetivo de apresentar conceitos para o entendimento da área;
- Estudos experimentais relacionados à área;
- Apenas trabalhos escritos em língua inglesa serão considerados.

Os seguintes estudos serão excluídos:

- Estudos que não estejam claramente relacionados à área de Ontologias e Informática Médica;
- Estudos que não respondem a nenhuma das questões de pesquisa;
- Artigos duplicados, ou seja, aqueles encontrados em mais de uma fonte da busca automática e/ou manual;
- Artigos convidados, tutoriais, relatórios técnicos, teses, dissertações e relatórios de workshops que sejam incompletos;
- Estudos não disponíveis para download nos engenhos de busca definidos.

A mesma equipe foi responsável pela aplicação dos critérios de inclusão e exclusão nos estudos potencialmente relevantes, também foram mantidas as avaliações por 2 pesquisadores diferentes e a ordem de consultas para situações de conflito de avaliações.

Durante a execução desta etapa da pesquisa foram consideradas análise do resumo e introdução, nesta ordem de importância. Deste modo os resultados da Figura 10 foram obtidos.

Tabela 9 - Aplicação dos Critérios de Inclusão e Exclusão

Desenvolvimento da Seleção	Total de Estudos
Total de Estudos Potencialmente Relevantes	695
Aplicação dos Critérios de Inclusão e Exclusão (-)	184
Total de Artigos Relevantes	511

Os artigos excluídos nesta etapa, vide Tabela 9, foram devidos não apenas a aplicação dos critérios, mas também por serem identificadas *surveys* relacionados a área. Segundo (KITCHENHAM E DYBA, 2004), o procedimento que deve ser

adotado na identificação de outras revisões dentre os artigos do Mapeamento Sistemático é reservá-las para comparações após a finalização da pesquisa, considerando que os escopos de ambas sejam semelhantes e passíveis de comparação.

De acordo com o protocolo desenvolvido para esta pesquisa, uma vez que os artigos considerados relevantes tenham sido selecionados a etapa posterior é a da extração de dados.

5.6. Extração de Dados

A referência sobre Mapeamento Sistemático, a saber, Kitchenham e Dyba (2004), Dyba, (2005) e Kitchenham, (2007), definem, de modo geral, procedimentos para a documentação dos resultados da pesquisa manual e utilizando os engenhos de busca, vide Tabela 10.

Tabela 10 - Diretrizes para a documentação dos resultados da pesquisa sugerido por Kitchenham

Fonte de dados	Documentação
Fontes Digitais	• Nome da Base de Dados • Estratégia de Busca Adotada • Data da Realização da Pesquisa • Período de Cobertura da Pesquisa
Jornais e Revistas Científicas	• Nome da publicação • Anos cobertos pela pesquisa • Questões não pesquisadas
Anais de conferências	• Títulos dos Anais • Nome da Conferência (se diferente do título dos anais) • Tradução do Título (se necessário) • Nome do Jornal (se publicado como parte de um jornal)
Esforços para identificar estudos não publicados	• Grupos de Pesquisa e Pesquisadores contatados (Nomes e detalhes do contato)

	• Sites de pesquisa (Data e URL)
Outras fontes	• Data da Pesquisa/Contato • URL • Qualquer condição específica pertinente a pesquisa

Fonte: Kitchenham, (2007), p. 16.

Portanto, baseado nas diretrizes presentes na Tabela 8, sugeridos por Kintchenham (2007), foram definidos os itens a serem preenchidos no formulário para avaliação quantitativa dos trabalhos selecionados na busca automática, vide Tabela 9. Como não foram selecionados artigos durante a busca manual, conforme já justificado na Sessão 5.3., só foi produzido o formulário para a extração de dados dos artigos relevantes selecionados a partir da busca automática.

Este formulário permitiu, entre outros fatores, Kitchenham e Dyba (2004), Dyba, (2005), Kitchenham, (2007), Jacinto, (2010), Almeida, (2011), Suassuna, (2011) e Cardozo, (2012), investigar quais as tendências de pesquisa, as principais lacunas nesta área de pesquisa, os períodos em que a evolução da área foi mais evidente, os países e instituições que produzem mais trabalhos nesta área, e seus principais autores, (KITCHENHAM, 2007).

Tabela 11 - Definição dos itens a serem preenchidos no formulário para análise quantitativa dos trabalhos selecionados

Dados	Descrição
Ano de publicação	Ano de publicação do estudo
Autores	Nomes dos autores
Organização de Pesquisa	Afiliação dos autores do estudo
Tipo do trabalho	Qual o tipo de trabalho analisado
País	Local da Organização
Tópico em Informática Médica	Categorização baseada na taxonomia de Informática Médica
Objetivos	Objetivos do estudo

Fonte: Autoria própria

Para a realização da extração dos dados foi necessária a utilização de uma taxonomia para a classificação dos trabalhos selecionados entre as subáreas existentes, (KITCHENHAM, 2007). A categorização dos artigos selecionados nas subáreas da Informática Médica é essencial para a visualização da evolução das mesmas, de um modo geral, embora esta pesquisa não seja classificada como histórica segundo Cooper, (1988).

Na literatura clássica sobre mapeamento sistemático, Kitchenham e Dyba (2004), Dyba, (2005) e Kitchenham, (2007) e Travassos (2007), é necessário que o pesquisador analise taxonomias existentes, visto que a criação de uma nova categorização de subáreas para uma determinada área do conhecimento não é uma atividade trivial, ou seja, deve-se priorizar o reuso de taxonomias. Desta forma evitando problemas relacionados a existência de diversas categorizações, que fragmentam o conhecimento produzido sobre a área em questão.

Portanto, dentre os itens dispostos na Tabela 11, o Tópico em Informática Médica é o menos trivial para ser definido, tendo em vista que a taxonomia precisa ser adequada ao trabalho e não foi encontrada durante a pesquisa uma única taxonomia clássica que categorize todas as subáreas da Informática Médica.

5.6.1. Criação da taxonomia utilizada na pesquisa

Durante a pesquisa foram encontradas taxonomias que definem a Informática Médica, porém, direcionadas para contextos específicos, (Dixon, 2007), como as propostas pelo HIMSS (*Healthcare Information Management System Society*) e pela MESH (*Medical Subject Headings*).

Segundo Dixon (2007), a taxonomia proposta pelo HIMSS aborda uma parte das principais subáreas da Informática Médica, porém, é estritamente direcionada para sistemas de prontuários médicos eletrônicos, desconsiderando certas aplicações da informática na medicina, como telemedicina e o compartilhamento de informações em saúde. Portanto, não foi considerada adequada para classificar os trabalhos selecionados para o Mapeamento Sistemático, uma vez que a proposta é o desenvolvimento de uma taxonomia geral que represente as principais subáreas existentes para esta grande área da ciência. Posteriormente esta hierarquia de termos foi atualizada e expandida pelo MEDLINE, incluindo a terminologia

“tecnologias da informação em saúde”. Na hierarquia proposta pelo *MESH*, (DIXON, 2007), faltam diversos conceitos da Informática Médica.

Portanto diante deste cenário foi identificada a necessidade de criação de uma taxonomia para a classificação dos estudos selecionados. Tendo em vista que não foi encontrada uma taxonomia que se adequasse ao escopo deste trabalho, uma vez que as que foram pesquisadas possuíam focos específicos, (DIXON, 2007). A taxonomia construída foi desenvolvida utilizando como base os estudos de Dixon (2007), a partir de uma comparação entre os termos presentes em outras taxonomias estudadas.

Observando os trabalhos de Feldt *et al.* (2008), que detalham a construção de sua própria taxonomia baseada principalmente em palavras-chaves encontradas nos resumos dos estudos selecionados, e Dixon *et al.* (2007), que sugerem uma categorização dos termos existentes na literatura médica; iniciou-se a definição de uma estratégia para a construção de uma taxonomia que represente as subáreas da Informática Médica. As taxonomias pesquisadas e estudadas estão dispostas na Tabela 12.

Foi percebido que determinadas categorizações encontradas são específicas para região em que a associação, congresso ou conferência se destina, por exemplo, categorias como *Informatics of Traditional Chinese Medicine* só foram identificadas na *Asia Pacific Association for Medical Informatics Conference*.

Também é possível notar que na taxonomia proposta pela *European Federation for Medical Informatics* existem categorias diferentes direcionadas para a mesma subárea da Informática Médica, mas com focos específicos, por exemplo: *Knowledge discovery in biomedical databases for decision support*, *Intelligent interoperability and telemedicine Data and knowledge management for decision support in forensic medical disciplines*, *eHealth decision support systems for GPs, clinicians, nurses, health care managers and patients*, *Evaluation of decision support systems Diagnostic, therapeutic and prognostic decision support*; ou seja, essas categorias podem ser estruturadas numa hierarquia da subárea *Clinical Decision Support*.

Tabela 12 - Taxonomias pesquisadas

Fonte	Sigla	Tipo
<i>European Federation for Medical Informatics</i>	EFMI	Federação
Sociedade Brasileira de Informática em Saúde	SBIS	Sociedade
<i>American Medical Informatics Association</i>	AMIA	Associação
Associação Portuguesa de Informática Médica	APIM	Associação
<i>Medical Informatics - Practical Guide for the Healthcare Professional, Hoyt (2009)</i>	X	Livro
<i>China-Japan-Korea Joint Symposium on Medical Informatics</i>	CJKMI	Symposium
<i>Asia Pacific Association for Medical Informatics Conference</i>	APAMI	Associação
<i>South African Health Informatics Association Conference</i>	SAHIAC	Associação

Posteriormente, foram selecionados os termos semelhantes ou correlacionados entre as taxonomias, de acordo com sua definição nas fontes citadas anteriormente, (FELDT, 2008), (RICHMOND, 2003).

Uma vez selecionados, os mesmos foram agregados para a construção de uma única categoria geral. Por exemplo, os seguintes termos: Prontuário Eletrônico do Paciente (SBIS), *Electronic Medical Records/Electronic Health Records* (CJKMI), foram agregados ao termo *Electronic Health Records*, Hoyt (2009), por considerarmos este como o termo mais abrangente e que poderia classificar qualquer trabalho direcionado ao desenvolvimento de Registros Eletrônicos aplicados a Saúde.

5.6.1.1. Resultado da análise comparativa entre as taxonomias pesquisadas

Categorias como *Electronic Prescribing* e *Mobile Technology* encontradas em Hoyt, (2009) foram aglutinadas num só termo, *Consumer Health Informatics*. Este novo termo reúne todo estudo que desenvolva produtos comerciais, ou não, voltados para pacientes. Esta decisão está relacionada à constatação durante a pesquisa de que trabalhos nestas áreas são desenvolvidos para esta finalidade.

Internet em Saúde (SBIS), *Networks* (HOYT 2009), *Online Medical Resources* (HOYT 2009), *Search Engines* (HOYT, 2009), *Internet of Things* (APAMI) e *Cloud Computing* (APAMI) foram reunidas na categoria *Online Medical Resources*. Dessa forma, os estudos relacionados aos usos de recursos web na Informática Médica são classificados nesta categoria, bem como serviços online.

As categorias *Formalization of knowledge, ontologies, clinical guidelines and standards of health care* (EFMI), *Standard in Health Informatics and cross-language solution* (CJKMI), *Standards in Medical Informatics* (APAMI) e Padronização da Informação em Saúde (SBIS) foram aglutinadas no termo *Standards in Medical Informatics*. Os estudos aqui classificados são direcionados para definição de padrões de grande abrangência para aplicações que utilizam ontologias em Informática Médica.

Information Technology Infrastructure (SAHIAC), *Privacy, Security, and Confidentiality Issues in Medical Information* (APAMI), *Medical safety management* (CJKMI), *Human Resources Management* (SAHIAC) e *Integrated Practice Management Systems* (HOYT 2009) foram agrupados na categoria *Infrastructure of Medical Information Systems*, direcionada para estudos que desenvolvem solução estrutural para sistemas de Informática Médica baseados em ontologias.

Intelligent interoperability and telemedicine (EFMI) e *Health Information Technology Interoperability* (HOYT 2009) foram resumidas na categoria *Health Management Information Systems*. Esta Categoria reunirá estudos direcionados para o desenvolvimento de sistemas de Informática Médica baseados em ontologias. *Pay for Performance* (P4P) (HOYT 2009) e *Financial management for improved health service delivery* (APAMI) foram reunidas no termo *Financial management and Health Service Delivery of Medical Informatics Resources*, cujo objetivo é reunir os

trabalhos voltados para a administração de recursos financeiros em sistemas de informática médica e também para o desenvolvimento de serviços.

O termo *Informatics of Traditional Chinese Medicine* foi dividido entre diferentes categoria, dependendo do uso dado a ontologias para este tipo de medicina tradicional. Um dos termos que mais classificou artigos deste tipo foi *Ontology Engineering in Medical Informatics*, dedicada a criação de ontologias, visto que em boa parte destes estudos foram criadas novas ontologias com o objetivo de digitalizar o conhecimento disperso neste ramo da medicina. A Tabela 13 contém a definição dos 19 termos criados para compor esta taxonomia

A validação da taxonomia ocorreu na etapa da análise qualitativa, na qual, foi realizada a leitura e posterior classificação de todos os artigos selecionados para o Mapeamento Sistemático de acordo com os termos vistos na Tabela 13.

É importante salientar que Mapeamento Sistemático precisa ser rigorosamente documentado para que, diferente de uma revisão bibliográfica convencional, seja passível de replicação e avaliação. Portanto é necessário que todo o protocolo seja bem definido e transparente para o leitor, (KITCHENHAM, 2007).

Tabela 13 - Definição dos Termos da Taxonomia desenvolvida

Termos da Taxonomia	Definição
<i>Applications of Medical Informatics in Bioinformatics</i>	Bioinformática é uma área da ciência direcionada para aplicação de tecnologias informáticas em biologia, especificamente em genética.
Ontology Engineering in Medical Informatics	O desenvolvimento de ontologias é uma técnica de engenharia do conhecimento, e por isso essa categoria é reservada para todos os estudos que propõem a criação de ontologias, bem com o desenvolvimento de metodologias para validações e análises das mesmas.
<i>Biomedical Electronic and Robotics Devices</i>	Esta subárea da Informática Médica tem como objetivo o desenvolvimento de

	sistemas biomecânicos ou robóticos que proporcionem melhor qualidade de vida do paciente.
<i>Clinical Informatics</i>	Consiste em uma das diversas aplicações da Informática Médica. Ou seja, na criação, gerenciamento, compartilhamento e aplicação de guias de práticas clínicas
<i>Decision Support Systems</i>	Qualquer software projetado para ajudar diretamente na tomada de decisão clínica em que as características de cada paciente são compatíveis com uma base de conhecimento informatizada.
<i>Disease Management and Disease Registries</i>	O principal objetivo desta subárea é auxiliar na gestão de pacientes crônicos através da utilização dos recursos aqui apresentados.
<i>Electronic Health Records</i>	Conhecidos como prontuários eletrônicos, pode ser considerado o principal foco do desenvolvimento da informática médica.
<i>Evidence Based Medicine</i>	Consiste no auxílio tecnológico a avaliação das subáreas da medicina baseada na busca por evidências. Ou seja, a avaliação da qualidade dos estudos nas diversas áreas da medicina
<i>Financial management and Health Service Delivery of Medical Informatics Resources</i>	Esta subárea consiste na definição da tributação e o retorno financeiro aos clínicos por aplicações e serviço na área clínica, ou médica. E o desenvolvimento de aplicações do tipo <i>Service Delivery</i> .
<i>Medical Informatics in education</i>	Consiste em estudos que utilizam ontologias com o objetivos educacionais direcionados para Informática Médica.
<i>Health Management in Medical Informatics</i>	Categoria direcionada para estudos cujo foco é o desenvolvimento e

	gerenciamento de recursos na área de saúde.
Health Information Systems	Esta categoria é dedicada ao desenvolvimento de sistema de informação voltados para a área de saúde.
Infrastructure of Medical Information Systems	Esta categoria é direcionada para estudos que propõe melhorias nas infra-estruturas dos sistemas em Informática Médica.
Medical Signal, Image and Data Annotation, Processing, Representation, Storage, Mining, Visualization, and Sharing	Subárea voltada para o processamento e análise de sinais, dados e imagens médicas, além de abranger a transmissão, armazenamento e formato de dados desta natureza.
Online Medical Informatics Resources	Consiste no desenvolvimento de bases de dados de informações médicas confiáveis, de modo que possam ser consultados para finalidades como educação, recuperação de informação por médicos e pacientes, entre outros.
Patient Informatics	Estudos cujo objetivo é melhorar o tratamento de pacientes em hospitais crônicos e terminais, inclusive evitando erros médicos.
Public Health and Medical Informatics	Subárea direcionada para uso de Informática Médica direcionada para as populações e não indivíduos.
Standards in Medical Informatics	A definição de padrões para a área da Informática Médica é importante para a interoperabilidade de sistemas diferentes.
Telemedicine	Consiste na transferência de dados administrativos relacionado a instituições de saúde e a monitoração remota de pacientes respectivamente.

Porém também é sugerido por Brereton (2007), que determinados detalhes precisam ser abordados quanto a estratégia de pesquisa e o desenvolvimento do protocolo da mesma, dentre eles podemos citar:

- Diferentes estratégias de busca permitem ao pesquisador alcançar diversos objetivos em sua pesquisa, ou seja, o protocolo deve ser rigoroso, porém maleável de modo a permitir sua adaptação de acordo com as necessidades da pesquisa. O pesquisador deve, portanto, desenvolver uma estratégia de pesquisa que seja apropriada para a resolução das suas questões de pesquisa.

E de maneira semelhante Kitchenham, (2007), Travassos, (2007), afirma que o Mapeamento Sistemático é um processo iterativo e evolutivo. Seu protocolo precisa ser bem definido, transparente e rigorosamente seguido. Porém, quando necessário, o mesmo precisa ser avaliado e modificado conforme as necessidades da pesquisa de modo que alcance os objetivos pretendidos. Portanto, é comum que o protocolo do Mapeamento Sistemático definido inicialmente seja alterado conforme a pesquisa é executada.

Dito isto, pode-se afirmar que a Taxonomia pode ser modificada mesmo após a conclusão do Mapeamento Sistemático caso seja identificada a necessidade uma vez ela apesar de ter sido desenvolvida para o presente trabalho pode ser utilizada para a classificação de estudos em Informática Médica de maneira geral.

5.7. Extração dos Dados

A extração foi realizada tendo como guia o formulário de extração de dados implementado em Microsoft Excel™, como pode ser visto na Figura 9.

Este formulário foi baseado nas diretrizes definidas por (KITCHENHAM, 2007) vistas na Tabela 9. As questões de pesquisa tiveram que ser modificadas para se adequar a extração de informações individual de cada estudo selecionado, a saber:

- QP1 - Qual(is) o(s) uso(s) de Ontologias neste trabalho?
- QP2 - Qual(is) o(s) benefícios do uso de Ontologias neste trabalho?
- QP3 - Qual(is) a(s) ontologia(s) utilizada(s)?

- QP4 - É considerado o reuso de ontologias neste projeto?
- QP5 - Qual(is) a(s) ferramenta(s) utilizada(s) para a criação e edição de ontologia(s)?
- QP6 - Quais bancos de ontologias são utilizados?
- QP7 - Como os trabalhos de Ontologias em Informatica Medica estão sendo validados/avaliados?

Figura 9 - Formulário para a extração de dados desenvolvido para esta pesquisa

Formulário para coleta de dados		
ID:		Data da Avaliação:
Área de I.M.:		
Pesquisador:		
Título do Trabalho:		
Autor(es):		
Instituição:		
País da Instituição:		
Fonte de Pesquisa:		Ano da Publicação:
Questões de Avaliação		
QP1	Qual(is) o(s) uso(s) de Ontologias neste trabalho?	
QP2	Qual(is) o(s) benefícios do uso de Ontologias neste trabalho?	
QP3	Qual(is) a(s) ontologia(s) utilizada(s)?	
QP4	É considerado o reuso de ontologias neste projeto?	
QP5	Qual(is) a(s) ferramenta(s) utilizada(s) para a criação e edição de ontologia(s)?	
QP6	Quais bancos de ontologias são utilizados?	
QP7	Como os trabalhos de Ontologias em Informatica Medica estão sendo validados/avaliados?	

Após o preenchimento de todos os formulários referentes aos artigos relevantes as respostas individuais foram sintetizadas respondendo assim as Questões de Pesquisa gerais elaboradas previamente.

5.7.1. Avaliação da qualidade dos trabalhos

Dentre outros benefícios que a análise qualitativa produz para o trabalho, podemos citar, (Kitchenham, 2007):

- Avaliar características, como os critérios de inclusão e exclusão adotados para esta pesquisa;

- Permite investigar as prováveis diferentes qualidades entre os resultados obtidos no estudo;
- Averiguar a importância dos estudos individuais para a síntese dos dados;
- Guiar a interpretação dos resultados e determinar as inferências feitas a este respeito;
- Guiar as recomendações para futuras pesquisas.

Segundo Kitchenham, (2007), existe uma dificuldade inicial quanto ao que representa a qualidade para a pesquisa. Pesquisadores como Khan, (2001), sugerem que a qualidade de uma pesquisa pode ser definida através da relação entre três termos, vide Tabela 14.

Tabela 14 - Conceitos que definem a qualidade dos resultados para a pesquisa

Termos	Sinônimo	Definição
Viés	Erro sistemático	A tendência para produzir resultados que se desviam sistematicamente dos verdadeiros resultados.
Validade Interna	Validade	Se os resultados encontrados são relevantes para a pesquisa. Ou seja, se respondem as questões de pesquisa.
Validade Externa	Aplicabilidade, Generalidade	Verifica se os resultados encontrados são aplicáveis fora do estudo.

De modo que, segundo Khan, (2001), a relação ideal entre estes termos é alcançada quando o Viés é minimizado e a Validade Interna e Externa são maximizadas. Porém esse não é o único método para avaliação da qualidade de um mapeamento sistemático, avaliações de qualidade são baseadas, geralmente, em instrumentos de qualidade, que são listas de verificação de características que precisam estar presentes em um estudo. Estas características podem ser avaliadas de maneira numérica ou nominal.

De acordo com Kitchenham, (2007), existem muitas formas diferentes para a avaliação de qualidade de estudos de Mapeamento Sistemático. Autores como Petticrew, (2005), e Fink, (2005), bem como, Greenhalgh, (2000), em seus trabalhos definem diferentes listas para este tipo de avaliação.

Fink, (2005), sugere que todos os estudos que compõem um Mapeamento Sistemático devem ser avaliados igualmente, portanto, os autores deste tipo de

estudo devem selecionar os itens a serem avaliados que pertencem ao domínio do estudo. Também é necessário que seja desenvolvido uma escala com o objetivo de quantificar cada item do formulário.

