

**CRIAÇÃO E IMPLANTAÇÃO DE UM SISTEMA DE  
INDEXAÇÃO DE LAUDOS RADIOLÓGICOS EM UM  
CENTRO DE REFERÊNCIA NO ENSINO E PESQUISA  
EM ONCOLOGIA**

**LETÍCIA SILVA SENE**

**Dissertação apresentada à Fundação Antônio  
Prudente para obtenção do título de Mestre  
em Ciências**

**Área de Concentração: Oncologia**

**Orientador: Dr. Rubens Chojniak**

**Co-Orientadora: Dra. Marcela Pecora Cohen**

**São Paulo**

**2014**

## FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Sene, Letícia Silva

**Criação e implantação de um sistema de indexação de laudos radiológicos em um centro de referência no ensino e pesquisa em oncologia** / Letícia Silva Sene - São Paulo, 2014.

46p.

Dissertação (Mestrado)-Fundação Antônio Prudente.

Curso de Pós-Graduação em Ciências - Área de concentração:  
Oncologia.

Orientador: Dr Rubens Chojniak

Descritores: 1. INDEXAÇÃO COMO ASSUNTO. 2. RECUPERAÇÃO DA INFORMAÇÃO. 3. RELATÓRIO. 4. SISTEMAS DE INFORMAÇÃO EM RADIOLOGIA

## DEDICATÓRIA

Dedico este trabalho aos meus pais Paulo e Sandra, que foram fundamentais na minha educação.

Ao meu irmão Túlio pelo incentivo na realização do mestrado.

E ao André pelo apoio durante mais esta etapa da minha vida.

## **AGRADECIMENTOS**

Agradeço ao meu orientador Dr. Rubens Chojniak que possibilitou a realização deste trabalho.

À minha co-orientadora Dra. Marcela Cohen pela grande contribuição nas sugestões e orientações desta dissertação.

Aos colegas radiologistas do A.C. Camargo Cancer Center, em especial a Paula Barbosa e Simone Liao, pelo tempo dedicado na alimentação do banco de dados e na avaliação da ferramenta.

Ao Departamento de Tecnologia da Informação do A.C. Camargo Cancer Center, em especial ao Bruno Muraca, pelo auxílio em diversos momentos da realização deste trabalho.

## RESUMO

Sene LS. **Criação e implantação de um sistema de indexação de laudos radiológicos em um centro de referência no ensino e pesquisa em oncologia.** São Paulo; 2014. [Dissertação de Mestrado-Fundação Antônio Prudente].

Por meio dos relatórios radiológicos produzidos no Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center é gerada uma quantidade significativa de dados em formato de texto livre, que habitualmente ficam apenas armazenados em arquivos pessoais, o que torna o armazenamento, o gerenciamento e a recuperação destas informações desafios complexos. O objetivo principal deste trabalho foi a criação de um modelo de indexação dos relatórios radiológicos no Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center, que é um centro de referência no ensino e pesquisa em oncologia. Desenvolveu-se um modelo de ferramenta de indexação de forma manual com termos do léxico CID-10. O modelo desenvolvido tornou possível a indexação dos relatórios radiológicos e a recuperação fidedigna dos relatórios armazenados a partir do termo inserido. Além de reafirmar a importância do tema o trabalho concluiu que é viável a criação de um banco de dados vinculado ao sistema de informação radiológica em uso, sem a necessidade de aquisição de novas tecnologias. Todavia, embora o modelo tenha sido bem sucedido em suas etapas iniciais algumas dificuldades foram percebidas durante a fase de implantação. A principal complicação esteve relacionada com a dificuldade dos usuários em assimilar o processo de indexação como parte da rotina de trabalho. O tempo despendido com esse novo processo foi outro fator que atuou como barreira para a implementação da ferramenta. Levando à conclusão de que melhorias que simplifiquem o processo ainda precisam ser alcançadas para que o a indexação dos relatórios radiológicos se torne parte da rotina dos médicos radiologistas do Departamento de Diagnóstico por Imagem.

## SUMMARY

Sene LS. [**Creation and implementation of a system for indexing radiology reports in a referral center in teaching and research in oncology**]. São Paulo; 2014. [Dissertação de Mestrado-Fundação Antônio Prudente].