Formulários que buscam avaliar a qualidade dos estudos selecionados para o Mapeamento Sistemático que apenas possuem uma escala binária, ou seja, cujas avaliações são apenas positivas ou negativas tende a revelar resultados enganosos e não confiáveis, (FINK, 2005). Dessa maneira, é importante, que a avaliação dos trabalhos seja bem definida antes que o mesmo seja publicado como uma visão geral de uma determinada área. A Tabela 15 mostra um exemplo de índice sugerido para avaliação da qualidade dos estudos selecionados para o Mapeamento Sistemático, de acordo com Kitchenham, (2007).

Tabela 15 - Relação entre valores e qualidade dos itens da lista de avaliação do mapeamento sistemático

Qualidade	Valor Numérico
Pobre	0
Baixa	0.33
Boa	0.67
Excelente	1.0

Fonte: Autoria própria

As questões definidas para a avaliação dos trabalhos selecionadas nesta pesquisa foram respondidas de maneira numérica, de acordo com a Tabela 16.

As perguntas foram elaboradas de modo que a avaliação qualitativa dos estudos é realizada tendo como base o objetivo do estudo, ou seja, serão avaliados os métodos utilizados para alcançar o objetivo final e se o mesmo foi de fato alcançado. Este tipo de avaliação é válida de acordo com Fink (2005), uma vez que estas questões de avaliação são sensíveis ao contexto dentro do domínio dos estudos selecionados. Também é importante relatar que segundo Kitchenham, (2007), a precisão da escala de valores numéricos depende apenas dos autores do Mapeamento Sistemático.

Portanto, observando a Tabela 16, para este trabalho foi considerado como um estudo de excelente qualidade aqueles que atingiram a média máxima igual a um. Obviamente dificilmente um estudo será avaliado com a média mínima tendo em

vista que os artigos selecionados estão de acordo com os critérios de inclusão e exclusão definidos na Sessão 5.2., ou seja, apenas estudos completos, finalizados e revisados.

Tabela 16 - Formulário para Avaliação Qualitativa dos estudos selecionados

Formulário Avaliação Qualitativa	
Avaliador:	
ID do Paper:	Data da Avaliação:
Critérios para Avaliação de Qualidade	Pontuação
Há uma uma descrição clara dos objetivos do estudo?	Não = 0; Em parte = 0,5; Sim = 1,0.
Os objetivos descritos foram alcançados?	Não = 0; Em parte = 0,5; Sim = 1,0.
A metodologia estabelecida para alcançar os objetivos propostos é claramente definida?	Não = 0; Em parte = 0,5; Sim = 1,0.
A metodologia está de acordo com a bibliografia utilizada como base para o desenvolvimento do estudo?	Não = 0; Em parte = 0,5; Sim = 1,0.
Os passos para chegar aos resultados do estudo foram bem documentados?	Não = 0; Em parte = 0,5; Sim = 1,0.
Os resultados divulgados foram satisfatórios para a proposta do estudo?	Não = 0; Em parte = 0,5; Sim = 1,0.
Os resultados são relevantes para a área de pesquisa?	Não = 0; Em parte = 0,5; Sim = 1,0.

Nos capítulos seguintes serão apresentadas as considerações finais, bem como os próximos objetivos para a conclusão do presente trabalho.

Capítulo 6 - Síntese dos Dados e Apresentação dos Resultados

Neste Capítulo serão apresentados a Síntese dos Dados e a Discussão dos resultados após a execução do protocolo definido no capítulo anterior. Este Capítulo está organizado da seguinte maneira:

- **Estatísticas Básicas:** Nesta Sessão serão reportados dados como: quantidade de artigos publicados durante os anos abordados nesta pesquisa, principais autores, países e instituições que mais desenvolveram estudos;
- **Mapeamento das evidências:** Sessão dedicada a analisar as respostas das Questões de Pesquisa;
- **Discussão dos Resultados:** aqui serão discutidos os resultados obtidos no Mapeamento Sistemático.

6.1. Estatísticas Básicas

Inicialmente foi coletada a quantidade de publicações através dos anos, foram encontradas publicações entre os anos de 1986 e 2011. Não foi definido nenhum intervalo entre anos nesta pesquisa, os resultados apresentados a seguir foram extraídos dos estudos considerados como potencialmente relevantes.

Na Figura 12 é possível perceber que a quantidade de publicações na área de Ontologias e Informática Médica têm crescido através dos anos, devido principalmente ao aumento da sua importância no cenário médico e tecnológico atual, (HOYT, 2009) (SARKAR, 2010).

É importante frisar que as publicações do ano de 2012 estão incompletas e não foram consideradas para este Mapeamento Sistemático tendo em vista que o esta pesquisa teve início em 2012 e até então o ano não havia sido concluído e, portanto não seria possível observar todas as publicações do mesmo. Porém, mesmo sem considera-lo para o presente estudo pode-se constatar que já no início de 2012 a quantidade de publicações foi por volta de 64% do ano anterior. Ou seja, é possível que a quantidade de publicações a partir de 2012 seja maiores que as dos anos anteriores.

Também é necessário relatar que para o desenvolvimento deste gráfico foram

utilizados apenas os estudos considerados sem a aplicação dos critérios de inclusão e exclusão.

Figura 10 - Quantidade de Publicações relevantes para a área através dos anos



Nesta pesquisa também foi possível identificar os principais países, pesquisadores e autores que mais publicaram estudos relacionados a Ontologias e Informática Médica, porém não foram considerados para as estas estatísticas foram apenas considerados os últimos 5 anos, ou seja, trabalhos, instituições e autores que possuem estudos publicados entre 2006 e 2011.

Dentre mais de 130 países os que mais publicaram estudos na área de Ontologias e Informática Médica foi: Estados Unidos, China, Espanha, Austrália, França, Alemanha, Reino Unido e Japão.

Também foram encontradas mais de 230 instituições distribuídas entre os mais de 130 países que produziram estudos neste sentido. Nesta estatística básica instituições já conhecidas por trabalhos no domínio desta pesquisa estão presentes, como: Stanford University, University of Freiburg e University of Manchester. Também podemos destacar as seguintes instituições: Dalhousie University, Shanghai University, Clemson University, Deakin University, Hong Kong Polytechnic University, National University of Tainan, Saarland University, University of Sydney, Zhejiang University, etc.

No que diz respeito aos autores de trabalhos sobre Ontologias e Informática Médica foram encontrados mais de 800 pesquisadores, dentre eles estão: Stefan

Schulz, Natalya F. Noy, Elizabeth Chang, Tharam S. Dillon, Daniel L. Rubin, Udo Hahn, David Feng, etc. Também pode-se afirmar que estão de acordo com a nacionalidade das instituições e os países que mais produziram estudos. Portanto, é possível afirmar que os países, as instituições, os pesquisadores citados estão alinhados, confirmando assim que a extração de dados está coerente.

Este tipo de estatística é importante para os pesquisadores do domínio deste trabalho uma vez que outros estudos podem utilizar o mesmo como referência e restringir sua pesquisa de acordo com os autores e instituições que estejam alinhados com suas investigações.

Trabalhos como Gruber (1995) e Guarino (1998), responsáveis pelo conceito e formalização atuais sobre ontologias e representação do conhecimento também foram indexados por esta pesquisa; quanto ao uso de tecnologias baseadas em conhecimento na medicina temos o estudo de Kulikowski, (1986), nessa pesquisa foram relacionadas essas duas áreas da ciência numa retrospectiva de 10 anos mostrando os principais avanços nesta área. O autor relata inovações como, o aprendizado de máquinas auxiliado pela criação de regras baseadas em conhecimento utilizado no sistema Metadendral. É interessante notar que nas referências para o trabalho de Kulikowski, (1986), foram utilizadas referências mais antigas que já propunham a o uso de tecnologias baseadas em conhecimento como em Szolovits, (1982), Kuipers, (1984), Reggia, (1985). Ou seja, é um tema que tem despertado o interesse de pesquisadores num período muito anterior ao registrado nesta pesquisa que não foi indexado pelos engenhos de busca selecionados. Isso pode indicar que estes estudos podem não estar preservados ou foram classificados em outras áreas, uma vez que neste período o termo “Ontologia”, ou “Informática Médica” não era utilizado neste meio.

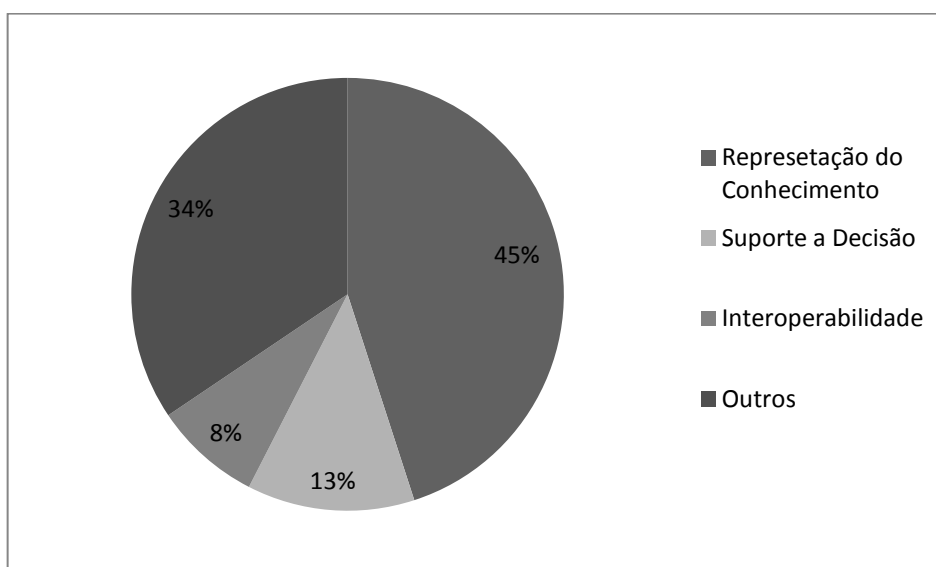
É importante relatar que as seguintes ferramentas foram utilizadas para sintetizar essas estatísticas: o software estatístico R (<http://www.r-project.org/>), Mendeley™ e Microsoft Excel™.

6.2. Mapeamento das Evidências

QP1 - Qual(is) o(s) uso(s) de Ontologias em Informática Médica?

Dentre vários usos de Ontologias em Informática Médica, a partir dos dados analisados neste Mapeamento Sistemático temos a seguinte distribuição, vide Figura 11.

Figura 11- Principais usos de Ontologias em Informática Médica



De acordo com os artigos relevantes para o Mapeamento Sistemático, grande parte dos trabalhos desenvolvidos utilizam ontologias para a representação de conhecimento semanticamente, cerca de 50% dos estudos; outro uso que se apresentou ser um uso comum para ontologias foi para auxiliar ou tornar mais eficiente sistemas de suporte a decisão, 13%; com o objetivo de promover a interoperabilidade entre sistemas diferentes ou mesmo entre ontologias diferentes, apenas 8% dos estudos; por fim 34% dos estudos estão divididos entre 24 outros usos diferentes que não tiveram individualmente uma porcentagem relevante, dentre estes usos temos: Ontologias e Informática Médica aplicadas na Educação, Desenvolvimento de Sistemas Multiagentes baseados em Ontologias, Estruturação de Bases de Dados Médicos.

São exemplos de trabalhos de estudos classificados:

- O trabalho de Cheung *et.al.*, (2007), por exemplo, relata o processo de criação de uma ontologia chamada BrainPharm capaz de representar

conhecimentos sobre neurologia para que a mesma faça parte de um sistema, AlzPharm, que tem como objetivo auxiliar neurologistas na criação de uma banco ontológico neste domínio. Os autores relatam que neste domínio há grande conhecimento distribuído em diversas bases de dados que não são integradas nem interoperáveis resultando numa barreira de difícil transposição, também afirmam que existem outras ontologias de domínio, mas são específicas dentro da neurologia, portanto são pouco reutilizáveis, sendo assim a criação de uma ontologia geral é de grande valia para os neurocientistas. A BrainPharm foi desenvolvida para ser uma proto-ontologia, passível de extensão, de fácil reutilização e compreensão, o banco ontológico em que a mesma estaria hospedada seria também uma espécie de *framework* que auxiliaria pesquisadores na criação de novas ontologias a partir do reuso da BrainPharm.

Embora as perspectivas futuras deste artigo seja criar um banco ontológico de ontologias voltadas para neurologia, o propósito atual é a representação do conhecimento dentro deste domínio, portanto foi classificada nesta categoria.

- O estudo de Altman *et al.*,(2006), propõe a criação de uma ontologia (*RNA Ontology*), bem como a *RNA Consortium* (RNAC) com os objetivos de: integrar diferentes bases de dados de RNA, promover a interoperabilidade entre sistemas distribuídos, servir como base para poderosas ferramentas computacionais com o objetivo de auxiliar os profissionais da área e facilitar as pesquisas sobre informações pertinentes relativas ao RNA. Para tanto, no desenvolvimento da *RNA Ontology* e a RNAC será priorizado a interoperabilidade através de um processo colaborativo e coletivo entre diferentes pesquisadores da área e também reutilizando ontologias já existentes como *Open Biomedical Ontology Consortia*. A interoperabilidade e integração entre bases de dados é o foco deste trabalho, os autores relatam a importância do compartilhamento do conhecimento para o desenvolvimento da área como um todo, além de ressaltar as vantagens de dispor de uma grande base de dados interoperável dentro deste domínio.

Claramente este artigo foi classificado como Interoperabilidade e Integração de bases de dados, embora também seja possível considerar que um dos usos que dá para as ontologias seja a criação de base de dados semânticas,

mas como o principal objetivo do estudo é bem definido pelos autores classificou-se o mesmo na categoria atual.

- Em Fujita, Hakura e Kurematsu, (2011), é desenvolvido um sistema de *Virtual Doctor System* (VDS) interativo que realiza o procedimento de tomada de decisão a partir da visualização do paciente através de câmeras de vídeo, outros sensores e relatos dos sintomas que o paciente está sentindo no momento. O VDS está integrado a uma base dados composta de perfis de pacientes anteriores, diagnósticos médicos e perfis emocionais de pessoas. Para tanto o sistema utiliza duas ontologias, a Physical Ontology e Mental Ontology, que são representações de estados físicos e emocionais de pacientes respectivamente. O algoritmo utilizado é uma rede bayesiana.

O processo de diagnóstico ocorre da seguinte forma, o sistema interage com o paciente colhendo os sintomas que o mesmo sente, além de capturar através dos sensores o perfil emocional do paciente bem como outros sintomas visíveis com o objetivo de confirmar o relato do paciente ou adquirir novas informações. O VDS então busca por diagnósticos que possuam características semelhantes as do paciente para informa-lo o posteriormente. De acordo com os autores, o algoritmo considera o perfil emocional do usuário para tentar identificar outros tipos de sintomas e compara-los com o que o mesmo está relatando.

Também é relatado que todo o sistema é desenvolvido com o auxílio de profissionais da área médica com diferentes especializações e que o VDS não foi desenvolvido para substituir o médico mas para agilizar o processo de consulta e diagnóstico fazendo com que o profissional ganhe tempo e que o paciente obtenha atendimento específico e eficaz.

Obviamente este artigo foi classificado em sistemas de suporte a decisão, seus objetivos são claros e bem definidos, embora seja possível relaciona-lo a outras categorias como desenvolvimento de sistemas de saúde distribuído.

- Crowley e Medvedeva, (2003), reportam o desenvolvimento de uma arquitetura para auxiliar o ensino de diagnóstico de diversas patologias na área de dermatologia. O sistema interage com o estudante da seguinte maneira: o usuário é apresentado a uma imagem de uma determinada patologia, o algoritmo do sistema gera uma árvore de passos que o estudante

deve seguir para diagnosticar corretamente a enfermidade. A medida que o estudante segue os passos ele avança no estágio do diagnóstico e novas árvores de decisão são geradas. Quando todos os passos são seguidos o estudante é bem sucedido e avaliado, caso ele cometa algum erro em certa etapa o sistema auxilia o estudante para que ele decida corretamente qual o próximo passo a ser tomado.

As autoras afirmam que o sistema é reutilizável, o sistema é interoperável e podem ser carregadas novas ontologias além das já presentes poderem ser atualizadas. Portanto este é adaptável ao contexto, ou seja, apto a ser utilizado em qualquer área da medicina. Uma das suas limitações é que ainda é uma ferramenta apenas visual, mas a arquitetura sob a qual foi desenvolvido é de propósito geral e pode ser adequar a outros tipos de ferramentas de auxílio ao ensino.

A ferramenta utiliza ontologias na definição de diferentes patologias dermatológicas, bem como nas atitudes do usuário, nos diferentes estágios de tomada de decisão, etc. As ontologias utilizadas são: *domain ontology*, *domain task ontology* and *case data*.

Creio que o objetivo deste estudo é bem definido, deste modo o mesmo foi incluído na categoria de Ontologias e Informática Médica aplicadas na Educação.

Como se pode perceber os artigos foram classificados de acordo com seus objetivos principais em pelo menos uma das 27 categorias criadas com base nas respostas a QP1. O outro tipo de classificação de artigos é de acordo com a taxonomia desenvolvida para este trabalho, definida na Sessão 5.6. Os trabalhos citados nesta sessão, Cheung *et.al.*, (2007), Altman *et. al.*, (2006), Fujita, Hakura e Kurematsu, (2011), Crowley e Medvedeva, (2003), foram classificados, respectivamente, sob os seguintes termos da taxonomia: *Medical Signal, Image and Data Annotation, Processing, Representation, Storage, Mining, Visualization, and Sharing; Decision Support Systems; Health Information Systems; Medical Informatics in Education*.

É evidente também que os usos das ontologias estão relacionados a sua categoria na taxonomia, porém nem sempre um estudo foi classificado em único

termo da categorização planejada. Este fato justifica a quantidade de respostas comparada com o número de termos da taxonomia.

São destaques também nesta Questão de Pesquisa trabalhos classificados como: *Standards in Medical Informatics*, *Ontology Engineering in Medical Informatics*, a exemplo dos estudos: Schulz, (2011a), (2011b), em que são analisadas ontologias já bastante utilizadas e validadas na comunidade de Informática Médica, mas que contém informações, ambíguas, equivocadas ou erradas. Em Schulz *et. al.*, (2011a), é proposta a ancoragem de certas ontologias com outras *BioTop Ontologies* que possuem a informação representada corretamente. No estudo Schulz *et. al.*, (2011b), a SNOMED CT, é analisada e também são encontradas informações erradas, este problema é mais difícil de resolver por ser uma ontologia muito grande. Para tanto é proposta o uso de um *Ontology Design Pattern* que facilitará a correção dos diversos termos incoerentes dispostos na SNOMED CT.

Em Schulz *et. al.* (2008), é discutido a discriminação entre técnicas de representação do conhecimento difundidas na Informática Médica e comumente confundidas, são elas: ontologias clínicas, terminologias clínicas e *information models*. Posteriormente é proposta a interoperabilidade e integração destas técnicas com o objetivo de reutilizar todo o conhecimento armazenado em diversas bases de dados.

Inicialmente pode-se concluir que diversos estudos tem sido desenvolvidos no sentido de representar diversos tipos de informações médicas, como em Altman *et. al.* (2006), Cheung *et.al.*, (2007), Fujita, Hakura e Kurematsu, (2011). Isto não significa, porém que estejam sendo criadas ontologias diferentes e não está sendo considerado o reuso, os trabalhos de Schulz (2011a), Schulz, (2011b), Schulz *et. al.* (2008) mostra que o reuso é priorizado. Outra tendência interessante é a da criação de design patterns direcionados para a criação de ontologias médicas, ou biológicas, com o objetivo de padronizar o método de criação garantindo sua consistência inclusive em futuras atualizações, (SCHULZ *et. al.*, 2011b).

QP2 - Qual(is) o(s) benefícios do uso de Ontologias neste trabalho?

Além dos principais usos também foi mapeado os benefícios através dos quais os autores dos estudos selecionados justificaram os usos de Ontologias em Informática Médica. Foram identificados 16 benefícios nesta questão de pesquisa, destes 60% dos estudos os autores afirmam que os principais benefícios são os clássicos dispostos na literatura sobre ontologias, (BREITMAN, 2010; GUARINO, 1998), a saber: são uma representação formal, explícita e consensual de conhecimento, passível de processamento automático, recuperação de informações e realização mecanismos de inferência.

Outras vantagens do uso de ontologias em informática médica são: A interoperabilidade e/ou integração entre sistemas diferentes facilitadas pelo uso de ontologias; A possibilidade de compartilhar conhecimento. Que respectivamente correspondem a 13% e 7.2% dos estudos selecionados. Os outros 19.8% estão distribuídos em 13 outras categorias, como: a adição de semântica a bases de dados existentes, diminuir erros durante o processo de extração de informações, etc. São exemplos de trabalho classificados nesta Questão de Pesquisa:

- No trabalho de Chute, Solbrig e Tao, (2011), é atualizada a ontologia *Clinical Narrative Temporal Relation Ontology 2.0* (CNTRO 2.0). Esta ontologia tem o objetivo de representar semanticamente informações temporais em bases de dados clínicas, o rápido crescimento da subárea da Informática Médica conhecida como *electronic health records* (EHR) torna possível a utilização deste tipo de informações em aplicações que tem como base as tecnologias semânticas. De acordo com os autores, esse tipo de informação é essencial para o diagnóstico e tratamento de pacientes acometidos das mais diversas enfermidades. Portanto ter uma ontologia que adiciona esta dimensão à sistemas semânticos é de grande valia para a comunidade da Informática Médica.

A partir deste estudo fica claro a categoria em que foi classificado de acordo com a respostas da QP2, esta ontologia permite que sejam aplicados algoritmos de inferência, a anotação de informações a partir de bases de dados clínicas, etc.

- Em Das *et. al.*, (2006), é desenvolvido um sistema baseado em conhecimento para servir a um grande consórcio de pesquisas, *Immune Tolerance Network* (ITN), uma grande rede de pesquisadores que através da integração e interoperabilidade de sistemas diferentes compartilham e desenvolvem estudos relacionados a *Clinical Trials*. Estes são estudos formais a respeito de condições de pacientes e suas reações a medicamentos durante os tratamentos aos quais são submetidos. O sistema desenvolvido também monitora os pacientes durante três fases distintas, o pré-tratamento, fase do tratamento e pós-tratamento.

O *Epoch* tem como base quatro ontologias que foram criadas para o projeto, estas são reutilizadas em todo a ITN, facilitando assim o compartilhamento e integração dos sistemas. Ele também suporta a integração de informações temporais, e diversas bases de conhecimentos diferentes.

Obviamente este artigos foi classificado, segundo a QP2, que evidencia a integração e interoperabilidade de bases de dados e sistemas através o *Epoch*.

- Bromberg *et. al.*, (2005), cria o *Rat Genome Databases* (RGD). Uma grande base de dados genômica capaz de integrar e compartilhar informações biológicas para que possam ser realizados estudos, utilizadas ferramentas d mineração de dados, criados mecanismos de inferência, etc. Inicialmente seu objetivo é fornecer um framework simples e de fácil entendimento para classificação, representação e navegação em informações genéticas e fenotípicas que relacionam dados genômicos com diferentes doenças. Para tanto, foram usadas quatro ontologias: *Gene Ontology* (GO), *Mammalian Phenotype Ontology* (MP), *Disease Ontology* (DO) e a *PathWay ontology* (PW).

Apesar do objetivo maior de se tornar uma grande base de dados integradora de diversos sistemas, seu objetivo inicial é bem menor e de acordo com a categoria na qual foi classificado. O principal benefício evidenciado nesta pesquisa é o compartilhamento deste tipo de informações, promovendo o crescimento da área e integração de diversas pesquisas de instituições diferentes utilizando a mesma base de dados.

Dentre outros trabalhos, destacamos o de Álvarez-Rodríguez *et. al.*, (2012), que utiliza ontologias aliada a mecanismos de inferência lógica para tornar ferramentas de diagnóstico de doenças psicológicas mais eficientes. De acordo com o autor este meio carece de ontologias que definam patologias psicológicas, as mesmas são de difícil diagnóstico e a definição de ontologias pode trazer benefícios junto a um maior desenvolvimento de sistemas de suporte a decisão neste domínio.

A partir deste estudo é possível perceber o crescimento da necessidade de se compartilhar informações, integrar sistemas diferentes, bem como poder analisar semanticamente o conhecimento clínico, ou médico, armazenado em bases de dados. Diversas pesquisas tem optado pela integração e reutilização de ontologias diferentes com este objetivo, como o trabalho de Schulz *et. al.* (2008) na QP anterior e Bromberg *et. al.*, (2005).

QP3 - Qual(is) a(s) ontologia(s) utilizada(s)?

Foram registradas por volta de 296 ontologias diferentes utilizadas no domínio da Informática Médica. Dentre essas as mais utilizadas são: *Gene Ontology* (8%), *Foundational Model of Anatomy Ontology* (FMA) (6.45%), *Open Biological and Biomedical* (OBO) (1.56%), SNOMED CT (6%) e *GALLEN Ontology* (4.3%), a porcentagem destas ontologias juntas representam 26.31% do total.

Estudos que não relatam qual a ontologia representam 12.45% do total de artigos. Esta postura foi comumente encontrada em trabalhos cuja proposta era o desenvolvimento de sistemas adaptáveis ao contexto, como Crowley e Medvedeva, (2003), entre outros.

Também foram registrados a partir desta Questão de Pesquisa trabalhos que propuseram a criação de novas ontologias foi um total de 15% de artigos. Quando somadas as porcentagens dos trabalhos que reutilizaram as ontologias citadas anteriormente com a dos estudos que não definem uma ontologia específica para seu sistema chegamos a 38.76%, ao comparar esta porcentagem com os estudos que desenvolveram novas ontologias é relativamente pequena. É importante salientar que não se pode afirmar que todos os estudos que criaram novas ontologias não consideram o reuso. Portanto, é possível que haja uma tendência ao reuso de ontologias.

As estatísticas anteriores referentes somadas correspondem a 53.76% do total de artigos analisados, ou seja, então 46.24% estão distribuídos entre 296 ontologias diferentes. Dentre essas podemos citar ontologias conhecidas que não foram tão representativas, mas foram registradas neste estudo, são elas: *Protein Ontology* (CHANG, DILLON, SIDHU, 2005), *OpenEHR Ontology* (FERNANDEZ-BREIS, 2008), *Clinical Pathway Ontology* (YE et. al. 2009), Generic Human Disease Ontology (CHANG, 2004), Breast Cancer Ontology (HU, 2003), etc.

São exemplos de estudos que não relaram a ontologia que foi utilizada, e trabalhos que criaram novas ontologias:

- Maragoudakis, Maglogiannis e Lymberopoulos, (2008), criam uma nova ontologia direcionada para a classificação de lesões de pele. Esta ontologia define diferentes atributos de uma lesão para um sistema de suporte a decisão baseado em imagens. De acordo com os autores foi reutilizada a *Foundational Model of Anatomy Ontology*, durante a criação desta nova ontologia. A mesma posteriormente será integrada a FMA para outros pesquisadores e estudos possam utiliza-la.

Este trabalho é um exemplo dos artigos que propõem a criação de novas ontologias, fica claro que o reuso está sendo considerado portanto bem como a integração com grandes ontologias do domínio da Informática Médica. Ou seja, pode-se concluir que mesmo que 15% dos estudos optem por desenvolver uma nova ontologia, não significa dizer que não será considerado o reuso.

- Abidi e Hussain, (2007), desenvolvem um *framework* baseado em web semântica para auxiliar sistemas de suporte a decisão. De acordo com os autores o procedimento de tomada de decisão depende da integração e avaliação de conhecimento de diferentes fontes e tipos, desenvolver um sistema que auxilie neste processo faz com que a tomada de decisão seja mais eficaz além de promover um maior compartilhamento e integração de dados médicos.

A automatização do procedimento de integração e avaliação de dados médicos é chamado pelos autores de *Knowledge Morphing*. O framework utiliza as ontologias para representar conhecimentos médicos dispersos além de definir os resultados do *Knowledge Morphing*.

O estudo não especifica nenhuma ontologia a ser utilizada, o sistema é adaptável a contextos e pode ser utilizado em qualquer subárea da Informática Médica.

Avaliando os resultados desta Questão de Pesquisa, pode-se afirmar que o reuso de ontologias e o desenvolvimento de sistemas adaptáveis a contextos são duas tendências evidenciadas nesta pesquisa no domínio da Informática Médica. Esta parece ser a próxima etapa diante do cenário atual, em que existem ontologias dedicadas a diversas áreas da medicina.

QP4 - É considerado o reuso de ontologias em Informática Médica?

Segundo a síntese das respostas para esta Questão de Pesquisa é confirmada a tendência de reuso de ontologias, visto que 64.26% dos artigos reusam ontologias existentes para criação de novas ontologias, ou as utilizam integralmente em seus trabalhos. Neste cenário, ontologias como SNOMED CT, *Gene Ontology* e *Foundational Model of Anatomy Ontology*, conforme pôde ser observado na QP3, podem ser consideradas como três grandes repositórios semânticos de definições do conhecimento médico. 20% dos estudos selecionados para este Mapeamento Sistemático utilizam uma delas, se comparado com o número de ontologias (296) descobertas juntamente com o universo de artigos (511), representa uma grande taxa de reuso. Esse tipo de comportamento dos pesquisadores desta área já era o esperado, uma vez que economiza-se tempo, recursos e evita a fragmentação do conhecimento, Breitman, (2010), Hu *et. al.* (2007), Noy (2004), afirmam que este é o cenário ideal.

Do total de trabalhos analisados 24.1% não reutilizaram ontologias existentes, dentre estes destaca-se como exemplo o estudo de Bruc, Magariu e Verlan, (2011), em que foi desenvolvida uma ontologia especificamente para a classificação de órgãos humanos em imagens de ultrassonografias. A *Ultrasound images (UI) ontology* não considerou o reuso. Os autores não relatam também se foram realizadas pesquisas em busca de outras ontologias para serem usadas como base, este é considerado o primeiro passo para este tipo de trabalho, como afirmam Noy (2004), Breitman (2010). Mesmo que não utilizadas completamente pensamos que a SNOMED CT e *Foundational Model of Anatomy Ontology* poderiam ter seus

conceitos reusados neste estudo, porém os autores também não relataram os motivos de não considerarem outras ontologias.

Os outros 11.64% foram classificados como trabalhos teóricos ou diferenciados como os de Aranguren *et. al.*, (2008), que propõe a criação de *Ontology Design Patterns* tendo como base *design patterns* de Engenharia de Software. De acordo com os autores as bio-ontologias são essenciais para o desenvolvimento da Informática Médica, Medicina, Bioinformática e áreas afins, portanto devem representar de maneira fiel e íntegra os conhecimentos do domínio de sua aplicação. A única forma de garantir esta fidelidade é através da criação de metodologias rigorosas que garantam a integridade, fidelidade, completude e corretude da ontologia. O estudo ainda analisa sob este ponto de vista a *Cell Cycle Ontology*, demonstrando que sua metodologia pode ser aplicada para verificar as características que o estudo considera que as bio-ontologias devem ter.

Então é possível concluir que, de acordo com esta QP a maior parte dos estudos desenvolvidos neste domínio considera o reuso de ontologias, este é considerado o cenário ideal por Schulz, *et. al.* (2011^a), Noy (2004), Breitman (2010), entre outros. Porém existem trabalhos como os de Bruc, Magariu e Verlan, (2011), que, apesar de poucos, podem contribuir com a fragmentação do conhecimento e a falta de interoperabilidade entre ontologias.

Também podemos destacar certa tendência no desenvolvimento de estudos sobre metodologias para a criação e avaliação de ontologias, como em Aranguren *et. al.*, (2008), Schulz, *et. al.* (2011^a).

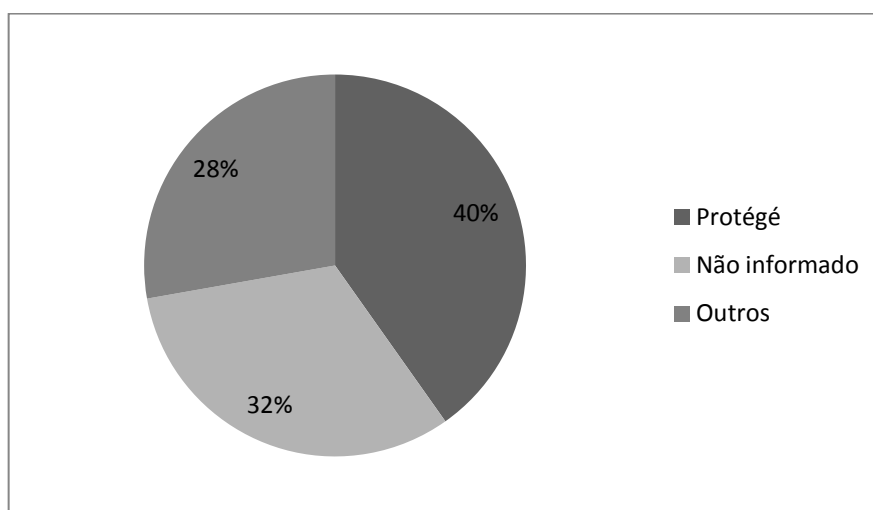
QP5 - Qual(is) a(s) ferramenta(s) e linguagem(ns) utilizada(s) para a criação e edição de ontologias?

Baseado nas respostas para esta Questão de Pesquisa foi possível mapear o cenário apresentado na Figura 12.

Dentre os 28% de estudos que não utilizaram o Protégé, podemos citar as seguintes ferramentas: NeOn Toolkit (KHAN JR., 1998); Java Ontology Editor (JOE), (MAHALINGAM e HUHNS, 1997); Swoop, (SAMWALD ET. AL., 2010); Semantic Editor (ZHEN, *et. al.* 2009); Differential Ontology Editor (DOE), (BARNEYX, CHARLET e JAULENT, 2006).

Os motivos por não utilizarem o Protégé são diversos, desde alegarem dificuldades em utilizá-lo como em Samwald *et. al.*, (2010), ou por julgarem que a ontologia criada será mais simples de integrar ao sistema proposto se for utilizado o Semantic Editor do estudo de Zhen, *et. al.*, (2009).

Figura 12 - Principais Ferramentas utilizadas para criação e edição de ontologias



Quanto as linguagens utilizadas para a representação de ontologias no domínio da Informática Médica, pode-se afirmar que como o Protégé é a ferramenta mais utilizada é provável que as linguagens as quais ela oferece suporte serão as usadas. Ao sintetizar apenas as linguagens mais usadas encontramos a seguinte distribuição: 30% dos estudos publicaram suas ontologias na linguagem OWL; 5% em RDF/RDFS; 3% em XML; 3% em UMLS; por volta de 30% não informaram a linguagem em que a ontologia foi publicada; por fim, 29% utilizaram outras linguagens como DAML+OIL, (HU *et. al.*, 2003), entre outras.

O que pode ser observado também é que a escolha dos editores e linguagens depende apenas do projeto e dos seus autores, algumas vezes justificativas não são relatadas para estas escolhas como em (Barneyx, Charlet e Jaulent, 2006). Nos trabalhos mais antigos como em Khan Jr., (1998), e Mahalingam e Huhns, (1997), o Protégé não é considerado mesmo existindo nesta época, mas não é apresentada uma justificativa pelos autores.

Novamente o cenário é considerado o esperado, está de acordo com as recomendações de Bailer *et. al.* (2012), em que de acordo com as recomendações do W3C, para facilitar a integração entre diferentes bases de dados e a reutilização

das ontologias criadas, é recomendado aos pesquisadores que desenvolvam ontologias em OWL. Também neste documento são destacadas ferramentas como Swoop, NeON Toolkit e Protégé, todas com suporte às principais linguagens recomendadas.

QP6 - Quais bancos de ontologias são utilizados?

Nesta questão obtivemos o seguinte cenário: 76% dos estudos não relataram os bancos de ontologias utilizados na pesquisa e 24% divulgaram as fontes, vide Figura 33.

De acordo com a literatura a respeito do Mapeamento Sistemático, (BRERETON, 2006; BRERETON, 2007; KITCHENHAM, 2007) é possível utilizar as informações adquiridas nas Questões de Pesquisa anteriores para inferir novos conhecimentos posteriormente.

Deste modo, observando a QP3, as seguintes ontologias foram as mais reusadas: *Gene Ontology*, *Foundational Model of Anatomy Ontology*, *Open Biological and Biomedical Ontology*, SNOMED CT e *GALLEN Ontology*. Como estas ontologias possuem bancos próprios pode-se afirmar que 26.31% destes trabalhos acessaram os bancos correspondentes às mesmas.

Então somando as porcentagens da QP3 referentes às ontologias reusadas com as porcentagens da QP6 referentes aos estudos que divulgaram os bancos ontológicos, temos que 50.31% destes estudos declararam direta ou indiretamente os bancos ontológicos consultados.

Dos 49.59% que não declararam, de acordo com as análises realizadas, foi possível inferir que o banco ontológico não é referenciado uma vez que tenha sido informada a ontologia que foi reusada. A partir dela é possível inferir qual o banco foi consultado.

O SNOMED CT, apesar de ser citado como um banco ontológico (Héja, 2008), ele é uma ontologia (MICHAEL, 2010). Por possuir mais de 300000 termos, é consultado de maneira semelhante a um banco de dados e isto faz com que seja confundido com um banco ontológico, (MICHAEL, 2010).

Então podemos citar os seguintes bancos como os mais utilizados: *Gene Ontology Consortium* (<http://www.geneontology.org/>), em Bromberg et. al., (2005),

OBO Foundry (<http://www.obofoundry.org/>) no trabalho de Kong *et. al.* (2011), BioPortal (<http://bioportal.bioontology.org/>), em Thomas, Pappu e Baker (2009).

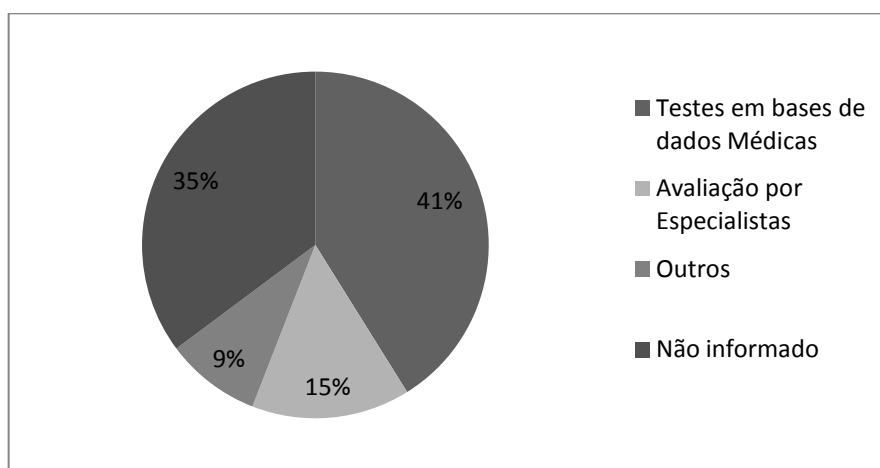
QP7 - Como os trabalhos de Ontologias em Informática Médica estão sendo validados/avaliados?

Segundo a análise das respostas para esta questão foi possível verificar que a maior parte dos estudos sobre o Uso de Ontologias em Informática Médica são avaliada a partir de testes realizados com conjuntos de dados já validados por especialistas do domínio do estudo.

A avaliação por especialistas é a segunda forma de avaliação/validação de trabalhos mais utilizada nesse meio, vide Figura 13. Foi verificado que uma grande parte dos estudos não informa como foram avaliados os trabalhos, isto também está relacionado fato de não existirem métodos de avaliação voltados para esta área do conhecimento.

Foram descobertos estudos que propõe métodos de avaliação para este tipo de estudo como Aranguren *et. al.*, (2008), Schulz, *et. al.* (2011^a), mas estes não representam 1% dos trabalhos selecionados para este Mapeamento Sistemático

Figura 13 - Métodos de Avaliação/Validação de Ontologias



Dentre os trabalhos que propõem novos métodos de avaliação de bio-ontologias, citamos o trabalho de Luther *et. al.* (2011), que consiste no uso de uma técnica conhecida como *Statistical Text Mining*, ela é utilizada para e selecionar os

melhores termos em uma base de dados médica pra posteriormente ser construída a ontologia necessária para o trabalho.

Este método de criação de ontologias permite que a mesma seja matematicamente testada e avaliada após sua construção. Evidentemente características como consistência e corretude dos dados ainda precisa de testes específicos, mas este trabalho é importante por demonstrar a possibilidade do uso de formalismo estatístico para garantir a completude de uma ontologia a ser construída.

Diante destes dados coletados, é possível concluir que este domínio necessita de técnicas de avaliação/validação de ontologias que garantam sua consistência, completude e corretude dos termos.

É válido alertar que mesmo ontologias bastante utilizadas pela comunidade científica possuem incoerências, conceitos errados e ambiguidades conforme apontados por Schulz, *et. al.* (2011^a), Schulz, *et. al.* (2011^b). Ele também afirma que este problema se deve as constantes atualizações e expansões das ontologias supracitadas, então podemos perceber que esta é uma área ainda deficiente quanto ao uso deste tipo de avaliação.

6.3. Mapa das evidências investigadas

O objetivo desta fase de análise e síntese é traçar um panorama geral de como a área está atualmente. Certos comportamentos dos estudos eram esperados, como a grande porcentagem de trabalhos que reusam ontologias, atribuímos isto a influência de grandes pesquisadores como Stefan Schulz e Natalya Noy que recomendam que sempre se deve priorizar o reuso de ontologias, (SCHULZ *et. al.*, 2006; NOY, 2004).

Outra característica interessante é o destaque quanto ao uso de três grandes ontologias nos trabalhos deste domínio, *SNOMED CT*, *Gene Ontology* e a *FMA Ontology* que tem se estabelecido como grandes repositórios de conhecimento. Diversos estudos têm reusado ou integrado novas ontologias a uma delas, como em Sui *et. al.* (2008).

Quanto às ferramentas também não tivemos nada fora do esperado, a W3C (BAILER *et. al.*, 2012) tem recomendado fortemente o uso de suas principais

ferramentas, entre elas o Protégé, e deixado evidente as vantagens das ontologias serem representados em linguagens como OWL, RDF, etc.

Devido ao grande número de ontologias também ficou claro a busca pela integração de bases de dados e ontologias diferentes como no estudo de Sujansky, (2001), Chen, Su e Chen, (2010).

Ficou evidente também a forma com que as Ontologias em Informática Médica tem sido avaliadas, ou seja, grande parte dos trabalhos tem sido avaliados por especialistas da área e mesmo assim erros tem sido encontrados em ontologias que já tenham sido validadas, Schulz, *et. al.* (2011^a).

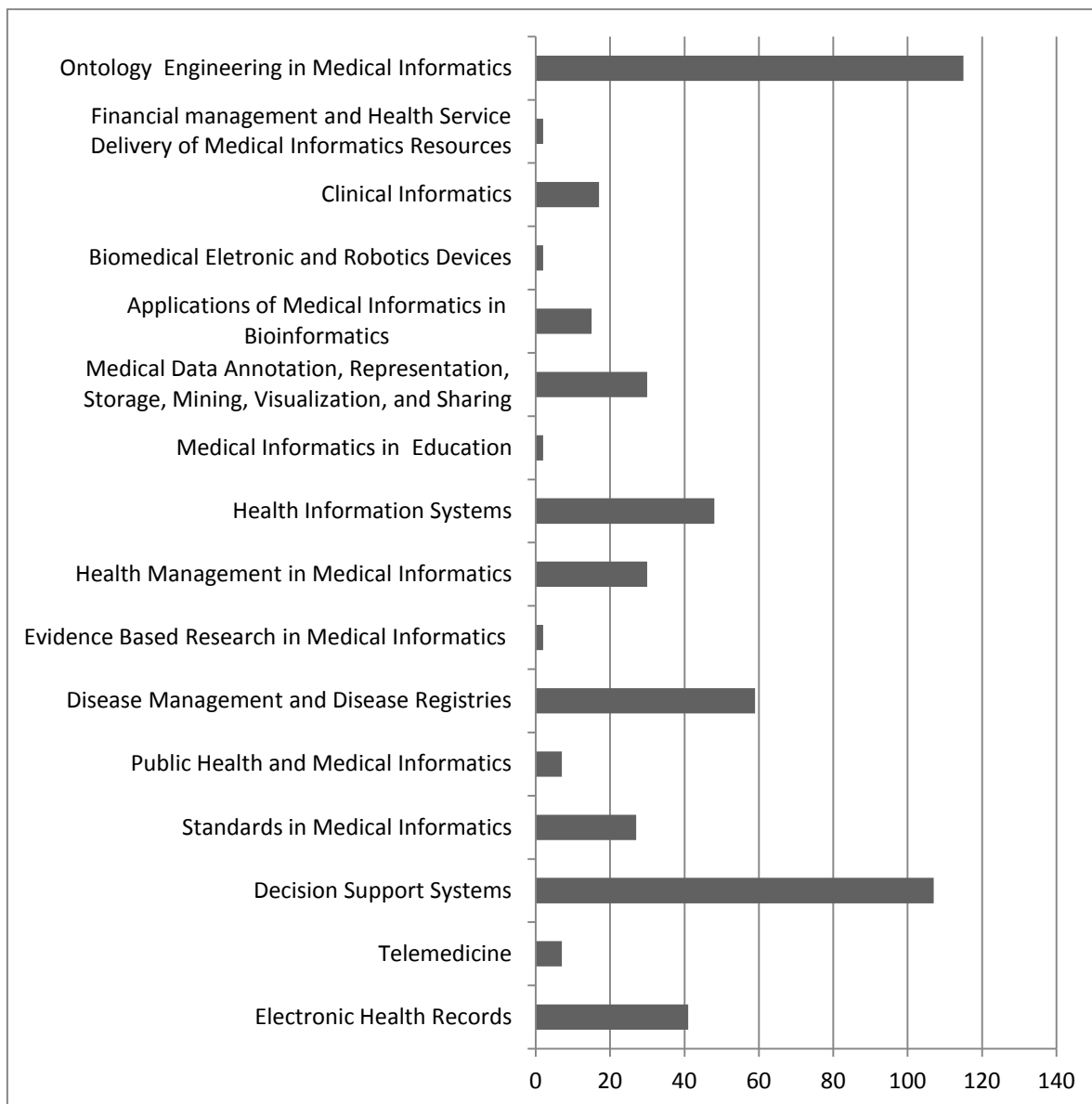
Yu, (2006), relatou em sua pesquisa sobre a utilização de métodos de criação, avaliação, alinhamento e manutenção de bio-ontologias. Sobre os métodos de avaliação mostra que apesar de existirem técnicas, os trabalhos não as têm utilizado. E isto é um problema, pois a Informática Médica tem evoluído, mas parece que as bio-ontologias no que diz respeito a métodos de avaliação não tem acompanhado esta evolução. Obviamente isto pode levar ao desenvolvimento de sistemas falhos e nesta área qualquer possibilidade de falha pode gerar grandes prejuízos.