Through radiology reports produced in A.C. Camargo Cancer Center image department, a significant amount of free text data is daily generated, which usually are stored in personal files, so that the storage, management and retrieval of such information become a complex challenges. The aim of this research was to create an indexation model that could be used to input data into an integrated system of radiology reports at the A.C. Camargo Cancer Center Image Department, which is a oncology center reference in teaching and research. A manual operated indexing model was developed using lexicon ICD-10 terms, and so It was possible to index the radiology reports and provide searchable results of all data information stored. Besides reaffirming the topic importance the work attested the possibility of creating a radiology database linked with the current radiological information system at image department without acquire new technologies. Although the model has reached good results at the early stages, complications were observed during the implementation phase. The main complication was related with users assimilating the new methodology on their work. The time spent in this process was another barrier factor for the toll implementation. Appointing that simplify the process is still necessary for the radiology reports indexing become a routine for radiologists imaging department.

## LISTA DE FIGURAS

<b>Figura 1</b>	Modelo de criação de uma ontologia usando Protege.....	17
<b>Figura 2</b>	Tela do sistema onde será incluído o ícone de busca.....	22
<b>Figura 3</b>	Modelo da Ferramenta de Busca.....	22
<b>Figura 4</b>	Ícone Múltiplos Diagnósticos.....	30
<b>Figura 5</b>	Léxico Criado.....	30
<b>Figura 6</b>	Interface para Observações.....	31
<b>Figura 7</b>	Tela para a Simulação de Busca de Diagnóstico.....	33
<b>Figura 8</b>	Filtro de Busca Manipulável.....	34
<b>Figura 9</b>	Lista com os Resultados da Pesquisa.....	35
<b>Figura 10</b>	Modelo de Pergunta.....	35

## LISTA DE SIGLAS E ABREVIATURAS

<b>ACC</b>	American College of Cardiology
<b>ACR</b>	American College of Radiology
<b>BI-RADS</b>	Breast Imaging Report and Data System
<b>CAP</b>	College of American Pathologists
<b>CID</b>	Classificação Internacional de Doenças
<b>CID-O</b>	Classificação Internacional de Doenças em Oncologia
<b>DICOM</b>	Digital Imaging and Communications in Medicine
<b>HIS</b>	Hospital Information System
<b>ISO</b>	International Organization for Standardization
<b>OMS</b>	Organização Mundial de Saúde
<b>PACS</b>	Picture Archiving and Communication System
<b>PI-RADS</b>	Structured Reporting System for Magnetic Resonance of the Prostate
<b>RIS</b>	Radiology Information System
<b>RSNA</b>	Radiological Society of North America
<b>SNOMED-CT</b>	Systematized Nomenclature of Medicine—Clinical Terms, from the College of American Pathologists
<b>TI</b>	Tecnologia da Informação
<b>TI-RADS</b>	Thyroid Imaging Reporting and Data System

# ÍNDICE

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>1</b>
1.1	Objetivo .....	3
1.1	Delineamento da Dissertação.....	3
<b>2</b>	<b>REVISÃO TEÓRICA</b> .....	<b>5</b>
2.1	Sistemas de Gerenciamento de Informações.....	5
2.2	Armazenamento e Recuperação das Informações.....	7
2.3	Léxicos Médicos .....	12
2.3.1	Classificação Internacional de Doenças.....	13
2.3.2	Snomed-CT .....	14
2.3.3	ACR Index .....	15
2.3.4	RadLex .....	16
2.4	Ontologias .....	17
<b>3</b>	<b>APLICAÇÃO PRÁTICA DO MODELO DE FERRAMENTA DE INDEXAÇÃO</b> .....	<b>19</b>
3.1	Análise do Sistema Atual de Laudos e Desenho da Ferramenta .....	20
3.2	Modelos de Desenvolvimento de Software.....	23
3.3	Estudo da Viabilidade de Inclusão e Integração.....	24
3.4	Verificação, Validação e Implantação.....	25
<b>4</b>	<b>RESULTADOS</b> .....	<b>28</b>
<b>5</b>	<b>DISCUSSÃO</b> .....	<b>36</b>
<b>6</b>	<b>CONCLUSÕES</b> .....	<b>40</b>
<b>7</b>	<b>REFERÊNCIAS</b> .....	<b>42</b>

## ANEXOS

**Anexo 1** Entrevista de Avaliação da Ferramenta Pelos Usuários

**Anexo 2** Carta de Aprovação do Comitê de Ética em Pesquisa

## 1 INTRODUÇÃO

O relatório radiológico é o registro oficial de um exame diagnóstico, um procedimento intervencionista ou terapêutico. O relatório é direcionado ao médico solicitante e é parte oficial dos prontuários médicos. Todo o relatório inclui informações demográficas básicas, o nome e a idade do paciente, bem como o número de seu registro. O relatório consiste na descrição dos achados de imagem de acordo com a interpretação do médico radiologista que, geralmente, é expressa em formato de texto de livre. Inclui, ainda, a comparação com os exames anteriores quando disponíveis.