6.4. Mapeamento das Evidências quanto as subáreas da Informática Médica

Conforme a Tabela 13 foi desenvolvida para este trabalho uma taxonomia composta por 18 termos que definem as diversas subáreas da Informática Médica. Após a síntese dos resultados é possível mapear a evolução das subáreas quanto as questões de pesquisa definidas. É importante salientar que este mapeamento é importante por validar a taxonomia criada para este trabalho.

Na Figura 14, é possível ver o número de publicações por subárea da Informática definidas na Tabela 13 Médica entre 1986 e 2011. Existem áreas pouco exploradas de acordo com os dados coletados, por exemplo: *Telemedicine, Public Health and Medical Informatics, Evidence Based Research in Medical Informatics, Medical Informatics in Education, Biomedical Electronic and Robotic Devices e Financial management and Health Service Delivery of Medical Informatics Resources.*

Figura 14 - Artigos classificados de acordo com a Taxonomia



A subárea *Public Health and Medical Informatics* tem tido dificuldades principalmente no que diz respeito a aquisição e acesso a dados da saúde pública dos países. Quanto a *Evidence Based Research in Medical Informatics*, ou seja, trabalhos como este Mapeamento Sistemático, são difíceis de encontrar, as revisões investigadas não possuem um protocolo que a sistematize.

Consideramos então estas como oportunidades de pesquisa, uma vez que temos o registrado de que há poucos estudos publicados nestas subáreas. Identificamos também que quatro subáreas tiveram um grande número de publicações foram: *Decision Suport Systems*, *Ontology Engineering in Medical Informatics*, *Electronic Health Records* e *Health Information Systems*.

Nestas subáreas identificamos as seguintes tendências:

- *Decision Support Systems*: Foi observado que boa parte dos estudos desenvolvidos são dedicados a classificação e diagnóstico de doenças, (ZHU *et. al.*, 2011), através principalmente da extração de informações de banco de dados clínicos digitalizados ou a partir de registros escritos, (TRAN, QUOC e KAMEYAMA, 2007; RUBIN *et. al.*, 2000) a partir do processamento de textos, relatórios e prognósticos médicos, (ZHU *et. al.*, 2007).

Foram encontrados estudos cujo o foco é providenciar compartilhamento de informações para apoio à decisão clínica, (CENAN, 2008), direcionando os sistemas de diagnóstico para a mais adequada referência do domínio de acordo com a classe de sintomas que recebe como entrada, (MICHIN *et. al.*, 2006), (JORM *et. al.*, 2009).

- *Ontology Engineering in Medical Informatics*: Os trabalhos classificados nesta categoria são os que mais têm considerado o reuso de ontologias (XIANG, *et. al.* 2011). Também encontramos trabalhos dedicados a integração de ontologias existentes como em Chen, Su e Chen, (2010), em que são integradas ontologias para auxiliar na pesquisa e diagnóstico de diabetes. Schulz, *et. al.*, (2006) são exemplos de estudos que consideram o reuso durante a criação de novas ontologias para evitar a fragmentação do conhecimento, bem como novos erros uma vez que ao considerar o reuso serão utilizadas ontologias já validadas pela comunidade científica. Existem também trabalhos que tem se dedicado ao desenvolvimento de métodos para avaliação de ontologias e que garantam a consistência das mesmas durante o processo de criação e sua manutenção, Schulz, *et. al.* (2011^a).
- *Electronic Health Records*: Esta categoria é considerada por Hoyt, (2009), como a coluna central de toda a Informática Médica. Todo o processo de geração e aquisição de conhecimento é facilitada por esta subárea. Trabalhos como o de Argüello, *et. al.*, (2011), propõe a geração de uma ontologia a partir de registros médicos eletrônicos. Também tem sido desenvolvido estudos sobre possíveis estratégias de investigação de registros médicos como em Hovenga, Garde e Heard, (2005), Lee, Tu e Das, (2009) e Ceusters e Smith, (2006). Visto sua importância para a Informática Médica, tem crescido o

número de trabalhos que tem como objetivo acrescentar semântica a este tipo de registro como aponta Hwang, *et. al.* (2010).

- *Health Information Systems*: a subárea de Sistemas de Informação é algo já estabelecido na Ciências da Computação, (HOYT, 2009), também é dito que a mesma tem crescido bastante em detrimento a possibilidade de compartilhar conhecimento e o desenvolvimento de sistemas distribuídos. Lenz, *et. al.* (2002), desenvolve um sistema de informação médica cujo objetivo é integrar os sistemas médicos de diferentes setores de um mesmo hospital de modo que as informações dispostas no sistema seja integrada, avaliada e auxilie no tratamento e diagnóstico dos pacientes. Também tem sido desenvolvidos sistemas de informação voltados para o gerenciamento de recursos médicos Anagnostakis, *et. al.*, (2005). Orgun e Vu, (2006), integram sistemas de informações médicos diferentes através do uso de ontologias e agentes inteligentes.

A partir dos dados coletados as principais tendências de pesquisa são direcionadas para a integração de informações médicas de natureza heterogênea, ou oriundas de sistemas diferentes. Também é possível perceber a preocupação com a qualidade dos sistemas desenvolvidos e a criação de métodos de avaliação das ontologias utilizadas como base dos estudos. A principal tendência observada também é aliar registros médicos eletrônicos a semântica provida pelas ontologias.

6.5. Avaliação de qualidade dos trabalhos

De acordo com as métricas estabelecidas na Tabela 16, baseadas nos trabalhos de Kitchenham, (2007), Petticrew, (2005), Fink, (2005) e Greenhalgh, (2000), os trabalhos selecionados foram avaliados dentro dos seguintes conceitos e intervalos de notas, vide a Tabela 17.

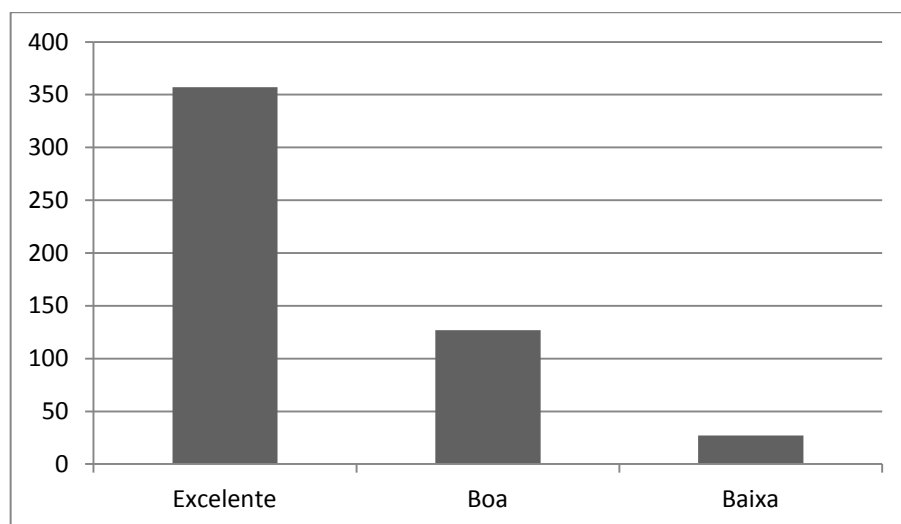
A Figura 15 mostra a qualidade dos trabalhos selecionados para este Mapeamento Sistemático, portanto é possível afirmar que a seleção dos trabalhos foi rigorosa e que os critérios para inclusão foram suficientes para garantir este resultado uma vez que 70% dos artigos foram classificados com excelente e 24,7% com boa qualidade.

Tabela 17 - Intervalos das qualidades dos trabalhos avaliados

Qualidade	Valor Numérico
Baixa	0 – 0.49
Boa	0.5 – 0.79
Excelente	0.8 - 1

Enquanto na pequena quantidade de trabalhos foram considerados de qualidade baixa, 5.3%. Este resultado era esperado por terem sido considerados apenas estudos finalizados e revisados, isto garantiu que os artigos selecionados produzissem este resultado. Este tipo de constatação demonstra a qualidade do Mapeamento Sistemático produzido neste trabalho, de um modo geral

Figura 15 - Qualidade das Trabalhos



Em média os artigos analisados possuem é considerada Boa, ou seja, até 0.79. Isto demonstra que boa parte dos estudos sobre Ontologias e Informática Médica estão de acordo com os itens analisados na Tabela 16, ou seja, os pesquisadores especificam os objetivos, estabelecem a metodologia claramente e de forma coerente, bem como os passos durante o desenvolvimento dos estudos e grande parte dos estudos alcançam os resultados pretendidos.

6.6. Avaliação do Mapeamento Sistemático

De acordo com Kitchenham, (2007) e Fink, (2005), um formulário deve ser elaborado para que o Mapeamento Sistemático seja avaliado pela equipe que o executou. Para tanto foi elaborado o seguinte formulário disposto na Tabela 18. As notas dadas estão de acordo com as sugestões de Fink, (2005) vistos na Tabela 16.

Tabela 18 - Formulário para Avaliação do Mapeamento Sistemático

Questões de Avaliação		Nota
1	Qual a credibilidade destas conclusões?	0.67
2	Quanto esta avaliação alcança seus objetivos iniciais e propósitos?	1
3	Como pode ser avaliada a coleta de dados para este trabalho?	1
4	Como este trabalho pode ser avaliado quanto a profundidade e complexidade dos resultados que foram divulgados.	0.67
5	Quão explícita é a relação entre a interpretação dos dados e as conclusões as quais os pesquisadores chegaram?	1
6	Quão claro e coerente é o mapeamento?	1
7	Quão claras são as premissas, perspectivas e valores que moldaram a forma e a saída da avaliação?	1
8	Quão adequadamente o processo de investigação tem sido documentado?	0.67
Total		0.88

De acordo com as notas apresentadas na Tabela 18, a qualidade geral do Mapeamento Sistemático foi considerada Boa. A profundidade das análises está

relacionada diretamente com as Questões de Pesquisa, por isso não foi possível uma análise mais aprofundada sobre o Uso de Ontologias e Informática Médica, porém as análises executadas estão de acordo com os objetivos definidos na Sessão 1.3. Então podemos declarar que o Mapeamento Sistemático foi executado com sucesso.

Segundo Kitchenham, (2007), a última etapa do Mapeamento Sistemático é a etapa de reporte do mesmo para que possa ser utilizado pela comunidade científica. Nós iniciamos esta etapa com a divulgação dos resultados em Mota, *et. al.*, (2012) e Mota, *et. al.*, (2013), estes artigos estão nos Anexos I e II deste trabalho. Também foram definidos os seguintes periódicos para a publicação dos resultados:

Tabela 19 - Periódicos selecionados para a divulgação do Mapeamento Sistemático

ID	Área	Nome	Qualis
1469-8005	CIÊNCIA DA COMPUTAÇÃO	Knowledge Engineering Review (Online)	B1
0950-7051	CIÊNCIA DA COMPUTAÇÃO	Knowledge-Based Systems	B1
1353-4505	CIÊNCIA DA COMPUTAÇÃO	International Journal for Quality in Health Care	B2
0926-9630	CIÊNCIA DA COMPUTAÇÃO	Studies in Health Technology and Informatics	B2
1570-5838	CIÊNCIA DA COMPUTAÇÃO	Applied Ontology	B2
1741-1009	CIÊNCIA DA COMPUTAÇÃO	International Journal of Knowledge and Learning	B2
0218-1940	CIÊNCIA DA COMPUTAÇÃO	International Journal of Software Engineering and Knowledge Engineering	B2
1741-1009	CIÊNCIA DA COMPUTAÇÃO	International Journal of Knowledge and Learning (Print)	B2
0218-1940	CIÊNCIA DA COMPUTAÇÃO	International Journal of Software Engineering and Knowledge Engineering	B2
1353-4505	CIÊNCIA DA COMPUTAÇÃO	International Journal for Quality in Health Care	B3
1744-2621	CIÊNCIA DA COMPUTAÇÃO	International Journal of Metadata, Semantics and Ontologies (Print)	B3
1555-3396	CIÊNCIA DA COMPUTAÇÃO	International Journal of Healthcare Information Systems and Informatics	B4

Os periódicos presentes na Tabela 19 foram selecionados de acordo com a afinidade com a área de Ontologias e Informática Médica e pelo Qualis disponibilizado pela CAPES.

Capítulo 7 - Conclusões

O principal objetivo desta pesquisa foi alcançado, ou seja, foi apresentada uma visão geral sobre Uso de Ontologias e Informática Médica, fornecendo informações úteis para o desenvolvimento de estudos neste contexto.

Dentre as características atuais de como se encontra a área de estudo deste trabalho pode-se afirmar que:

- 45% dos estudos desenvolvidos sobre o uso de Ontologias em Informática Médica são direcionados para a subárea de *Ontology Engineering in Medical Informatics*. Enquanto 34% são direcionados para a subárea de *Decision Support Systems*, essas duas categorias englobam cerca de 79% dos estudos investigados;
- 60% dos estudos analisados recorrem ao uso de Ontologias em Informática Médica devido as suas principais características e benefícios, a saber: são especificações formais, explícitas de um conhecimento consensual e passível de processamento automático, (BREITMAN, 2010). 14% dos trabalhos afirmam que as ontologias facilitam a integração e interoperabilidade de sistemas clínicos diferentes, como em Das *et. al.*, (2006);
- Quanto às ontologias mais utilizadas, foi uma das questões de pesquisa em que mais foi encontrada fragmentação de respostas, de modo que *Gene Ontology* (8%) é a ontologia mais reutilizada documentada por este Mapeamento Sistemático. Isto se deve ao fato de que foram encontradas por volta de 296 ontologias diferentes, e também está relacionado com o grande número de trabalhos (cerca de 12.45% do total) não especificaram uma ontologia que será utilizada em seu sistema como em Crowley e Medvedeva, (2003), principalmente por desenvolverem aplicações adaptáveis a diferentes contextos.
- 64.26% dos estudos em Informática Médica consideram o reuso de ontologias. É possível relacionar essa estatística a anterior somando todas as porcentagens da questão de pesquisa 3 que declaram o reuso de ontologias e possuem uma relevância estatística a partir de 1%,

deste modo chegamos a um percentual de 62%, ou seja, uma diferença de apenas 2% que provavelmente está relacionado às estatísticas inferiores a 1%; Ou seja, as respostas das questões de pesquisa podem ser relacionadas e estão coerentes em seus resultados finais;

- Quanto às ferramentas utilizadas para edição e construção de ontologias é possível afirmar que o Protégé foi utilizado por 40% dos estudos analisados. 28% declaram o uso de ferramentas diferentes como NeOn Toolkit (KHAN JR., 1998); Java Ontology Editor (JOE), (MAHALINGAM E HUHNS, 1997); Swoop, (SAMWALD ET. AL., 2010); Semantic Editor (ZHEN, et. al. 2009); Differential Ontology Editor (DOE), (BARNEYX, CHARLET E JAULENT, 2006);
- A questão de pesquisa 6 também revelou uma grande fragmentação em suas respostas, estas referentes aos bancos de ontologias mais utilizados na Informática Médica. 24% dos estudos declararam que utilizam um dos seguintes bancos de ontologias: *Gene Ontology*, *Foundational Model of Anatomy Ontology*, *Open Biological and Biomedical Ontology*, SNOMED CT e *GALLEN Ontology*. Esta estatística está diretamente relacionada a da questão de pesquisa três, uma vez que se somadas as porcentagens das ontologias mais utilizadas temos 26.31%. Ou seja, pode-se afirmar com certa segurança que há coerência entre os dados relacionados destas duas questões.
- Quanto a questão de pesquisa sete, talvez seja a que menos possui relação com as anteriores, pode-se afirmar que 41% dos estudos são avaliados e validados através de testes em bases de dados médicas, da área de saúde, que 15% são avaliados ou validados por especialistas, 8% são avaliados por outros métodos como em Luther et. al. (2011) que utiliza a técnica de *Statistical Text Mining* para selecionar termos para a criação de ontologias permitindo sua validação e avaliação posterior através de métodos estatísticos. Também é possível relacionar as duas estatísticas desta questão da seguinte forma: uma vez que as bases de dados médicas foram

validadas e avaliadas por profissionais da área de saúde, pode-se assumir que 56% dos estudos são validados por especialistas da medicina.

Também foram identificadas oportunidades atuais de pesquisa, áreas com maior quantidade de estudos, e as áreas que mais carecem de pesquisas relacionadas. São elas: *Telemedicine, Public Health and Medical Informatics, Evidence Based Research in Medical Informatics, Medical Informatics in Education, Biomedical Electronic and Robotic Devices e Financial management and Health Service Delivery of Medical Informatics Resources*.

Quanto ao caráter inovador desta pesquisa, podemos afirmar que não foram encontrados estudos semelhantes, nos principais engenhos e periódicos pesquisados definidos nas Tabelas 1, 3 e 6, desta forma, existe a possibilidade deste ser um trabalho pioneiro na temática especificada.

Se existirem outros estudos anteriores, ou atuais, nesta linha de pesquisa ou semelhantes, é importante salientar que estudos deste tipo podem ser comparados mas com as devidas adequações tendo em vista que o que define a profundidade e semelhança entre Mapeamentos Sistemáticos são as Questões de Pesquisa Kitchenham e Dyba (2004), Dyba, (2005) e Kitchenham, (2007) e Travassos (2007).

Foram encontradas revisões dentro desta área nos artigos selecionados, por volta de 65 artigos eram deste tipo, porém as revisões encontradas são bastante específicas como: Kashifi (2011), cujo estudo consiste numa revisão sobre a interseção entre sistemas de suporte a decisão e registros eletrônicos de saúde; Liao (2005), que faz uma revisão sobre sistemas especialistas suas metodologias e aplicações entre os anos de 1995 e 2004; Burgun (2002), em que é feito um estudo de caso relacionado a ontologias voltadas para a genética.

Ou seja, foram encontrados artigos de revisões, porém ou não estavam diretamente relacionados especificamente com a área de pesquisa deste trabalho de um modo geral, ou eram bastante específicos dentro da Informática Médica. Apenas Sun *et. al.* (2012) é uma revisão sistemática de literatura, mas não está relacionado ao escopo desta pesquisa.

No mais, por não serem revisões sistematizadas este tipo de comparação deve ser realizada com cuidado, Kitchenham (2007). Os artigos de revisão

encontrados nesta pesquisa podem ser visto no Apêndice B.

7.1. Principais Contribuições

As contribuições esperadas eram as seguintes:

- Apresentar uma visualização geral de como está sendo desenvolvida a área pesquisada, através da resposta das questões de pesquisa definidas na sessão 1.2;
- Verificar os principais autores, instituições e países que estão desenvolvendo estudos nesta linha;
- Identificar quais as subáreas da Informática Médica com o maior desenvolvimento, e quais as que se beneficiariam mais com o uso das tecnologias semânticas, como as Ontologias;
- Principais deficiências da área pesquisada e as dificuldades encontradas para o desenvolvimento de trabalhos;

Concluimos que as contribuições esperadas foram alcançadas, visto que:

- Foi possível mapear, de acordo com as questões de pesquisa, o estado atual sobre o Uso de Ontologias em Informática bem como as tendências de pesquisa dentro da área;
- Foram mapeados os autores, instituições e países que mais desenvolveram estudos em um intervalo de 5 anos. Ou seja, entre 2006 e 2011;
- Também foram identificadas as principais subáreas da Informática Médica, como também elencamos as que possuem maior quantidade de trabalhos publicados e as tendências de pesquisa nas mesmas.
- Foram reveladas as subáreas que possuem uma menor quantidade de publicações e as oportunidades de pesquisa nelas;

Resta apenas divulgar este trabalho completo para que o mesmo contribua com o desenvolvimento desta área, portanto falta concretizar a seguinte contribuição:

- Servir de base para o desenvolvimento de estudos relacionados a Ontologias e Informática Médica, tanto para pesquisadores em geral quanto especializados, de acordo com Cooper, (1988).

No mais, a partir das terminologias estudadas e na taxonomia final, também foi desenvolvido o protótipo de uma Taxonomia Ontológica, como pode ser visto na Figura 16. Uma Taxonomia Ontológica pode ser definida como uma taxonomia acrescida da semântica proporcionada pela ontologia, (BREITMAN, 2010).

Figura 16 - Protótipo d Taxonomia Ontológica



Uma taxonomia serve para classificar informações em uma hierarquia (árvore) utilizando o relacionamento pai-filho (generalização ou “tipo-de”). Uma Ontologia permite atribuir características a cada termo existente na taxonomia, fazendo uma distinção clara entre suas definições, Breitman, (2010). Dessa maneira, nenhum estudo poderá ser agrupado, classificado ou relacionado erroneamente com algum termo incompatível com suas características. Portanto, através do desenvolvimento da Taxonomia Ontológica, é possível um compartilhamento e organização mais eficientes de informações presentes em bases de dados relacionadas a Informática Médica. Por ser concebida em OWL, de acordo com os padrões do W3C, sua integração com projetos como os propostos por Borges *et. al.*, (2011) e Kamura *et. al.*, (2011) bem como com ferramentas de buscas baseadas em tecnologias semânticas, além de ser de grande utilidade para o desenvolvimento da Web Semântica, Hausenblas, *et. al.*, (2010).