A maioria das instituições mundiais, incluindo o Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center, não recorre a laudos estruturados ou semiestruturados, mas, sim, documentos de texto livre para representar essa descrição (REINER 2013). Por meio dos relatórios radiológicos é gerada diariamente uma quantidade significativa de dados em formato de texto livre, que, habitualmente, ficam apenas armazenados em bases de dados ou em arquivos pessoais, quando o médico emissor julga importante arquivá-las.

Considerando o volume de informação que é gerado diariamente no setor de diagnóstico por imagem por meio dos relatórios médicos radiológicos, o armazenamento, o gerenciamento e a recuperação destas informações tornam-se desafios complexos, já que estas informações

encontram-se dispersas em formato de texto livre, sem uma adequada formalização e estruturação, o que dificulta a sua extração e uso.

Os relatórios radiológicos patológicos e as suas imagens são fonte de enriquecimento do ensino para os médicos em formação, assim a disponibilização destes casos para estudo e discussão entre a equipe médica agrega conhecimento na curva de aprendizado. Os casos patológicos podem contribuir também para a pesquisa clínica com o objetivo de estabelecer novas diretrizes e aprimorar a prática médica e o atendimento aos futuros pacientes.

O A.C. Camargo Cancer Center (2012) é pioneiro no tratamento, ensino e pesquisa do câncer, sendo, atualmente, um dos maiores centros oncológicos mundiais, atendendo a cerca de 15 mil novos pacientes ao ano. A instituição tem à disposição o que há de mais avançado em diagnóstico por imagem e realizou, neste ano, mais de 225 mil exames de diagnóstico por imagem, incluindo os exames de raio x, ultrassonografia, tomografia computadorizada e ressonância magnética. Associado à assistência médica o A.C. Camargo Cancer Center é também o maior produtor de conhecimento científico em oncologia do país. No ano de 2012, seus pesquisadores publicaram 181 trabalhos científicos em revistas indexadas, o que representa 74% de toda a produção científica de alto impacto relacionada à oncologia.

Poucas áreas da medicina melhoraram tanto quanto a de diagnóstico por imagem, em virtude da evolução tecnológica dos equipamentos. O Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center

acompanhou esta evolução e, hoje, é considerado um centro de ensino. O departamento atrai médicos de todo o país e também do exterior, que vêm até a instituição para aprender e extrair novos conhecimentos em relação ao diagnóstico e ao tratamento do câncer. Além disso, os profissionais dessa equipe estão em constante atualização nos principais centros oncológicos do mundo, a fim de introduzir novas técnicas e tecnologias para o Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center.

## **1.1 OBJETIVO**

O objetivo deste trabalho foi a criação de um modelo de indexação dos relatórios radiológicos no Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center, que é um centro de referência no ensino e pesquisa em oncologia.

A partir da hipótese de que é possível criar um banco de dados no sistema de informação em radiologia (RIS) em uso atualmente no Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center. A pergunta a ser respondida é se o banco de dados criado será eficiente na recuperação rápida e fidedigna dos casos armazenados.

## **1.2 DELINEAMENTO DA DISSERTAÇÃO**

A delimitação da pesquisa, em termos espaciais, portanto, foi o Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center.

Em relação ao critério temporal, a revisão teórica e o trabalho empírico foram realizadas de setembro de 2012 a agosto de 2013. Quanto ao objeto da pesquisa, procurou-se avaliar especificamente o processo de indexação de relatórios radiológicos de tomografia computadorizada e ressonância magnética e a sua contribuição na organização de um banco de dados.

Além desta introdução o trabalho foi organizado em mais três capítulos. No segundo capítulo é apresentada uma contextualização do local de realização da pesquisa, o surgimento dos sistemas de gerenciamento das informações como ferramenta de uso rotineiro nos serviços de imagem, incluindo o Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center e o programa em uso neste serviço. Nesse capítulo também é realizada uma revisão da literatura que trata do armazenamento de informações e as iniciativas de padronização de léxicos médicos, destacando-se os principais léxicos disponíveis na atualidade.