7.2. Trabalhos Futuros

Dentre os trabalhos que poderão ser desenvolvidos a partir deste Mapeamento Sistemático, podemos citar os seguintes:

- Desenvolvimento de estudos relacionados às subáreas da Informática Médica deficientes identificadas no mapeamento sistemático;
- Desenvolvimento de pesquisas seguindo as principais tendências reveladas pelo presente trabalho;
- Estudos desenvolvidos com o objetivo de oferecer soluções para as lacunas encontradas nas pesquisas selecionadas e avaliadas;
- Servir de base para o desenvolvimento de projetos de graduação e pós-graduação na área investigada;
- Refinar a Taxonomia Ontológica, aqui desenvolvida, para que a mesma permita a estruturação de bases de dados de estudos relacionados à Informática Médica e o compartilhamento de conhecimento mais eficiente através de aplicações que utilizem tecnologias semânticas Hausenblas, *et. al.*, (2010) e Heath (2011);
- Utilizar técnicas estatísticas entre os dados permitindo que sejam inferidos novos conhecimentos a partir desta base de dados estabelecida.

7.3. Limitações deste Trabalho

Todo Mapeamento Sistemático tem as seguintes limitações de acordo com Kitchenham e Dyba (2004), Dyba, (2005) e Kitchenham, (2007) e Travassos (2007): cobertura limitada, possível viés introdutório no processo de seleção, imprecisões na extração de dados e na avaliação de qualidade. Portanto, entendemos que este Mapeamento Sistemático deve possuir estas mesmas limitações.

Porém o protocolo foi desenvolvido com base em diretrizes bem definidas para que fossem reduzidas as limitações já citadas.

Nenhum problema foi identificado quantos as terminologias utilizadas, mas salientamos que talvez este trabalho tenha um viés devido aos engenhos selecionados. Como boa parte destes estudos vieram de engenhos como PMC e PubMed, dedicados principalmente a genética e biologia, talvez as categorias como: *Telemedicine, Public Health and Medical Informatics, Evidence Based Research in*

Medical Informatics, Medical Informatics in Education, Biomedical Electronic and Robotic Devices e Financial management and Health Service Delivery of Medical Informatics Resources; tenham tido poucos trabalhos registrados.

Com relação a profundidade do Mapeamento, esta depende das Questões de Pesquisa definidas e por isso pode ser considerada uma visão bastante geral, não aprofundada. Porém de acordo com a literatura, (KITCHENHAM e DYBA, 2004; DYBA, 2005; KITCHENHAM, 2007; TRAVASSOS, 2007), isto não torna este trabalho inválido nem inferior a outros mapeamentos nesta área caso existam. Também é sabido que trabalhos deste tipo podem ser reavaliados e estendidos, portanto ainda podem-se utilizar os dados obtidos para uma análise mais aprofundada no domínio especificado.

Capítulo 8 - Referências

- ALMEIDA, A., T. **Um Mapeamento Sistemático de Mecanismos para Guiar Estudos Empíricos em Engenharia de Software** Dissertação (Mestrado em Ciências da Computação). Universidade Federal de Pernambuco. Recife, Pernambuco. Março, 2011.
- ABIDI, S. S. R.; HUSSAIN, S.; **Medical Knowledge Morphing via a Semantic Web Framework**. Twentieth IEEE International Symposium on Computer-Based Medical Systems, 554-562, Junho, 2007.
- ALTMAN, R. B.; BERMAN, H. M.; BRENNER, S. E.; BROWN, J. W.; ENGELKE, D. R.; HARVEY S. C.; HOLBROOK, S. R.; JOSSINET, F.; LEONTIS, N. B.; LEWIS, S. E.; MAJOR, F. O.; MATHEWS, D. H.; RICHARDSON, J. S.; WESTHOF, E.; WILLIAMSON, J. **The RNA Ontology Consortium: An open invitation to the RNA community**. Cold Spring Harbor Laboratory Press, USA, 2006.
- ÁLVAREZ-RODRÍGUEZ, J. M.; CASADO-LUMBRERAS, C.; COLOMO-PALACIOS, R.; RODRÍGUEZ-GONZÁLEZ, A. **PsyDis: Towards a diagnosis support system for psychological disorders**. Expert Systems with Applications, 39: 11391–11403, 2012.
- ANAGNOSTAKIS, A. G.; Tzima, M.; Sakellaris, G. C.; Fotiadis, D. I.; Likas, A. C.; **Semantics-based information modeling for the health-care administration sector: the citation platform**. IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society, v. 9, n. 2, p. 239-47, jun. 2005.
- ARANGUREN, M. E.; ANTEZANA, E.; KUIPER, M.; STEVENS, R.; **Ontology Design Patterns for bio-ontologies: a case study on the Cell Cycle Ontology**. BMC Bioinformatics, 9(Suppl 5), Abril, 2008.
- ARGUELLO, M.; DES, J.; FERNANDEZ-PRIETO, M.J.; PEREZ, R.; LEKKAS, S.; **An Ontology-Based Approach to Natural Language Generation from Coded Data in Electronic Health Records**. UKSim 5th European Symposium on Computer Modeling and Simulation, 366-371, 2011.
- ARKSEY, H; O'MALLEY, L. **Scoping Studies: towards a methodological framework**. International Journal of Social Research Methodology, vol. 8, número 1, p. 19-32. 2005.
- BAILER, W.; BÜRGER, T.; CHAMPIN, P., A.; EVAIN, J., P.; LEE, WonSuk; MALAISÉ, V.; MICHEL, T.; SASAKI, F.; SÖDERBERG, J.; STEGMAIER, F.; STRASSNER, J.; **Ontology for Media Resources 1.0 W3C Recommendation**. Novembro, 2009. Disponível em <http://www.w3.org/TR/mediaont-10/>. Acesso em: 15 mai. de 2012, 10:30.

BANEYXA, A.; CHARLETA, J.; JAULENT, M-C.; **Building an ontology of pulmonary diseases with natural language processing tools using textual corpora.** International journal of medical informatics, 76: 208–215, 2007.

BORGES, M. R. S., CAMPOS, M. L. M., CORDEIRO, K. F. AND MARINO, T.; **Use of Linked Data in the Design of Information Infrastructure for Collaborative Emergency Management System.** In Proceedings of 15th International Conference on Computer Supported Cooperative Work in Design, 764-771, 2011.

BREITMAN, K, K. **Web Semântica: A Internet do Futuro.** LTC – Livros Técnicos e Científicos Ltda. Rio de Janeiro, 2010.

BRERETON, P.; KITCHENHAM, B., A.; BUDGEN, D., T.; MOHAMED, M.; MOHAMED, K.. **Lessons from applying the systematic literature review process within the software engineering domain.** JSS 80, 571-583, 2007.

BRERETON, P.; BUDGEN, D.; **Performing Systematic Literature Reviews in Software Engineering.** International Conference on Software Engineering; Shanghai, China, Maio, 2006.

BROMBERG, S.; CHEN, J.; JACOB, H.; KWITEK, A.; NENASHEVA, N.; PASKO, D.; PETRI, V.; SHIMOYAMA, M.; TWIGGER, S.; WU, W.; **Using multiple ontologies to integrate complex biological data.** *Comparative and functional genomics*, 6:373–378, 2005.

BRUC, N.; MAGARIU, G.; VERLAN, T.; **Elaborating of Ultrasound Images Ontology in Ultrasound Diagnostics.** 3rd International Conference on E-Health and Bioengineering, 24-27, Iasi, România, 2011.

Burgun, A.; Bodenreider, O.; Duff, F. L.; Moussouni, F.; Loreal, O.; **Representation of roles in biomedical ontologies: a case study in functional genomics.** Annual AMIA Symposium, 86-90, E.U.A., 2002.

CARDOZO, E., S., F. **Mapeamento Sistemático sobre o uso do Auto-Gerenciamento em Equipes de Desenvolvimento de Software.** Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco. 2012.

CEUSTERS, W.; SMITH, B.; **Strategies for referent tracking in electronic health records.** *Journal of Biomedical Informatics*. 39 (Issue 3), 362–378, 2006.

CHANG, E.; HADZIC, M.; WONGTHONGTHAM, P.; MEERSMAN, R.; **Disease Ontology based Grid Middleware for Human Disease Research Study.** 30th Annual Conference of the IEEE Industrial Electronics Society, 480-486, November, Busan, Korea, 2004.

CHANG, E.; DILLON, T. S.; SIDHU, A. S.; **An Ontology for Protein Data Models.** International Conference of the IEEE Engineering in Medicine and Biology Society, 6:6120-3, Shanghai, China, 2005.

CHEUNG, K-H.; CLARK, T.; CRASTO, C.; GAO, Y.; KINOSHITA J.; LAM, H. YK.; LIU, N.; MARENCO, L.; MILLER, P.; MORSE, T.; SHEPHERD, G.; STEPHENS, S.; WONG, G. T.; WU, E. **AlzPharm: integration of neurodegeneration data using RDF.** *BMC Bioinformatics*; Maio, 2007.

CHEN, J-X. ; SU, S-L.; CHEN, Z-L.; **Sharing and Integration of Profession Ontology by Web Service - An Example of Diabetes Care.** International Conference on Electronics and Information Engineering, V2, 136-140, 2010.

CHUTE, C. G.; SOLBRIG, H. R.; TAO, C.; **CNTRO 2.0: A Harmonized SemanticWeb Ontology for Temporal Relation Inferencing in Clinical Narratives.** AMIA - Annual Symposium on biomedical and health informatics, Washington – DC, USA, 2011.

CONDORI-FERNANDEZ, N.; DANEVA, M.; DIESTE, O.; PASTOR, O.; SIKKEL, K.; WIERINGA, R.; **A Systematic Mapping Study on Empirical Evaluation of Software Requirements Specifications Techniques.** Third International Symposium on Empirical Software Engineering and Measurement. Flórida, E.U.A., Outubro, 2009.

COOPER, H. M. **Organizing knowledge synthesis: A taxonomy of literature reviews.** *Knowledge, Technology & Policy*, vol. 1, nº1, p. 104-206, 1988.

CROWLEY, R. S.; MEDVEDEVA, O.; **A General Architecture for Intelligent Tutoring of Diagnostic Classification Problem Solving.** AMIA - Annual Symposium on biomedical and health informatics; Washington – DC, USA, Novembro, 2003.

DACONTA, M.; SMITH , K.; OBRST, L. **The Semantic Web: The Future of XML, Web Services, and Knowledge Management.** John Wiley, Inc., Maio, 2003.

DANG, J; HEDAYATI, A; HAMPEL, K; TOKLU, C. **An ontological knowledge framework for adaptive medical workflow.** *Journal of Biomedical Informatics*, volume 41, Junho, 2008. P. 829–836.

DAS, A. K.; MARTINS, S. B.; O'CONNOR, M.; PARRISH, D. B.; SHANKAR, R. D. **A Knowledge-Based System for Managing Complex Clinical Trials.** 19th IEEE Symposium on Computer-Based Medical Systems, 270-278, Maribor, Slovenia, 2006.

DIXON, B. E., MCGOWAN, J. J. AND ZAFAR, A.; **Development of a Taxonomy for Health Information Technology.** In Proceedings of World Congress on Medical and Health Informatics, Australia, 2007.

DYBA, T.; JORGENSEN, M.; KINTCHENHAM, B. **Evidence-based software engineering for practitioners.** *IEEE Software*, p. 58 – 65. Janeiro, 2005.

FELDT, R., PETERSEN, K., MATTSSON, M. AND MUJTABA, S.; **Systematic Mapping Studies in Software Engineering.** In Proceedings of 12th International Conference on Evaluation and Assessment in Software Engineering, Bari, Italy, 71-80, 2008.

FERNANDEZ-BREIS, J. T.; MENARGUEZ-TORTOSA, M.; MARTINEZ-COSTA, C.; FERNANDEZ-BREIS, E.; HERRERO-SEMPERE, J.; MONER, D.; SANCHEZ, J.; VALENCIA-GARCIA, R.; ROBLES, M.; **A SemanticWeb-based SystemforManaging Clinical Archetypes**. Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1482-5, Vancouver, Canadá, 2008.

FINK, A. **Conducting Research Literature Reviews**. From the Internet to Paper, Sage Publication Inc., 2005.

FUJITA, H.; HAKURA, J.; KUREMATSU, M.; **Multiviews Ontologies based Reasoning for Medical Diagnosis in VDS**. IEEE 9th International Symposium on Intelligent Systems and Informatics; Subotica, Serbia, Setembro, 2011

GÓMEZ-PÉREZ, A.; FERNÁNDEZ-PÉREZ, M.; CORCHO, O.; **Ontological Engineering**. Springer Verlag. 2004.

GREENHALGH, T. **How to read a paper: The Basics of Evidence-Based Medicine**. BMJ Books, 2000.

GRUBER, T., R. **Toward principles for the design of ontologies used for knowledge sharing**. International Journal of Human-Computer Studies, Vol. 43, Issues 4-5, November 1995, p. 907-928.

GUARINO, N. **Formal Ontology and Information Systems**: Proceedings of the 1st International Conference. Junho 6-8, Trento, Itália. IOS Press Amsterdam, Holanda, 1998.

GUARINO, N.; WELTY, C. **Evaluating Ontological Decisions with Ontoclean**. Communications of the ACM, Vol. 45, No.2 – 2002. Pp.61-65.

HAUSENBLAS, M. AND KARNSTEDT, M.; **Understanding Linked Open Data as a Web-Scale Database** .In Proceedings of Second International Conference on Advances in Databases, Knowledge, and Data Applications, 56-61, 2010.

HEATH, T.; **Linked Data — Welcome to the Data Network**. IEEE Internet Computing, vol. 11, 70-73, 2011.

HÉJA, G.; SURJÁN, G.; VARGA, P.; **Ontological analysis of SNOMED CT**. BMC Medical Informatics and Decision Making, 8(Suppl 1):S8, 2008.

HOVENGA, E.; GARDE, S.; HEARD, S.; **Nursing constraint models for electronic health records: A vision for domain knowledge governance**. International Journal of Medical Informatics, 74 (Issue 11-12), 886—898, 2005.

HOYT, R. E.; SUTTON, M.; YOSHIHASHI, A. **Medical Informatics - Practical Guide for the Healthcare Professional**. 3ª Edição, University of West Florida, School of Allied Health and Life Sciences, 2009.

HU, B.; DASMAHAPATRA, S.; DUPPLAW, D.; LEWIS, P.; SHADBOLT, N.; **Reflections on a medical ontology**. International Journal of Human-Computer Studies. 65: 569-582, 2007.

HU, B.; DASMAHAPATRA, S.; LEWIS, P.; SHADBOLT, N.; **Ontology-based Medical Image Annotation with Description Logics**. International Conference on Tools with Artificial Intelligence, 77-82, E.U.A., Novembro ,2003.

HUANG, J.; DOU, D.; HE, LEI, H. P.; DANG, J.; **Ontology-Based Knowledge Discovery and Sharing in Bioinformatics and Medical Informatics: A Brief Survey**. Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2203-2208, China, 2010.

HWANG, K. H.; CHUNG, K-I.; CHUNG, M-A.; CHOI, D.; **Review of Semantically Interoperable Electronic Health Records for Ubiquitous Healthcare**. Healthcare informatics research, 16 (Issue 1), 1-5, 2010.

ISERN, D.; SÁNCHEZ, D.; MORENO, A.; **Ontology-driven execution of clinical guidelines**. Journal of Computer methods and programs in biomedicine, 107: 122-39, 2012.

JACINTO, S., S. **Um Mapeamento Sistemático de Pesquisa sobre a Influência da Personalidade na Engenharia de Software**. Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco. Setembro, 2010.

JORM, L.; GRUSZIN, S.; CHURCHES, T.; **A multidimensional classification of public health activity in Australia**. In: Aust N Z Health Policy, 6:9, 2009.

KAMURA, T., KATO, F., OHMUKAI, I., TAKAHASHI, T., TAKEDA, H. AND UEDA, H.; **Study support and integration of cultural information resources with Linked Data**. In Proceedings of Second International Conference on Culture and Computing, 177 – 178, 2011.

KASHFI, H.; **The intersection of clinical decision support and electronic health record: A literature review**. Computer Science and Information Systems, 347 - 353, Polônia, Setembro, 2011.

KHAN, K. S.; TER RIET, G.; GLANVILLE, J.; SOWDEN, A. J; KLEIJNEN, J. **Undertaking Systematic Review of Research on Effectiveness. CRD's Guidance for those Carrying Out or Commissioning Reviews**. CRD Report Number 4 (2nd Edition), NHS Centre for Reviews and Dissemination, University of York, March 2001.

KAHN JR., C. E.; **An Internet-based ontology editor for medical appropriateness criteria**. Journal of Computer Methods and Programs in Biomedicine, 56:31–36, Dezembro, 1998.

KITCHENHAM, B.; DYBA, T.; JORGENSEN, M. **Evidence-based software engineering**. In ICSE'04: Procendings of the 26th International Conference on Software Engineering, p. 273-281, Washington, DC, USA, 2004.

KITCHENHAM, B. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. Vol 2.3 EBSB Technical Report, EBSE-2007-01, 2007.

KONG, Y. M.; DAHLKE, C.; XIANG, Q.; QIAN, Y.; KARP, D.; SCHEUERMANN, R. H.; **Toward an ontology-based framework for clinical research databases.** Journal of biomedical informatics, 44: 48–58, 2011.

KUIPERS, B., AND KASSIRER, J.; **Causal Reasoning in Medicine: Analysis of a Protocol.** Cognitive Science, v.8, p.363, 1984.

KULIKOWSKI, C. A.; **Knowledge-Based Systems In Biomedicine : A 10 Year Retrospective.** Department of Computer Science; New Rutgers University; Brunswick, New Jersey, p. 423-424, 1986.

LASSILA, O., McGuinness, D. **The Role of Frame-Based Representation on the Semantic Web.** KSL Technical Report No. KSL-01-02. Disponível em: <http://www.ksl.stanford.edu/people/dlm/etai/lassila-mcguinnessfbrsw.html>. Acesso: 22 abr. de 2012, 20:15.

LEE, W-N.; TU, S. W.; DAS, A. K.; **Extracting Cancer Quality Indicators from Electronic Medical Records: Evaluation of an Ontology-Based Virtual Medical Record Approach.** Annual AMIA Symposium, 349-53, 2009.

LENZ, R.; ELSTNER, T.; SIEGELE, K.; KHUN, K. A.; **A Practical Approach to Process Support in Health Information Systems.** Journal of the American Medical Informatics Association v. 9, n. 6, p. 571-585, 2002.

LIAO, S.-H.; **Expert system methodologies and applications — a decade review from 1995 to 2004.** Expert Systems with Applications. Vol. 28, Issue 1, 93–103, Janeiro 2005.

LUTHER, S.; BERNDT, D.; FINCH, D.; RICHARDSON, M.; HICKLING, E.; HICKAM, D.; **Using statistical text mining to supplement the development of an ontology.** Journal of Biomedical Informatics, 44, S86–S93, 2011.

MARAGOUDAKIS, M.; MAGLOGIANNIS, I.; LYMBEROPOULOS D. **A Medical, Description Logic based, Ontology for Skin Lesion Images.** 8th IEEE International Conference on BioInformatics and BioEngineering, Atenas, Grécia, 2008.

MAHALINGAM, K.; HUHNS, M. N.; **An Ontology Tool for Query Formulation in an Agent-Based Context.** 2nd IFCIS - Conference on Cooperative Information Systems, 170-178, E.U.A., 1997.

MARAGOUDAKIS, M.; MAGLOGIANNIS, I.; LYMBEROPOULOS, D. **A Medical, Description Logic based, Ontology for Skin Lesion Images.** Conference on Bio-Informatics and Bio-Engineering 1 – 6. Grécia, Outubro, 2008.

MARCONI, M. A.; LAKATOS, E. M. **Metodologia Científica: Ciência e Conhecimento Científico.** Atlas, 2007.

MOTA, A. R. M.; LINO, N. C. Q.; TAVARES, E. A.; CAMPOS, S. P. R.; **Uma Taxonomia Ontológica para a Classificação de Estudos em Informática Médica.** Simpósio Brasileiro de Sistemas de Informação, João Pessoa, Paraíba, 2013

MOTA, A. R. M.; LINO, N. C. Q.; TAVARES, E. A.; CAMPOS, S. P. R.; **Um Estudo Sobre o Desenvolvimento de uma Taxonomia para a Classificação de Trabalhos de um Mapeamento Sistemático sobre o uso de Ontologias em Informática Médica.** Computer on the Beach, Florianópolis, Santa Catarina, 2012.

MICHAEL, L. Q. A.; TIONG, A. W.; KANAGASABAI, R.; **Towards Large Scale Modeling and Realization of SNOMED CT in OWL-DL.** IEEE 23rd International Symposium on Computer-Based Medical Systems, 401-407, Australia, 2010.

MICHIN, R.; PORTO, F. A. M.; VANGENOT, C.; HARTMANN, S.; **Symptoms Ontology for Mapping Diagnostic Knowledge Systems.** In: Computer-Based Medical Systems, p. 593 – 598, 2006.

NOY, N., KLEIN, M.; **Ontology evolution: not the same as schema evolution.** Knowledge and Information Systems 6 (4), 428–440, 2004.

ORGUN, B.; VU, J. **HL7 ontology and mobile agents for interoperability in heterogeneous medical information systems.** Computers in biology and medicine, v. 36, n. 7-8, p. 817-36, 2006.

PETTICREW, M.; HELEN. R. **Systematic Reviews in the Social Sciences: A Practical Guide**, Blackwell Publishing, 2005.

RANDOLPH, J.J. **A Guide to Writing the Discussion Literature Review.** Pratical Assesment, Research & Evaluation, vol. 14, nº13, p. 1-13, Junho, 2009.

REGGIA, J.; NAU, D.; WANG, P.; **A Formal Model of Diagnostic Inference.** Information Sciences, v. 37, p.227, 1985.

RICHMOND, B. D. H.; **The Truth About Taxonomies.** Information Management. Março/Abril, 44-53, 2003.

RUBIN, D.L.; GENNARI, J.H.; MUSEN, M.A.; **Knowledge Representation and Tool Support for Critiquing Clinical Trial Protocols.** In: Proceedings of the AMIA Symposium 2000.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach.** 3rd Edition. New Jersey: Prentice Hall; 2009.

SAMWALDA, M.; CHENA, H.; RUTTENBERG, A.; LIMA, E.; MARENCOA, L.; MILLER, P.; SHEPHERD, G.; CHEUNG, K-H.; **Semantic SenseLab: Implementing the vision of the Semantic Web in neuroscience.** Journal of Artificial Intelligence in Medicine ,48: 21–28, 2010.

SARKAR, I. N. **Biomedical informatics and translational medicine.** Journal of Translational Medicine, 2010. p. 1 – 12.

SCHULZ, S.; BEISSWANGER, E.; WERMTER, J.; HAHN, U.; **Towards an Upper-Level Ontology for Molecular Biology.** AMIA Annual Symposium Symposium, 694-8, 2006.

SCHULZ, S.; SCHOBER, D.; TUDOSE, I.; HASTINGS, J.; BOEKER, M.; **Unintended consequences of existential quantifications in biomedical ontologies.** BMC Bioinformatics, 2011^a.