No capítulo seguinte é detalhada a escolha da metodologia utilizada no trabalho empírico, além das etapas de criação de uma ferramenta de *software* e o modelo desenvolvimento estabelecido proposto. Por fim, são apresentados os resultados, as contribuições e limitações do trabalho, ressaltando os objetivos alcançados e as contribuições científicas, além das considerações finais.

## **2 REVISÃO TEÓRICA**

Neste capítulo será feita uma breve revisão da literatura sobre os sistemas de gerenciamento de informações e o processo de armazenamento e recuperação de informações. Serão revistos também alguns léxicos médicos e a Classificação Internacional de Doenças que foi a terminologia empregada para a indexação dos relatórios nesta pesquisa.

### **2.1 SISTEMAS DE GERENCIAMENTO DE INFORMAÇÕES**

A rotina de trabalho dos profissionais na área da saúde tem passado, nos últimos anos, por uma transformação, já que os documentos físicos têm sido substituídos por arquivos eletrônicos. Essa transformação modificou a forma de trabalho e de armazenamento das informações nos serviços de radiologia e diagnóstico por imagem. Os registros que antes eram realizados em papel, filmes impressos e outros meios físicos, hoje é armazenado em arquivos digitais.

Para auxiliar nesse trabalho, surgiram os Sistemas de Informações de Radiologia - *Radiology Information Systems* (RIS), que são programas de computadores criados para dar suporte ao fluxo de trabalho no Departamento de Diagnóstico por Imagem de um hospital ou a uma clínica que realiza exames de imagem. Estes sistemas possuem diversos tipos de ferramentas para o registro eletrônico de pacientes, a geração de listas de

trabalho, digitação de relatórios médicos, a integração com as imagens dos exames e a transmissão direta de dados para o faturamento.

Os RIS são integrados aos Sistemas de Informação Hospitalar - *Hospital Information Systems* (HIS), que por sua vez são programas de computadores desenvolvidos para propiciar o gerenciamento de pacientes e médicos, de prontuários eletrônicos, faturamento e leitos de um hospital. As informações demográficas dos pacientes que são utilizadas no RIS podem ser importadas do HIS.

A integração das imagens radiológicas no RIS é possível através dos sistemas de *Picture Archiving and Communication System* (PACS), estes possibilitam a armazenagem e recuperação das imagens em uma rede de computadores, além do pós-processamento e visualização das mesmas. As imagens são armazenadas em formato *Digital Imaging and Communications in Medicine* (DICOM), que consiste conjunto de dados compilado em um único arquivo. Este padrão de formatação possibilita a transmissão das informações veiculadas às imagens médicas entre os equipamentos geradores das imagens, as estações de trabalho e os computadores. As imagens em formato DICOM não podem ser visualizadas diretamente com os sistemas operacionais de computadores pessoais, é necessário um visualizador, o PACS. Estes visualizadores possuem ferramentas para exportar as imagens armazenadas em DICOM num formato compatível com estes sistemas operacionais, tipo JPEG, para que as imagens possam ser utilizadas sem o visualizador específico (VARMA 2012).

Os sistemas PACS associam as imagens de exames às informações de pacientes por meio de chaves de consultas textuais e numéricas, não suportando consultas baseadas no conteúdo pictórico das imagens. Essa integração possibilita a criação de sistemas de gerenciamento de imagens e redes locais de larga escala, proporcionando que as informações dos pacientes e as imagens sejam compartilhadas e visualizadas local e remotamente (SEELEY et al. 1987; BOEHME e CHOPLIN 1992).

A combinação de diversos sistemas de registros eletrônicos visa, também, à maior agilidade aos dados e informações do paciente, porém estes sistemas têm enfoques mais administrativos e financeiros do que acadêmicos e científicos (AZEVEDO-MARQUES et al. 2005). O RIS Digital Clinic (versão 1.0.2.299), da empresa Microdata, em uso, atualmente, no A.C. Camargo Cancer Center viabiliza a integração de imagens com dados demográficos dos pacientes e seus respectivos exames, disponibilizando uma gama de recursos para automatizar o processo de geração, arquivamento e recuperação dos relatórios radiológicos. O Digital Clinic gera relatórios de exames, listas de trabalhos, busca de exames por data, médico solicitante e médico emissor do relatório. Mas, não dispõe de uma ferramenta para a busca dos relatórios vinculada ao diagnóstico.