SCHULZ, S.; SMITH, B.; STENZHORN, H.; BOEKER, M.; **Adapting Clinical Ontologies in Real-World Environments.** Journal of Universal Computer Science; **14(22): 3767–3780, 2008.**

SCHULZ, S.; SPACKMAN, K.; JAMES, A.; COCOS, C.; BOEKER, M. **Scalable representations of diseases in biomedical ontologies.** Journal of Biomedical Semantics, 2 (Suppl 2):S6, 2011^b.

SIDDIQI, A. A.; AHMED, M.; ALGINAHI, Y. M.; ALHARBY, A. **Use of Information and Mobile Computing Technologies in Healthcare Facilities of Saudi Arabia.** International Conference on Information and Communication Technologies, Karachi, Pakistan, 2009. p. 1 -6.

SILACHAN, K.; TANTATSANAWONG, P. **Domain Ontology Health Informatics Service From Text Medical Data Classification.** Department of Computing, Faculty of Science, Silpakorn University. Nakornpratom, Thailand. Annual SRII Global Conference. 2011.

SMITH, B; WELTY, C. **Ontology-towards a new synthesis.** Proceedings of the international conference on Formal Ontology in Information Systems. New York, NY, 2001. p. 3 – 9.

SUASSUNA, M. **Mapeamento Sistemático sobre Replicação de Estudos Empíricos em Engenharia de Software.** Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco. Junho, 2011.

SUI, Z.; ZHAO, J.;KANG, W.; ZHAO, Q.; **The Building of a CBD-based Domain Ontology in Chinese.** IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 303-306, 2008.

SUJANSKY, W.; **Heterogeneous Database Integration in Biomedicine.** 34 (Issue 4), 285-98, 2001.

SUN, Y.; YANG, Y.; ZHANG, H.; ZHANG, W.; WANG, Q.; **Towards evidence-based ontology for supporting Systematic Literature Review.** Evaluation & Assessment in Software Engineering, 171 – 175, Espanha, May 2012.

SZOLOVITS, PETER; **Artificial Intelligence in Medicine.** AAAS Selected Symposium Series, V. 51, Westview Press, Boulder, Colorado, 1982.

THOMAS, D. G.; PAPPU, R. V.; BAKER, N. A.; **Ontologies for Cancer Nanotechnology Research.** International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society, 2-6, USA, September, 2009.

TRAVASSOS, G., BIOLCHINI J. **Revisões Sistemáticas Aplicadas a Engenharia de Software.** In: XXI SBES – Brazilian Symposium on Software Engineering, João Pessoa, PB, Brasil, 2007.

XIANG, Z.; COURTOT, M.; BRINKMAN, R. R.; RUTTENBERG, A.;HE, Y.; **OntoFox: web-based support for ontology reuse.** BMC Research Notes, 3:175, 2010.

YE, Y.; JIANG, Z.; DIAO, X.; DU, G.; **Knowledge-Based Hybrid Variance Handling for Patient Care Workflows based on Clinical Pathways.** Service Operations, Logistics and Informatics, 13-18, Chicago, E.U.A., 2009.

YU, C.; **Methods in biomedical ontology Alexander.** Journal of Biomedical Informatics 39: 252–266, Dezembro, 2006.

ZHEN, H.; JING-SONG, L.; HAI-YAN, Y.; XIAO-GUANG, Z.; MUNEOU, S.; KENJI, A.; **Modeling of Clinical Pathways Based on Ontology.** IEEE International Symposium on IT in Medicine & Education, 1170-1174, China, 2009.

ZHU, W.; FU, F.; XU, L.; ZHANG, B.; **A TCM Diagnosis System Based on Textbook Information Extraction.** In: IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing, 2011.

ZHU, X. YU, Y., JIA, T., ZHAO, D. AND LIU, J.; **An Information Extraction Strategy for Medical Imaging Diagnostic Report.** In: IEEE/ICME International Conference on Complex Medical Engineering, p. 388 – 392, 2007.

APÊNDICE A – Artigos de Revisão encontrados neste Mapeamento Sistemático

Tabela A1 – Artigos de Revisão indexados por esta pesquisa

Identificador	Título
REV01	Adaptive and Intelligent Systems for Collaborative Learning Support: A Review of the Field
REV02	Knowledge engineering for medical decision making: A review of computer-based clinical decision aids
REV03	The InterMed approach to sharable computer-interpretable guidelines: a review
REV04	Expert system methodologies and applications—a decade review from 1995 to 2004
REV05	Translational cognition for decision support in critical care environments: A review
REV06	A Review and Prospects on Collaborative Ontology Editing Tools
REV07	Cognitive and learning sciences in biomedical and health instructional design: A review with lessons for biomedical informatics education
REV08	Knowledge management technologies and applications—literature review from 1995 to 2002
REV09	Technology management methodologies and applications: A literature review from 1995 to 2003
REV10	Temporal representation and reasoning in artificial intelligence: A review
REV11	Knowledge based engineering: Between AI and CAD. Review of a language based technology to support engineering design
REV12	Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review
REV13	A knowledge-based taxonomy of critical factors for adopting electronic health record systems by physicians: a systematic literature review
REV14	Formal Representations of Eligibility Criteria: A Literature Review
REV15	Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review
REV16	Translational Cognition for Decision Support in Critical Care Environments: A Review
REV17	Cross-Topic Learning for Work Prioritization in Systematic Review Creation and Update
REV18	Inter-rater reliability and review of the VA unresolved narratives

REV19	Review of Informatics in Medical Imaging
REV20	A systematic review of research on integration of ontologies with the model-driven approach
REV21	Formal modelling, knowledge representation and reasoning for design and development of user-centric pervasive software: a meta-review
REV22	A review of interoperability and possibilities for data analysis from virtual world environments
REV23	Antonios-Konstantinos Thanellas and Mika Pollari. Compact Review of Structural and Microstructural Brain Analysis Methods
REV24	The intersection of clinical decision support and electronic health record: A literature review
REV25	Clinical decision support systems: a review on knowledge representation and inference under uncertainties
REV26	Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: A literature review of guideline representation models
REV27	A Review and Prospects on Collaborative Ontology Editing Tools
REV28	A review of auditing methods applied to the content of controlled biomedical terminologies
REV29	Situation identification techniques in pervasive computing: A review
REV30	Adapting Clinical Ontologies in Real-World Environments.
REV31	Review of Semantically Interoperable Electronic Health Records for Ubiquitous Healthcare
REV32	Conceptual Knowledge Acquisition in Biomedicine: A Methodological Review
REV33	Automatic de-identification of textual documents in the electronic health record: a review of recent research
REV34	Consumer language, patient language, and thesauri: a review of the literature
REV35	Knowledge bases in medicine: a review.
REV36	Detailed Clinical Models: A Review
REV37	Computerized decision support for concurrent utilization review using the HELP system.
REV38	Decision support for concurrent utilization review using a HELP-embedded expert system.
REV39	Clinical Information Systems—A Review
REV40	Histopathological Image Analysis: A Review
REV41	Medical Knowledge Base Acquisition: The Role of the Expert Review

	Process in Disease Profile Construction
REV42	A critical review of PASBio's argument structures for biomedical verbs
REV43	Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies
REV44	Agents applied in health care: A review
REV45	Computer-based execution of clinical guidelines: A review
REV46	Temporal reasoning with medical data—A review with emphasis on medical natural language processing
REV47	Formal representation of eligibility criteria: A literature review
REV48	An overview of the ONIONS project: Applying ontologies to the integration of medical terminologies
REV49	Conceptual knowledge acquisition in biomedicine: A methodological review
REV50	European research efforts in medical knowledge-based systems
REV51	A review of ontology based query expansion
REV52	Context awareness in health care: A review
REV53	Web information fusion: A review of the state of the art
REV54	Knowledge based engineering: Between AI and CAD. Review of a language based technology to support engineering design
REV55	Review of biomedical knowledge and data representation with conceptual graphs.
REV56	Medication-related Clinical Decision Support in Computerized Provider Order Entry Systems: A Review
REV57	Access to data: comparing AccessMed with Query by Review.
REV58	Overcoming Structural Constraints to Patient Utilization of Electronic Medical Records: A Critical Review and Proposal for an Evaluation Framework
REV59	Outreach activities of the National Library of Medicine: a five-year review.
REV60	Clinical information behavior of rehabilitation therapists: a review of the research on occupational therapists, physical therapists, and speech-language pathologists
REV61	Methods for the guideline-based development of quality indicators--a systematic review
REV62	A Review of Fuzzy Cognitive Maps research during the last decade
REV63	Harnessing Ontologies for Argument-Based Decision-Making in Breast Cancer
REV64	A versatile context-aware pervasive monitoring system: Validation and

	characterization in the health-care domain
REV65	Towards evidence-based ontology for supporting Systematic Literature Review

ANEXO I

Uma Taxonomia Ontológica para a Classificação de Estudos em Informática Médica

Moises Roberto de A. Mota¹, Esaú A. Tavares¹, Sidney Patrício R. Campos¹, Natasha C. Q. Lino¹

¹ Centro de Informática – Universidade Federal da Paraíba (UFPB)
CEP 58059-900 – João Pessoa – PB – Brasil

{esau, sidneypatrezio, moises.mota}@di.ufpb.br, [\[natasha\]@ci.ufpb.br](mailto:{natasha}@ci.ufpb.br)

Abstract. In view of the importance that Medical Informatics has acquired through the use of Artificial Intelligence techniques in their applications, especially concerning to Knowledge Engineering and Ontologies, it is necessary to map the evolution of this area thoroughly, impartially and systematically. Therefore, we built a Systematic Mapping that aims documenting about the use of Ontologies in Medical Informatics. By the fact that we don't find a general taxonomy that defines the subfields of Medical Informatics and that allows sorting the selected studies for Systematic Mapping, this paper presents a study about the construction of a taxonomy that characterizes all subareas of Medical Informatics.

Resumo. Tendo em vista a importância que a Informática Médica adquiriu, através da utilização de técnicas de Inteligência Artificial em suas aplicações, especialmente em relação ao Engenharia do Conhecimento e Ontologias, é necessário mapear a evolução desta área de modo completamente, imparcial e sistemático. Por isso, foi desenvolvido um mapeamento sistemático que tem como objetivo documentar o uso de ontologias em Informática Médica. Por não ter sido encontrada uma taxonomia geral que defina as subáreas da Informática Médica e permita a classificação dos estudos selecionados para mapeamento sistemático, este trabalho apresenta um estudo sobre a construção de uma taxonomia que caracterize as subáreas de Informática Médica.

1. Introdução

Representação do Conhecimento é a área da Inteligência Artificial que lida como o conhecimento deve ser representado simbolicamente e manipulado de forma automática por programas de raciocínio, Russell e Norvig (2009). Na Ciência da Computação, Ontologia é uma técnica de Representação do Conhecimento que se refere a um artefato de engenharia constituído por um vocabulário específico usado para descrever uma determinada realidade e por uma série de suposições sobre o significado pretendido do vocabulário. Ontologias descrevem uma

especificação formal sobre certo domínio, sendo um entendimento compartilhado sobre o mesmo e um modelo formal e executável por máquina.

Informática Médica surge como uma nova área de pesquisa que combina as áreas de Ciência da Computação, Tecnologias da Informação e Medicina, Silachan (2011); lida com o armazenamento, recuperação e uso da informação, dados e conhecimentos biomédicos para a resolução de problemas e tomadas de decisão.

Neste contexto, pesquisas em Informática Médica (IM) e Inteligência Artificial podem usufruir de inúmeras oportunidades para avançar o estado da arte em vários aspectos. Tecnologias como as Ontologias facilitam a integração de dados heterogêneos, bem como tornam possível a comunicação entre sistemas diferentes, além de tornar certos tipos de dados passíveis de processamento automático. Este tipo de recurso aplicado à Informática Médica torna possível certas operações antes impraticáveis, por exemplo: o diagnóstico médico baseado em várias fontes diferentes de dados, a digitalização de prontuários médicos, a mineração de dados em bases de clínicas, etc.

Devido ao uso crescente de Ontologias na Informática Médica, tornou-se necessária a obtenção e documentação de informações, de maneira eficiente, sobre o estado da arte do uso de Ontologias na Informática Médica, com o objetivo de promover uma visão detalhada de como as aplicações são desenvolvidas nessa área, quais seus benefícios e principais técnicas utilizadas. Assim, contribuindo para o crescimento da área, com ênfase nas suas necessidades e deficiências, evitando a criação de estudos duplicados ou não-relevantes. Dessa forma, optou-se pelo desenvolvimento de um Mapeamento Sistemático sobre o uso de Ontologias na Informática Médica.

Segundo Kitchenham (2007), o Mapeamento Sistemático é uma metodologia de pesquisa e avaliação de literatura que tem como objetivo identificar todas as pesquisas relacionadas a um tópico específico, ou seja, responder a questões mais amplas relacionadas à evolução da investigação de uma determinada área da ciência. Deste modo, tem-se uma visão geral da evolução de uma determinada área de pesquisa até o momento da construção do mapeamento. Além de documentá-la de maneira sistemática permitindo sua replicação, avaliação e posterior atualização.

O Mapeamento Sistemático (MS) é composto por três etapas, Kitchenham (2007), são elas: Etapa de Planejamento, Etapa de Condução e Etapa de Reporte. Todas essas etapas precisam ser bem documentadas, Arksey e O'Malley (2005). Os trabalhos de Jacinto (2010), Almeida (2011), Cardozo (2012), abordam que o processo para escolha, ou desenvolvimento, de uma taxonomia adequada para a classificação dos trabalhos de um MS nem sempre é detalhado, visto que, geralmente é utilizada uma taxonomia validada e bem conceituada pela comunidade acadêmica para qual é direcionado o MS. Por exemplo, no trabalho de Suassuna (2011) foi utilizado o SWEBOK, seus capítulos e sessões para classificar os estudos selecionados para o Mapeamento Sistemático.

Para o presente trabalho não foi encontrada, até o momento, uma taxonomia com os requisitos necessários. Desse modo, foi necessário o desenvolvimento de uma taxonomia que

proporcione uma estruturação de dados em estudos cujo foco é o uso de Ontologias em Informática Médica.

2. Planejamento do Mapeamento Sistemático

Dentro da etapa de Planejamento do Mapeamento, a atividade mais importante, de acordo com Kitchenham (2007), é a criação das questões de pesquisa. As questões de pesquisa são o foco do mapeamento, e são elas o guia da pesquisa. Além disso, as questões de pesquisa atuam como um dos fatores de inclusão e exclusão na busca automática de estudos e trabalhos, pois apenas são considerados artigos que abordem o tema de alguma dessas questões. As questões de pesquisa que orientaram este trabalho são:

- Qual(is) o(s) principal(is) uso(s) de Ontologias em Informática Médica?
- Qual(is) o(s) benefícios do uso de Ontologias em Informática Médica?
- Quais as ontologias utilizadas em Informática Médica?
- É considerado o reuso de ontologias nos projetos de Informática Médica?
- Quais as ferramentas utilizadas para a criação e edição de Ontologias?
- Quais repositórios de Ontologias são utilizados?
- Como os trabalhos de Ontologias em Informática Médica estão sendo validados/avaliados?

O próximo passo consiste em especificar os engenhos de busca para a execução da busca automática. Os engenhos de busca apresentados na Tabela 1 foram selecionados por possuírem, segundo a literatura pesquisada, Dyba (2005), Kitchenham (2007), Travassos (2007), Brereton (2007), boa parte das pesquisas desenvolvidas no domínio da Ciência da Computação. Foram excluídos engenhos pela indisponibilidade de seus artigos, por não possuir convênio com a Universidade Federal da Paraíba ou por abordarem temas que não são de interesse desta pesquisa.

Também foram definidos periódicos para compor a busca manual, como: International Journal of Metadata, Semantics and Ontologies (Qualis B2), International Journal of Medical Informatics (Qualis A2), Journal of Biomedical Informatics (Qualis A1), Journal of the American Medical Informatics Association (Qualis A1), Medical Informatics and the internet in Medicine (Qualis B2) e Applied Ontology (Qualis B2).

Em seguida, é preciso construir uma String de Busca, Kitchenham (2007) e Travassos (2007), que deve ser expressivo para representar o conjunto de termos encontrados nas questões de pesquisa e recuperar os trabalhos relevantes para o mapeamento com precisão e cobertura.

Tabela 1. Engenhos utilizados na busca automática

Engenho de busca	Endereço Web
ACM	dl.acm.org/dl.cfm
IEEE	ieeexplore.ieee.org/
Science Direct	sciencedirect.com/
PubMed	www.ncbi.nlm.nih.gov/pubmed
PMC	www.ncbi.nlm.nih.gov/pmc/
Google Scholar	scholar.google.com.br/

A String de Busca é construída a partir da concatenação dos termos encontrados nas questões de pesquisa através de operadores booleanos (“AND” e “OR”). Uma vez desenvolvida a String de Busca, vide tabela 2, é possível iniciar a pesquisa nos engenhos selecionados. A String de Busca para este trabalho em específico esta definida abaixo.

Tabela 2. String de Busca

("Knowledge Engineering" OR "Knowledge Representation" OR "Ontology Engineering") AND ("Health" OR "E-Health" OR "Electronic Health" OR "Medical Informatics" OR "Medical Systems" OR "Medical" OR "Clinical Decision Support" OR "Medical Decision Support") AND ("Ontology" OR "Ontologies")

Para auxiliar a seleção dos estudos relevantes para um Mapeamento Sistemático é preciso que existam critérios de inclusão e exclusão, Kitchenham (2007), Travassos (2007).

- Os seguintes tipos de estudos serão incluídos nesta avaliação sistemática:
- Estudos completos publicados em revistas ou conferências revisadas sobre o uso de Ontologias em Informática Médica e que respondam por uma das questões de pesquisa;
- Estudos secundários, ou seja, que dependam de estudos primários;
- Estudos teóricos com o objetivo de apresentar conceitos para o entendimento da área;
- Estudos experimentais relacionados à área;
- Apenas trabalhos escritos em língua inglesa serão considerados.

Os seguintes estudos serão excluídos:

- Estudos que não estejam claramente relacionados à área de Ontologias e Informática Médica;
- Estudos que não respondem a nenhuma das questões de pesquisa;
- Artigos duplicados, ou seja, aqueles encontrados em mais de uma fonte da busca automática e/ou manual;
- Artigos convidados, tutoriais, relatórios técnicos, teses, dissertações e relatórios de workshops que sejam incompletos;
- Estudos não disponíveis para download nos engenhos de busca definidos.

Os resultados das buscas foram os seguintes: ACM: 55 trabalhos encontrados dos quais 15 foram aceitos; IEEE: 2088 foram encontrados, 215 aceitos; Science Direct: 1712 encontrados, 255 aceitos; Google Scholar: 18300 encontrados, 13 aceitos; PubMed: 64 encontrados, 29 aceitos e PMC: 1375 encontrados, 263 aceitos.

A aplicação da string de busca nos periódicos trouxe resultados em menor número e de pequena relevância para o trabalho.

A seguir, apresenta-se o processo de desenvolvimento da taxonomia para classificação dos artigos aceitos no Mapeamento Sistemático.

3. Estudos sobre Taxonomias

Observando os trabalhos de Feldt et al (2008), que detalham a construção de sua própria taxonomia baseada principalmente em palavras-chaves encontradas nos resumos dos estudos selecionados, e Dixon et al (2007), que sugerem uma categorização dos termos existentes na literatura médica; iniciou-se a definição de uma estratégia para a construção de uma taxonomia que represente as subáreas da Informática Médica. Para a execução do presente trabalho foram pesquisadas e estudadas as taxonomias das seguintes fontes atuantes em diversas partes do mundo, são elas: *European Federation for Medical Informatics* (EFMI); Sociedade Brasileira de Informática em Saúde (SBIS); *American Medical Informatics Association* (AMIA); Associação Portuguesa de Informática Médica (APIM); *Medical Informatics - Practical Guide for the Healthcare Professional*, Hoyt (2009); *China-Japan-Korea Joint Symposium on Medical Informatics* (CJKMI); *Asia Pacific Association for Medical Informatics Conference* (APAMI); *South African Health Informatics Association Conference* (SAHIAC)

Foram considerados também termos presentes nos principais congressos, periódicos e associações, que se distribuem em diversos continentes e definem as subáreas da Informática Médica.

3.1. Taxonomias

Ao analisar as fontes definidas na sessão anterior, foi percebido que determinadas categorizações encontradas são específicas para região em que a associação, congresso ou conferência se destina, por exemplo, categorias como *Informatics of Traditional Chinese Medicine* só foram identificadas na *Asia Pacific Association for Medical Informatics Conference*. Também é possível notar que na taxonomia proposta pela *European Federation for Medical Informatics* existem categorias direcionadas para a mesma subárea da Informática Médica, mas com focos específicos, por exemplo: *Knowledge discovery in biomedical databases for decision support*, *Intelligent interoperability and telemedicine Data and knowledge management for decision support in forensic medical disciplines*, *eHealth decision support systems for GPs, clinicians, nurses, health care managers and patients*, *Evaluation of decision support systems Diagnostic, therapeutic and prognostic decision support*; ou seja, essas categorias podem ser estruturadas numa hierarquia da subárea *Clinical Decision Support*. Entre outras, a taxonomia da *American Medical Informatics Association*, apresenta este mesmo detalhe com as seguintes categorias: *Clinical Research Informatics*, *Clinical Informatics*.

A taxonomia construída neste trabalho foi desenvolvida a partir da comparação entre as diversas categorias estudadas, como veremos na sessão posterior.

4. Análises comparativas e desenvolvimento da Taxonomia

É importante salientar a existência de outras taxonomias que definem a Informática Médica, porém, direcionadas para contextos específicos, Dixon (2007), como as propostas pelo HIMSS (*Healthcare Information Management System Society*) e pela MESH (*Medical Subject Headings*).

Segundo Dixon (2007), a taxonomia proposta pelo HIMSS aborda uma parte das principais subáreas da Informática Médica, porém, é estritamente direcionada para sistemas de prontuários médicos eletrônicos, desconsiderando certas aplicações da informática na medicina, como telemedicina e o compartilhamento de informações em saúde. Portanto, não foi considerada adequada para classificar os trabalhos selecionados para o Mapeamento Sistemático, uma vez que a proposta é o desenvolvimento de uma taxonomia geral que represente as principais subáreas existentes para esta grande área da ciência. Posteriormente esta hierarquia de termos foi atualizada e expandida pelo MEDLINE, incluindo a terminologia “tecnologias da informação em saúde”. Entretanto, na hierarquia proposta pelo MESH há falta de diversos conceitos da Informática Médica principalmente no que diz respeito à implementação de prontuário eletrônico em saúde.

4.1. Metodologia

Conforme citado na sessão anterior, a taxonomia construída foi desenvolvida utilizando como base os estudos de Dixon (2007), para tanto, foi realizada uma comparação entre os termos presentes nas taxonomias estudadas. Posteriormente, foram selecionados os termos semelhantes ou correlacionados entre as taxonomias, de acordo com sua definição nas fontes citadas anteriormente, Feldt (2008), Richmond (2003). Uma vez selecionados, os mesmos foram agregados para a construção de uma única categoria geral. Por exemplo, os seguintes termos: Prontuário Eletrônico do Paciente (SBIS), *Electronic Medical Records/Electronic Health Records* (CJKMI), foram agregados ao termo *Electronic Health Records*, Hoyt (2009), por considerarmos este como o termo mais abrangente e que poderia classificar qualquer trabalho direcionado ao desenvolvimento de Registros Eletrônicos aplicados a Saúde.

4.2. Resultados da Análise Comparativa

Categorias como *Electronic Prescribing e Mobile Technology* encontradas em Hoyt, (2009) foram aglutinadas num só termo, *Consumer Health Informatics*. Este novo termo reúne todo estudo que desenvolva produtos comerciais, ou não, voltados para pacientes. Esta decisão está relacionada à constatação durante a pesquisa de que trabalhos nestas áreas são desenvolvidos para esta finalidade.

Internet em Saúde (SBIS), *Networks* [Hoyt 2009], *Online Medical Resources* [Hoyt 2009], *Search Engines* [Hoyt 2009], *Internet of Things* (APAMI) e *Cloud Computing* (APAMI) foram

reunidas na categoria *Online Medical Resources*. Dessa forma, os estudos relacionados aos usos de recursos web na Informática Médica são classificados nesta categoria, bem como serviços online.

As categorias *Formalization of knowledge, ontologies, clinical guidelines and standards of health care* (EFMI), *Standard in Health Informatics and cross-language solution* (CJKMI), *Standards in Medical Informatics* (APAMI) e Padronização da Informação em Saúde (SBIS) foram aglutinadas no termo *Standards in Medical Informatics*. Os estudos aqui classificados são direcionados para definição de padrões de abrangência mundial para a aplicação de ontologias em Informática Médica.

Information Technology Infrastructure (SAHIAC), *Privacy, Security, and Confidentiality Issues in Medical Information* (APAMI), *Medical safety management* (CJKMI), *Human Resources Management* (SAHIAC) e *Integrated Practice Management Systems* [Hoyt 2009] foram agrupados na categoria *Infrastructure of Medical Information Systems*, direcionada para estudos que desenvolvem solução estrutural para sistemas de Informática Médica baseados em ontologias.

Intelligent interoperability and telemedicine (EFMI) e *Health Information Technology Interoperability* [Hoyt 2009] foram resumidas na categoria *Health Management Information Systems*. Esta Categoria reunirá estudos direcionados para o desenvolvimento de sistemas de Informática Médica baseados em ontologias.

Pay for Performance (P4P) [Hoyt 2009] e *Financial management for improved health service delivery* (APAMI) foram reunidas no termo *Financial management for improved health service delivery*, cujo objetivo é reunir os trabalhos voltados para a administração de recursos financeiros para o desenvolvimento de sistemas de Informática Médica.

O termo *Informatics of Traditional Chinese Medicine* foi transformado em uma categoria exclusiva, tendo em vista a especificidade dos estudos desenvolvidos relacionado a este termo encontrado entre os estudos selecionados para compor o Mapeamento Sistemático.

5. Resultados e Discussão

Foram 23 termos considerados ideais para classificação dos estudos selecionados para o Mapeamento Sistemático. Acredita-se que esta taxonomia abrange todas as subáreas da Informática Médica, desde as mais gerais, como *Telemedicine*, até as mais específicas, como *Nursing Informatics*. O processo de desenvolvimento desta taxonomia foi baseado nos trabalhos de Dixon (2007) e Feldt (2008), de modo semelhante a estes estudos, foram selecionados, criados e agrupados termos de diversas fontes de acordo com a necessidade para o desenvolvimento do Mapeamento Sistemático.

Na Tabela 3 foram definidos os termos da taxonomia, baseadas nas definições de Hoyt (2009), visto que foi a única definição formal das subáreas da Informática Médica as quais tivemos acesso:

Tabela 3. Taxonomia Final e Definição dos termos

Termos da Taxonomia	Definição
<i>Applications of Bioinformatics in Medical Informatics</i>	Bioinformática é uma área da ciência direcionada para aplicação de tecnologias informáticas em biologia, especificamente em genética.
<i>Applications of Knowledge Engineering in Medical Informatics</i>	O desenvolvimento de ontologias é uma técnica de engenharia do conhecimento, e por isso essa categoria é reservada para todos os estudos que propõem a criação de ontologias, bem com o desenvolvimento de metodologias para validações e análises das mesmas.
<i>Artificial Intelligence in Medicine and Its Applications</i>	Categoria direcionada para estudos que propõem a aplicação de técnicas de Inteligência Artificial em sistema na Informática Médica.
<i>Biomedical Eletronic and Robótics Devices</i>	Esta subárea da Informática Médica tem como objetivo o desenvolvimento de sistemas biomecânicos ou robóticos que proporcionem melhor qualidade de vida do paciente.
<i>Clinical Informatics</i>	Consiste em uma das diversas aplicações da Informática Médica. Ou seja, na criação, gerenciamento, compartilhamento e aplicação de guias de práticas clínicas para tratamento de determinadas doenças
<i>Consumer Health Informatics</i>	Esta subárea é voltada para o desenvolvimento de aplicações e serviços móveis direcionados para o consumidor.
<i>Decision Support Systems</i>	Qualquer software projetado para ajudar diretamente na tomada de decisão clínica em que as características de cada paciente são compatíveis com uma base de conhecimento informatizada.
<i>Disease Management and Disease Registries</i>	O principal objetivo desta subárea é auxiliar na gestão de pacientes crônicos através da utilização dos recursos aqui apresentados.
<i>Electronic Health Records</i>	Conhecidos como prontuários eletrônicos, pode ser considerado o principal foco do desenvolvimento da informática médica.
<i>Evidence Based Medicine</i>	Consiste no auxílio tecnológico a avaliação das subáreas da medicina baseada na busca por evidências. Ou seja, a avaliação da qualidade dos estudos nas diversas áreas da medicina
<i>Financial management for improved health service delivery</i>	Esta subárea consiste na definição da tributação e o retorno financeiro aos clínicos por aplicações na área clínica, ou médica.
<i>Health informatics education</i>	Consiste em estudos que utilizam ontologias com o objetivos educacionais direcionados para Informática Médica.
<i>Health Management Information Systems</i>	Categoria direcionada para estudos cujo foco é o desenvolvimento e gerenciamento de sistemas informação na área de Informática Médica.
<i>Informatics of Traditional Chinese Medicine</i>	Categoria reservada para a medicina tracional chinesa.

<i>Infrastructure of Medical Information Systems</i>	Esta categoria é direcionada para estudos que propõe melhorias nas infra-estruturas dos sistemas em Informática Médica.
<i>Medical Data Representation, Mining, Visualization, and Sharing</i>	Consiste no desenvolvimento de técnicas que permitam o aprimoramento das tecnologias atuais de representação, mineração, visualização e compartilhamento de dados médicos.
<i>Medical Signal and Image Processings</i>	Subárea voltada para o processamento e análise de sinais e imagens médicas, além de abranger a transmissão, armazenamento e formato de dados desta natureza.
<i>Nursing Informatics</i>	Esta categoria é reservada para estudos voltados para o auxílio de profissionais da área de enfermagem.
<i>Online Medical Resources</i>	Consiste no desenvolvimento de bases de dados de informações médicas confiáveis, de modo que possam ser consultados para finalidades como educação, recuperação de informação por médicos e pacientes, entre outros.
<i>Patient Informatics</i>	Estudos cujo objetivo é melhorar o tratamento de pacientes em hospitais crônicos e terminais, inclusive evitando erros médicos.
<i>Public Health Informatics</i>	Subárea direcionada para uso de Informática Médica direcionada para as populações e não indivíduos.
<i>Standards in Medical Informatics</i>	A definição de padrões para a área da Informática Médica é importante para a interoperabilidade de sistemas diferentes.
<i>Telemedicine</i>	Consiste na transferência de dados administrativos relacionado a instituições de saúde e a monitoração remota de pacientes respectivamente.

A validação da taxonomia ocorreu na etapa da análise qualitativa, na qual, foi realizada a leitura de todos os artigos selecionados para o Mapeamento Sistemático com o objetivo de identificar suas contribuições e classificar sua(s) área(s) de atuação. Dessa maneira, pode-se identificar que o intuito da taxonomia abranger todos os trabalhos selecionados para o MS foi atingido. No entanto, é importante deixar claro que todo o processo de desenvolvimento de um Mapeamento Sistemático evolui e se adapta conforme as necessidades apresentadas durante o estudo, deste modo, esta taxonomia ainda poderá ser modificada se for identificada esta necessidade.

5.1. Taxonomia Ontológica

Com base na análise das terminologias estudadas e na taxonomia final, também foi desenvolvido o protótipo de uma Taxonomia Ontológica, como pode ser visto na Figura 1.

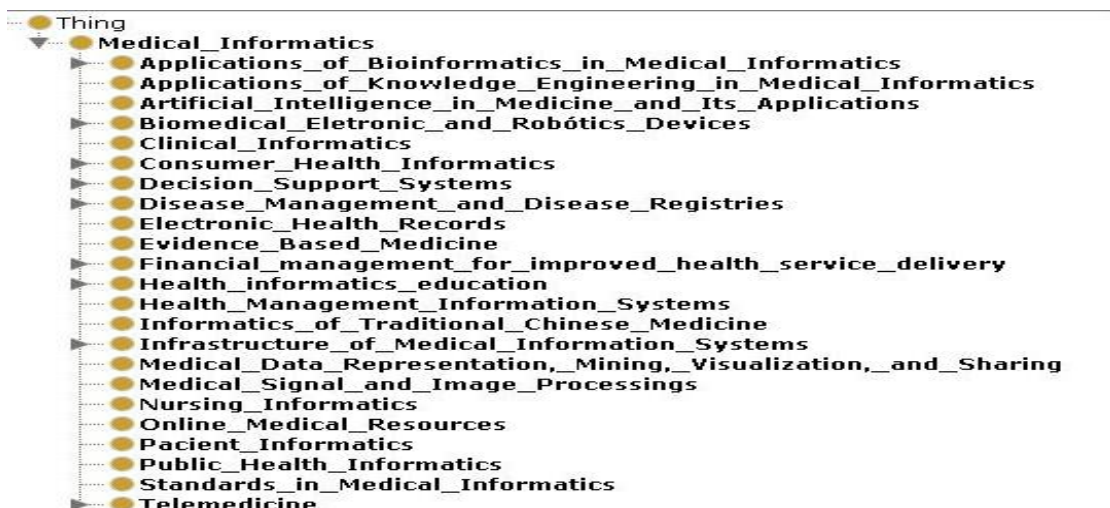


Figura 17. Taxonomia Ontológica desenvolvida no presente trabalho

Uma taxonomia serve para classificar informações em uma hierarquia (árvore) utilizando o relacionamento pai-filho (generalização ou “tipo-de”). Uma Ontologia permite atribuir características a cada termo existente na taxonomia, fazendo uma distinção clara entre suas definições, Breitman (2010). Dessa maneira, nenhum estudo poderá ser agrupado, classificado ou relacionado erroneamente com algum termo incompatível com suas características. Portanto, através do desenvolvimento da Taxonomia Ontológica, é possível um compartilhamento e organização mais eficientes de informações presentes em bases de dados relacionadas a Informática Médica. Por ser concebida em OWL, de acordo com os padrões do W3C, sua integração com projetos como os propostos por Borges (2011) e Kamura (2011) bem como com ferramentas de buscas baseadas em tecnologias semânticas, além de ser de grande utilidade para o desenvolvimento da Web Semântica, Hausenblas (2010).

6. Conclusões

O desenvolvimento de uma taxonomia que defina as principais subáreas da Informática Médica para a classificação de estudos de um Mapeamento Sistemático sobre Ontologias nesta área se torna mais expressivo, uma vez que o trabalho avançou no sentido do desenvolvimento de uma Taxonomia Ontológica, além de cumprir com seu objetivo inicial. Deste modo os resultados foram além do pretendido. Uma taxonomia que categorize a Informática Médica em subáreas é importante para estruturação de bases de dados relacionadas a área, já que não foi encontrada até o momento uma que se adequasse a necessidade deste estudo, bem como fosse abrangente o suficiente para permitir a classificação de estudos concentrados em áreas inovadoras e específicas que as terminologias utilizadas como base não permitiram.

No mais, a Taxonomia Ontológica não é apenas importante para este trabalho e sua divulgação, como também para o compartilhamento e estruturação de informações na web. Por ser concebida com base em tecnologias semânticas e dentro dos padrões definidos pela W3C, é

possível uma fácil integração com projetos como os citados por Heath (2011), permitindo também a integração de diferentes bases de dados distribuídas pela web, Hausenblas (2010).

7. Perspectivas futuras

O principal objetivo deste estudo foi o desenvolvimento de uma taxonomia que represente todas as subáreas da Informática Médica, para a classificação dos estudos selecionados para o Mapeamento Sistemático sobre o uso de Ontologias em Informática Médica que se encontra em desenvolvimento.

Após a conclusão do Mapeamento Sistemático, a eficácia da taxonomia, aqui proposta, será validada e posteriormente publicada, visando sua avaliação diante da comunidade científica relacionada, permitindo que outros tipos de estudos sejam desenvolvidos utilizando esta taxonomia como uma das suas referências.

Com a classificação dos estudos, da área de Informática Médica, de acordo com taxonomia proposta e após a finalização do Mapeamento Sistemático, pode-se identificar quais as subáreas da Informática Médica com maior desenvolvimento, e quais as que se beneficiam mais com o uso das tecnologias semânticas, como as Ontologias. Contribuindo, deste modo, para o desenvolvimento igual da área através da demonstração das suas necessidades e lacunas.

Pretendemos também validar, estender e refinar a Taxonomia Ontológica, aqui desenvolvida, para que a mesma permita a estruturação de bases de dados de estudos relacionados à Informática Médica e o compartilhamento de conhecimento mais eficiente através de aplicações que utilizem tecnologias semânticas Hausenblas (2010) e Heath (2011).

8. Referências

- Almeida, A. T. (2011). Um Mapeamento Sistemático de Mecanismos para Guiar Estudos Empíricos em Engenharia de Software. Dissertação (Mestrado em Ciências da Computação). Universidade Federal de Pernambuco. Recife, Pernambuco.
- Arksey, H and O'Malley, L. (2005). Scoping Studies: towards a methodological framework. *International Journal of Social Research Methodology*, vol. 8, n. 1, 19-32.
- Borges, M. R. S., Campos, M. L. M., Cordeiro, K. F. and Marino, T. (2011). Use of Linked Data in the Design of Information Infrastructure for Collaborative Emergency Management System. In *Proceedings of 15th International Conference on Computer Supported Cooperative Work in Design*, 764-771.
- Breitman, K. K. (2010). *Web Semântica: A Internet do Futuro*. LTC – Livros Técnicos e Científicos Ltda. Rio de Janeiro, Brasil.
- Brereton, P., Kitchenham, B. A., Budgen, D. T., Mohamed, M. and Mohamed, K. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *JSS* 80, 571-583.

- Cardozo, E. S. F. (2012). Mapeamento Sistemático sobre o uso do Auto-Gerenciamento em Equipes de Desenvolvimento de Software. Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco.
- Dixon, B. E., McGowan, J. J. and Zafar, A. (2007). Development of a Taxonomy for Health Information Technology. In Proceedings of World Congress on Medical and Health Informatics, Australia.
- Dyba, T., Jorgensen, M. and Kintchenham, B. (2005). Evidence-based software engineering for practitioners. IEEE Software, 58 – 65.
- Feldt, R., Petersen, K., Mattsson, M. and Mujtaba, S. (2008). Systematic Mapping Studies in Software Engineering. In Proceedings of 12th International Conference on Evaluation and Assessment in Software Engineering, Bari, Italy, 71-80.
- Hausenblas, M. and Karnstedt, M. (2010). Understanding Linked Open Data as a Web-Scale Database .In Proceedings of Second International Conference on Advances in Databases, Knowledge, and Data Applications, 56-61.
- Heath, T. (2011); Linked Data — Welcome to the Data Network. IEEE Internet Computing, vol. 11, 70-73.
- Hoyt, R. E., Sutton, M., and Yoshihashi, A. (2009). Medical Informatics - Practical Guide for the Healthcare Professional. 3ª Edição, University of West Florida, School of Allied Health and Life Sciences. Lulu.com.
- Jacinto, S. S. (2010). Um Mapeamento Sistemático de Pesquisa sobre a Influência da Personalidade na Engenharia de Software. Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco.
- Kamura, T., Kato, F., Ohmukai, I., Takahashi, T., Takeda, H. and Ueda, H. (2011). Study support and integration of cultural information resources with Linked Data. In Proceedings of Second International Conference on Culture and Computing, 177 - 178.
- Kitchenham, B. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. Vol 2.3 EBSB Technical Report, EBSE-2007-01.
- Richmond, B. (2003). D. H. The Truth About Taxonomies. Info Mgmt. (Mar/Apr. 2003), 44-53.
- Russell, S. and Norvig, P. (2009). Artificial Intelligence: A Modern Approach. 3rd Edition. New Jersey: Prentice Hall.
- Silachan, K. (2011). Tantatsanawong, P. Domain Ontology Health Informatics Service From Text Medical Data Classification. Department of Computing, Faculty of Science, Silpakorn University. Nakornpratom, Thailand. In Proceedings of Annual SRII Global Conference.
- Suassuna, M. (2011). Mapeamento Sistemático sobre Replicação de Estudos Empíricos em Engenharia de Software. Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco.
- Travassos, G. and Biolchini J. (2007). Revisões Sistemáticas Aplicadas a Engenharia de Software. In Proceedings of XXI SBES – Brazilian Symposium on Software Engineering, João Pessoa, PB, Brasil.

ANEXO II

Um Estudo Sobre o Desenvolvimento de uma Taxonomia para a Classificação de Trabalhos de um Mapeamento Sistemático sobre o uso de Ontologias em Informática Médica

Moises Roberto de A. Mota¹, Esaú A. Tavares¹, Sidney Patrício R. Campos¹, Natasha C. Q. Lino¹

¹ Centro de Informática – Universidade Federal da Paraíba (UFPB)
CEP 58059-900 – João Pessoa – PB – Brasil

{esau, sidneypatrezo}@di.ufpb.br, {moises.mota, natasha}@ci.ufpb.br

Abstract. *In view of the importance that Medical Informatics has acquired through the use of Artificial Intelligence techniques in their applications, especially concerning to Knowledge Engineering and Ontologies, it is necessary to map the evolution of this area thoroughly, impartially and systematically. Therefore, we built a Systematic Mapping that aims documenting about the use of Ontologies in Medical Informatics. By the fact that we don't find a general taxonomy that defines the subfields of Medical Informatics and that allows sorting the selected studies for Systematic Mapping, this paper presents a study about the construction of a taxonomy that characterizes all subareas of Medical Informatics.*

Resumo. *Tendo em vista a importância que a Informática Médica tem adquirido juntamente com a utilização de técnicas de Inteligência Artificial em suas aplicações, principalmente no que diz respeito à Engenharia do Conhecimento e Ontologias, faz-se necessário mapear a evolução desta área de maneira completa, imparcial e sistemática. Para tanto foi construído um Mapeamento Sistemático sobre o uso de Ontologias em Informática Médica. Como não há uma taxonomia geral que defina as subáreas da Informática Médica, este trabalho apresenta um estudo sobre a construção de uma taxonomia que caracterize todas as subáreas da Informática Médica que permita classificar os estudos selecionados para o Mapeamento Sistemático.*

Palavras-Chave. *Informática Médica, Ontologias, Mapeamento Sistemático, Taxonomia*

1. Introdução

Informática Médica surge como uma nova área de pesquisa que combina as áreas de Ciência da Computação, Tecnologias da Informação e Medicina. Integra elementos como Ciência da Informação, Tecnologia Informática e Históricos Médicos relacionados à saúde e serviços, [Silachan and Tantatsanawong 2011], lida com o armazenamento, recuperação e uso da informação, dados e conhecimentos biomédicos para a resolução de problemas e tomadas de decisão. Dentre as linhas de pesquisas na área de Informática Médica, pode-se destacar: digitalização de prontuários médicos, extração de informações em prontuários digitais, reconhecimento de documentos em linguagem natural, ferramentas de apoio ao diagnóstico clínico/médico, mineração de dados em bancos de

informações médicas, reconhecimento e diagnósticos baseados em imagens (radiografias, mamografias, etc.), reconhecimento de linguagem natural, etc. Ou seja, é perceptível que as técnicas de Inteligência Artificial são parte integrante da Informática Médica através das linhas de pesquisa relacionadas a mineração de dados, suporte ao diagnóstico clínico, etc.

A Representação do Conhecimento é a área da Inteligência Artificial que lida como o conhecimento pode ser representado simbolicamente e manipulado de forma automática por programas de raciocínio, [Russel and Norvig 2009]. Na Ciência da Computação, Ontologia é uma técnica de Representação do Conhecimento que se refere a um artefato de engenharia constituído por um vocabulário específico usado para descrever uma especificação formal sobre certo domínio, sendo um entendimento compartilhado sobre o mesmo e um modelo formal e executável por máquina.

Neste contexto, as pesquisas em Informática Médica (IM) agregadas a Inteligência Artificial, tem usufruído de inúmeras oportunidades, mediante a utilização de ontologias, para avançar o estado da arte em vários aspectos, principalmente no tocante a diagnósticos mais eficientes de determinadas doenças, redução de erros médicos, entre outros. Tecnologias como as Ontologias facilitam a integração de dados heterogêneos, bem como tornam possível a comunicação entre sistemas diferentes, além de tornar certos tipos de dados passíveis de processamento automático.

Este tipo de recurso aplicado a Informática Médica torna possível certas operações antes impraticáveis, por exemplo: o diagnóstico médico baseado em várias fontes diferentes de dados, a digitalização de prontuários médicos, a mineração de dados em bases de clínicas, etc. Diversas pesquisas têm sido desenvolvidas neste sentido, [Sarkar 2010], [Dang *et al.* 2008].

Portanto é necessário documentar, de maneira eficiente, o estado da arte do uso de Ontologias na Informática Médica, promovendo uma visão detalhada de como estas aplicações são desenvolvidas, quais seus benefícios, principais técnicas utilizadas: para deste modo contribuir para o crescimento da área mostrando suas necessidades e deficiências, bem como evitando a criação de estudos duplicados ou não-relevantes.

Segundo [Kitchenham 2007], o Mapeamento Sistemático é um método sistemático de pesquisa e avaliação de literatura que tem como objetivo identificar todas as pesquisas relacionadas a um tópico específico, ou seja, responder a questões mais amplas relacionadas à evolução da investigação. Deste modo, tem-se uma visão geral da evolução de uma determinada área de pesquisa, apresentando sua evolução até o momento da construção do mapeamento além de documentar a área investigada de maneira sistemática permitindo sua replicação, avaliação e posterior atualização.

O Mapeamento Sistemático (MS) é composto por três etapas, [Kitchenham 2007], a saber: Etapa de Planejamento; Etapa de Condução; Etapa de Reporte. Todas as etapas que compõem um MS precisam ser bem documentadas, [Arksey and O'Malley 2005], porém durante a Etapa de Condução percebemos, na literatura utilizada como base, [Jacinto 2010], [Almeida 2011], [Cardozo 2012], que a escolha, ou desenvolvimento, de uma taxonomia adequada para a classificação dos trabalhos a serem avaliados para o MS não é detalhada, visto que geralmente é utilizada uma taxonomia validada e bem conceituada pela comunidade acadêmica tida como principal público do MS, como no trabalho de [Suassuna 2011], no qual foi utilizado o SWEBOK, seus capítulos e sessões para classificar os estudos selecionados para o Mapeamento Sistemático.

No entanto conforme pode ser verificado no trabalho de [Feldt *et al.* 2008], o processo de desenvolvimento de uma taxonomia para a classificação de estudos de um Mapeamento Sistemático precisa ser documentado em determinados casos. Portanto, para o presente trabalho visto que não foi

encontrada, até o presente momento, uma taxonomia validada e bem conceituada no que diz respeito às principais subáreas da Informática Médica, foi necessário o desenvolvimento de uma taxonomia que permitisse a classificação dos estudos selecionados para o Mapeamento Sistemático sobre o uso de Ontologias em Informática Médica.

2. Estudos sobre taxonomias

[Feldt *et al.* 2008], desenvolveu sua própria taxonomia baseados principalmente nas palavras-chaves encontradas nos resumos dos estudos primários selecionados para o Mapeamento Sistemático juntamente com categorizações propostas por [Wieringa *et al.* 2006], as quais são definidas pelas contribuições dos estudos para a área a que se destinam.

No presente trabalho foi identificada a necessidade de uma taxonomia para a classificação dos estudos selecionados para o Mapeamento Sistemático, visto que após uma vasta pesquisa em livros, congressos, conferências e associações nacionais e internacionais, não foi encontrada uma taxonomia geral, validada e bem conceituada que defina as subáreas da Informática Médica. Para a execução do presente trabalho foram pesquisadas as taxonomias das seguintes fontes, vide tabela 1.

Tabela 1. - Fontes para a construção da taxonomia proposta

Nome	Abreviação	Tipo	Fonte	Acessado em
European Federation for Medical Informatics	EFMI	Federação	http://www.helmholtz-muenchen.de/ibmi/efmi/	Outubro 2012
Sociedade Brasileira de Informática em Saúde	SBIS	Sociedade	http://www.sbis.org.br/	Outubro 2012
American Medical Informatics Association	AMIA	Associação	http://www.amia.org/	Outubro 2012
Associação Portuguesa de Informática Médica	APIM	Associação	http://apim.med.up.pt/	Outubro 2012
Medical Informatics - Practical Guide for the Healthcare Professional	MI-PGHP	Livro	[Hoyt, Sutton, Yoshihashi 2009]	Outubro 2012
China-Japan-Korea Joint Symposium on Medical Informatics	CJKMI	Simpósio	http://jcmi2010.e-rad.jp/cjkmi2010/index.html	Outubro 2012
Asia Pacific Association for Medical Informatics Conference	APAMI	Conferência	http://www.apami2012.org/index.php?mod=en	Outubro 2012
South African Health Informatics Association Conference	SAHIAC	Conferência	uth-african-health-informatics-association-conference-2004	Outubro 2012

Decidimos selecionar fontes dos principais centros de desenvolvimento de estudos em Informática Médica do mundo, deste modo a taxonomia construída seria geral e sua abrangência para

a classificação de trabalhos seria maior visto que não se basearia apenas na fonte de um único país. Em seguida foi feita uma análise de cada uma das categorias presentes nas taxonomias selecionadas.

2.1 Taxonomias

A Tabela 2 nos mostra os termos encontrados nas taxonomias selecionadas, bem como suas fontes. É perceptível que determinadas categorizações encontradas são específicas para região em que a associação, congresso ou conferência se destina, como a divulgadas pela *Asia Pacific Association for Medical Informatics Conference* em que, categorias como *Informatics of Traditional Chinese Medicine* só foram identificadas nesta conferência.

Tabela 20. - Taxonomias estudadas

Congressos, Livros e Sociedades	Taxonomias
<i>European Federation for Medical Informatics</i>	- <i>Assistive diagnostic technologies for medical decision support</i> - <i>Knowledge discovery in biomedical databases for decision support</i> - <i>Knowledge management</i> - <i>Formalization of knowledge, ontologies, clinical guidelines and standards of health care</i> - <i>Intelligent interoperability and telemedicine</i> - <i>Data and knowledge management for decision support in forensic medical disciplines</i> - <i>eHealth decision support systems for GPs, clinicians, nurses, health care managers and patients</i> - <i>Education and training for decision support</i> - <i>Evaluation of decision support systems</i> - <i>Diagnostic, therapeutic and prognostic decision support</i>
Sociedade Brasileira de Informática Médica	- Sistemas de Informação em Saúde - Prontuário Eletrônico do Paciente - Telemedicina - Sistemas de Apoio à Decisão - Processamento de sinais biológicos - Processamento de Imagens Médicas - Internet em Saúde - Padronização da Informação em Saúde
American Medical Informatics Association	- <i>Translational Bioinformatics</i> - <i>Clinical Research Informatics</i> - <i>Clinical Informatics</i> - <i>Consumer Health Informatics</i> - <i>Public Health Informatics</i>
Associação Portuguesa de Informática Médica	- <i>Biomedical Electronics and Devices</i> - <i>Bioinformatics Models, Methods and Algorithms</i> - <i>Bio-inspired Systems and Signal Processing</i> - <i>Health Informatics</i>
Medical Informatics Practical Guide for the Healthcare Professional, (Hoyt, 2009)	- <i>Electronic Health Records</i> - <i>Integrated Practice Management Systems</i> - <i>Health Information Technology Interoperability Networks</i> - <i>Patient Informatics</i> - <i>Online Medical Resources</i> - <i>Search Engines</i> - <i>Mobile Technology</i> - <i>Evidence Based Medicine</i> - <i>Clinical Practice Guidelines</i> - <i>Disease Management and Disease Registries</i>

	- Pay for Performance (P4P) - Patient Safety and Technology - Electronic Prescribing - Telehealth and Telemedicine - Picture Archiving and Communication Systems (PACS) - Bioinformatics - Public Health Informatics
China-Japan-Korea Joint Symposium on Medical Informatics	- Medical Informatics - Medical imaging & image management - Tele-medicine and tele-surgery - EMR/EHR - Standard in Health Informatics and cross-language solution - Computer Aided Diagnostic (CAD), Treatment and Surgery - Artificial Intelligence in Medicine and Its Applications - Medical data representation, visualization, and sharing - Health informatics education - Biomechanics, modeling and computing - Public Health informatics - Nursing Informatics - Bioinformatics - Ontology Development - Medical safety management
Asia Pacific Association for Medical Informatics Conference	- e-Health and e-Hospital - Medical Signal and Image Processing - Telemedicine and Telesurgery - Standards in Medical Informatics - Computer Aided Diagnostic (CAD), Treatment and Surgery - Artificial Intelligence in Medicine and its Applications - Medical Data Representation, Mining, Visualization, and Sharing - Robotics and Virtual Reality - Biomechanics, Modeling and Computing - Public Health Informatics and the Health Information Infrastructure - Nursing Informatics - Informatics of Traditional Chinese Medicine - Information Retrieval and Digital Libraries - Clinical Decision Support Systems and Expert Systems - Privacy, Security, and Confidentiality Issues in Medical Information - Training, Education and Cognitive Sciences - Evidence-based Medicine and Guideline - Cloud Computing - Internet of Things - Other related topics
South African Health Informatics Association	- Disease Surveillance and Epidemiology - Health Management Information Systems - Financial management for improved health service delivery - Human Resources Management - Information Technology Infrastructure - Monitoring and Evaluation - Tele-Health / E-Health / Telemedicine/ M-Health

A taxonomia construída neste trabalho foi desenvolvida a partir da comparação entre as diversas categorias estudadas, como veremos na sessão posterior.

3. Análises Comparativas e Desenvolvimento da Taxonomia

É importante salientar a existência de outras taxonomias que definem a Informática Médica, porém desenvolvidas direcionadas para contextos específicos, como as propostas pelo HIMSS (*Healthcare Information Management System Society*) e pela MESH (*Medical Subject Headings*), [Dixon, McGowan, Zafar 2007].

Segundo [Dixon, McGowan, Zafar 2007], a taxonomia proposta pelo HIMSS aborda uma parte das principais subáreas da Informática Médica, porém é estritamente direcionada para sistemas de prontuários médicos eletrônicos, desconsiderando certas aplicações da informática na medicina como telemedicina e o compartilhamento de informações em saúde. Portanto não foi considerada adequada para classificar os trabalhos selecionados para o Mapeamento Sistemático uma vez que a proposta é o desenvolvimento de uma taxonomia geral que represente as principais subáreas existentes para esta grande área da ciência.

[Dixon, McGowan, Zafar 2007] também relata que a taxonomia desenvolvida pelo MESH é derivada dos estudos de John Shaw Billings na publicação *Index Medicus* de 1979, considerado um dos pioneiros a sugerir uma categorização dos termos existentes na literatura médica organizadas hierarquicamente de acordo com uma estrutura de árvore. Posteriormente esta hierarquia de termos foi atualizada e expandida pelo *MEDLINE* incluindo a terminologia “tecnologias da informação em saúde”. Na hierarquia proposta pelo MESH há falta de diversos conceitos da Informática Médica principalmente no que diz respeito à implementação de prontuário eletrônico em saúde.

3.1 Metodologia

Conforme citado na sessão anterior a taxonomia construída foi desenvolvida utilizando como base nos estudos de [Dixon, McGowan, Zafar 2007], para tanto optamos por comparar os termos presentes nas taxonomias estudadas, vide tabela 2. Posteriormente foram selecionados os termos semelhantes ou correlacionados entre as taxonomias estudadas de acordo com sua definição nas fontes citadas na tabela 1, [Feldt et al. 2008], [Richmond 2003]. Uma vez selecionados esses termos, os mesmos foram agregados para a construção de uma única categoria geral. Por exemplo, os seguintes termos: Prontuário Eletrônico do Paciente (SBIS), *Electronic Medical Records/Electronic Health Records* (CJKMI), foram agregados ao termo *Electronic Health Records* [Hoyt, Sutton, Yoshihashi 2009] por considerarmos este como o termo mais geral que poderia abranger qualquer trabalho direcionado ao desenvolvimento de Registros Eletrônicos aplicados a Saúde; quanto ao termo: *Medical Signal and Image Processing* [Proakis, Manolakis 2006] representa a subárea da Engenharia Elétrica voltada para o estudo de técnicas de processamento e análise de sinais e imagens, que nos trabalhos analisados foram aplicadas na medicina [Semmlow 2008], esta categoria também abrange a transmissão, armazenamento e formato de dados desta natureza. Este procedimento para o desenvolvimento dos termos da taxonomia foi baseado no trabalho de [Feldt et al. 2008].

3.2 Resultados da Análise Comparativa

Na tabela 3 podemos ver os termos que originaram as categorias finais para nossa taxonomia.

Tabela 3. - Construção dos termos para a taxonomia

Termos Originais	Termos Finais
Prontuário Eletrônico do Paciente (SBIS) Electronic Health Records (Hoyt, 2009) EMR/EHR (CJKMI)	<i>Electronic Health Records</i>
Telemedicina (SBIS) Telehealth and Telemedicine (Hoyt, 2009) e-Health and e-Hospital (APAMI) Telemedicine and Telesurgery (APAMI) Tele-medicine and tele-surgery (CJKMI) Tele-Health / E-Health / Telemedicine/ M-Health (SAHIAC)	<i>Telemedicine</i>
Assistive diagnostic technologies for medical decision support (EFMI) eHealth decision support systems for GPs, clinicians, nurses, health care managers and patients (EFMI) Diagnostic, therapeutic and prognostic decision support (EFMI) Sistemas de Apoio à Decisão (SBIS) Computer Aided Diagnostic (CAD) and Treatment and Surgery (CJKMI) Computer Aided Diagnostic (CAD) and Treatment and Surgery (APAMI) Clinical Decision Support Systems and Expert Systems (SAHIAC)	<i>Decision Support Systems</i>
Formalization of knowledge, ontologies, clinical guidelines and standards of health care (EFMI) Standard in Health Informatics and cross-language solution (CJKMI) Standards in Medical Informatics (APAMI) Padronização da Informação em Saúde (SBIS)	<i>Standards in Medical Informatics</i>
Public Health Informatics (AMIA) Public Health Informatics (Hoyt, 2009) Public Health informatics (CJKMI) Public Health Informatics and the Health Information Infrastructure (APAMI)	<i>Public Health Informatics</i>
Processamento de sinais biológicos (SBIS) Processamento de Imagens Médicas (SBIS) Bio-inspired Systems and Signal Processing (APIM) Picture Archiving and Communication Systems (Hoyt, 2009) Medical imaging & image management (CJKMI) Medical Signal and Image Processing (APAMI)	<i>Medical Signal and Image Processings</i>
Internet em Saúde (SBIS) Networks (Hoyt, 2009) Online Medical Resources (Hoyt, 2009) Search Engines (Hoyt, 2009) Internet of Things (APAMI) Cloud Computing (APAMI)	<i>Online Medical Resources</i>
Disease Management and Disease Registries (Hoyt, 2009) Disease Surveillance and Epidemiology (SAHIAC)	<i>Disease Management and Disease Registries</i>
Evidence Based Medicine (Hoyt, 2009) Evidence-based Medicine and Guideline (APAMI)	<i>Evidence Based Medicine</i>
Medical data representation, visualization, and sharing (CJKMI) Medical Data Representation, Mining, Visualization, and Sharing (APAMI)	<i>Medical Data Representation, Mining, Visualization, and Sharing</i>
Intelligent interoperability and telemedicine (EFMI) Health Information Technology Interoperability (Hoyt, 2009)	<i>Health Management Information Systems</i>
Education and training for decision support (EFMI) Health informatics education (CJKMI) Training, Education and Cognitive Sciences (APAMI)	<i>Health informatics education</i>
Artificial Intelligence in Medicine and Its Applications (CJKMI) Artificial Intelligence in Medicine and its Applications (APAMI)	<i>Artificial Intelligence in Medicine and Its Applications</i>
Translational Bioinformatics (AMIA) Bioinformatics Models, Methods and Algorithms (APIM) Bioinformatics (Hoyt, 2009) Bioinformatics (CJKMI)	<i>Applications of Bioinformatics in Medical Informatics</i>
Biomedical Electronics and Devices Robotics and Virtual Reality	<i>Biomedical Eletronic and Robótics Devices</i>

Biomechanics, Modeling and Computing	
Clinical Research Informatics (AMIA) Clinical Informatics (AMIA)	Clinical Informatics
Pay for Performance (P4P) (Hoyt, 2009) Financial management for improved health service delivery (APAMI)	Financial management for improved health service delivery
Electronic Prescribing (Hoyt, 2009) Mobile Technology (Hoyt, 2009)	Consumer Health Informatics
Information Technology Infrastructure (SAHIAC) Privacy, Security, and Confidentiality Issues in Medical Information (APAMI) Medical safety management (CJKMI) Human Resources Management (SAHIAC) Integrated Practice Management Systems (Hoyt, 2009)	Infrastructure of Medical Information Systems
Patient Safety and Technology (Hoyt, 2009) Patient Informatics (Hoyt, 2009)	Patient Informatics
Nursing Informatics (CJKMI)	Nursing Informatics
Knowledge discovery in biomedical databases for decision support (EFMI) Knowledge management (EFMI) Data and knowledge management for decision support in forensic medical disciplines (EFMI)	Applications of Knowledge Engineering in Medical Informatics
Informatics of Traditional Chinese Medicine (APAMI, 2012)	Informatics of Traditional Chinese Medicine

É importante esclarecer que termos como *Informatics of Traditional Chinese Medicine* foram transformados em uma categoria exclusiva tendo em vista a especificidade dos estudos desenvolvidos relacionado a este termo encontrado entre os estudos selecionados para compor o Mapeamento Sistemático.

4. Resultados Finais e Discussão

A Tabela 4 mostra os termos definitivos da nova taxonomia para as subáreas da Informática Médica. São 23 termos que consideramos como ideais para classificação dos estudos selecionados para o Mapeamento Sistemático. Acreditamos que esta taxonomia abrange todas as subáreas da Informática Médica, desde as mais gerais como *Telemedicine* até as mais específicas como *Nursing Informatics*.

Tabela 4. - Taxonomia Final

Termos da Taxonomia
<i>Electronic Health Records</i>
<i>Telemedicine</i>
<i>Decision Support Systems</i>
<i>Standards in Medical Informatics</i>
<i>Public Health Informatics</i>
<i>Medical Signal and Image Processings</i>
<i>Online Medical Resources</i>
<i>Disease Management and Disease Registries</i>
<i>Evidence Based Medicine</i>
<i>Medical Data Representation, Mining, Visualization, and Sharing</i>
<i>Health Management Information Systems</i>
<i>Health informatics education</i>

<i>Artificial Intelligence in Medicine and Its Applications</i>
<i>Applications of Bioinformatics in Medical Informatics</i>
Biomedical Electronic and Robotics Devices
Clinical Informatics
Financial management for improved health service delivery
Consumer Health Informatics
Infrastructure of Medical Information Systems
Patient Informatics
Nursing Informatics
Applications of Knowledge Engineering in Medical Informatics
Informatics of Traditional Chinese Medicine

Dentre as principais prioridades consideradas durante a construção da taxonomia podemos citar sua generalidade e simplicidade dos termos, esta pode ser verificada por serem termos que consideramos intuitivos e consistentes que permitem uma classificação de estudos eficiente e simples. Quanto a definição formal destes termos, podem ser verificados em suas respectivas fontes. No mais, é importante deixar claro que todo o processo de desenvolvimento de um Mapeamento Sistemático evolui e se adapta conforme as necessidades apresentadas durante o estudo, deste modo esta categoria ainda pode ser modificada se identificada esta necessidade.

5. Perspectivas Futuras

O principal objetivo deste estudo foi o desenvolvimento de uma taxonomia que caracterize todas as subáreas da Informática Médica, para a classificação dos estudos selecionados para o Mapeamento Sistemático sobre o uso de Ontologias em Informática Médica que se encontra em desenvolvimento.

Pretende-se publicar este estudo também após a conclusão do Mapeamento Sistemático, visto que sua eficácia será validada para este estudo, além de permitir sua avaliação através de sua utilização prática diante da comunidade científica relacionada, permitindo deste modo que outros tipos de estudos sejam desenvolvidos utilizando esta taxonomia como uma das suas referências.

Outra possibilidade seria a utilização deste estudo como uma das fontes necessárias para a construção de uma ontologia que defina as principais subáreas da Informática Médica permitindo assim o compartilhamento de conhecimento mais eficiente.

6. References

- Dixon, Brian E.; McGowan, Julie J.; Zafar, Atif. **Development of a Taxonomy for Health Information Technology**. World Congress on Medical and Health Informatics. Australia 2007.
- Feldt, R.; Petersen, K.; Mattsson, M.; Mujtaba, S. **Systematic Mapping Studies in Software Engineering**. 12th International Conference on Evaluation and Assessment in Software Engineering, pp. 71-80, Bari, Italy, June 26-27, 2008.

- Hoyt, R. E.; Sutton, M.; Yoshihashi, A. **Medical Informatics - Practical Guide for the Healthcare Professional**. 3ª Edição, University of West Florida, School of Allied Health and Life Sciences. Lulu.com. 2009.
- Kitchenham, B. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. Vol 2.3 EBSB Technical Report, EBSE-2007-01, 2007.
- Arksey, H; O'Malley, L. **Scoping Studies: towards a methodological framework**. International Journal of Social Research Methodology, vol. 8, número 1, 2005. p. 19-32.
- Richmond, Bruno D. H.; **The Truth About Taxonomies**. Info Mgmt. Mar/Apr. 2003. Pages 44-53.
- Travassos, G.; Biolchini J. **Revisões Sistemáticas Aplicadas a Engenharia de Software**. In: XXI SBES – Brazilian Symposium on Software Engineering, João Pessoa, PB, Brasil, 2007.
- Wieringa, R.; Maiden, N. A. M.; Mead, N. R; Rolland, C. (2006). **Requirements engineering paper classification and evaluation criteria: a proposal and a discussion**. *Requir. Eng.* **11**(1), 102–107.
- Almeida, A., T. **Um Mapeamento Sistemático de Mecanismos para Guiar Estudos Empíricos em Engenharia de Software** Dissertação (Mestrado em Ciências da Computação). Universidade Federal de Pernambuco. Recife, Pernambuco. Março, 2011.
- Cardozo, E., S., F. **Mapeamento Sistemático sobre o uso do Auto-Gerenciamento em Equipes de Desenvolvimento de Software**. Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco. 2012.
- Jacinto, S., S. **Um Mapeamento Sistemático de Pesquisa sobre a Influência da Personalidade na Engenharia de Software**. Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco. Setembro, 2010.
- Proakis, J. G.; Manolakis, D. K. **Digital Signal Processing**. 4th Edition; Prentice-Hall, 2006.
- Semmlow, J. L. **Biosignal and Medical Image Processing**. 2nd Edition. CRC Press, 2008.
- Suassuna, M. **Mapeamento Sistemático sobre Replicação de Estudos Empíricos em Engenharia de Software**. Dissertação (Mestrado em Ciências da Computação). Centro de Informática. Universidade Federal de Pernambuco. Recife, Pernambuco. Junho, 2011.
- Dang, J; Hedayati, A; Hampel, K; Toklu, C. **An ontological knowledge framework for adaptive medical workflow**. *Journal of Biomedical Informatics*, volume 41, Junho, 2008. P. 829–836.
- Silachan, K.; Tantatsanawong, P. **Domain Ontology Health Informatics Service From Text Medical Data Classification**. Department of Computing, Faculty of Science, Silpakorn University. Nakornpratom, Thailand. Annual SRII Global Conference. 2011.
- Russell, S.; Norvig, P. **Artificial Intelligence: A Modern Approach**. 3rd Edition. New Jersey: Prentice Hall; 2009.
- Sarkar, I. N. **Biomedical Informatics And Translational Medicine**. *Journal Of Translational Medicine*, 2010. P. 1 – 12.