## **2.2 ARMAZENAMENTO E RECUPERAÇÃO DE INFORMAÇÕES**

Uma das grandes vantagens do armazenamento digital é a possibilidade de se realizarem buscas e consultas às informações. A

recuperação de informações é o ramo da ciência cujo objetivo é a procura de elementos em algum meio em resposta a uma consulta de um usuário (STEVESON e WILKS 2003). Esta área de cresceu em importância, nos últimos anos, com o aumento da disponibilidade de informação em formato digital (GARFIELD 2001).

O processo de recuperação da informação consiste em identificar, em um conjunto de documentos de um sistema, quais atendem às necessidades do usuário. A busca pode ser realizada mediante uma pesquisa destas informações em texto livre, porém, com o aumento da quantidade de documentos, mesmo os computadores mais modernos não são capazes de recuperar a grande quantidade de informação armazenada, a busca deve limitar-se a um número restrito de palavras chave (DUQUE 2006). Outra dificuldade do usuário de busca está em predizer, por meio de uma expressão de busca, as palavras ou expressões que foram usadas para representar os documentos que satisfarão a sua necessidade, já que existe uma grande variabilidade em relação à terminologia utilizada nos relatórios radiológicos. Esta variabilidade pode ser exemplificada com diferentes termos que o radiologista pode designar para o mesmo achado: “litíase vesicular”, “cálculo na vesícula” ou “colecistopatia calculosa”.

A indexação visa dar destaque às informações necessárias ao usuário, objetivando o que será armazenado e, posteriormente, recuperado pelo mesmo (RODRIGUES e CRIPPA 2011). Durante o processo de indexação, são extraídos termos relevantes do documento que será armazenado, por meio da análise de seu conteúdo, e estes termos são

alterados para o correspondente lexical, que será a chave de acesso para a busca. Assim, o usuário precisa ter algum grau de conhecimento sobre o texto que está sendo analisado a fim de poder extrair as palavras chave e escolher um termo substituto dentro de um léxico pré-definido.

Os textos ou os documentos dos quais serão extraídas as informações de interesse podem ser encontrados em três formatos, dependendo do nível de estruturação. O texto é considerado estruturado, quando apresenta regularidade no formato de apresentação das informações, que permitem uma extração informatizada da informação, como podemos encontrar nos formulários preenchidos. Os textos semi-estruturados são aqueles que evidenciam alguma regularidade na disposição dos dados, pois algumas partes do texto apresentam uma formatação, enquanto outras partes aparecem de forma irregular, como exemplo podemos citar a primeira página de um artigo científico. Os textos não estruturados, também conhecidos como textos livres, não devotam qualquer regularidade na exposição dos dados. Nestes casos, as informações não são facilmente detectadas (a não ser que se tenha um conhecimento linguístico) e podem ser citados os relatórios médicos radiológicos. Assim sendo, o tipo de texto do qual é feita a extração dos dados tem grande influência na técnica de extração a ser utilizada, se manual ou automatizada. Técnicas de processamento da linguagem natural têm sido desenvolvidas com o objetivo de extrair informações relevantes de textos livres (GARFIELD 2001; DO et al. 2010). O processamento da linguagem natural trata-se de um conjunto de técnicas computacionais,

baseadas em algoritmos que permitem a compreensão e o processamento da linguagem humana por computadores. Esta forma de indexação apresenta resultados promissores, porém trata-se de uma abordagem complexa dos fundamentos teóricos de linguagem computacional (SPYNS 1996).

A recuperação da informação a partir do conteúdo descritivo não é facilmente acessível pelos programas de computador, configurando uma barreira entre as imagens e os *RIS* (MONGKOLWAT et al. 2012). O *RIS* possibilita a busca dos relatórios mediante dados demográficos básicos associados a estes: a data do exame, o médico solicitante ou o médico executor (RAMASWAMY et al. 1996).

Na tentativa de unificação da linguagem e da extração de informações chave do texto, além de um maior controle da qualidade do atendimento nos serviços de radiologia, os participantes da Conferência Intersociedade de 2007, promovida pelo *American College of Radiology* (ACR), recomendaram o uso de relatórios estruturados. Porém poucos modelos de relatórios estruturados estão disponíveis, com exceção dos exames de mama e os cardiológicos, para os quais as subespecialidades já criaram relatórios estruturados. Em 2008, o *American College of Cardiology* (ACC), juntamente com o ACR, RSNA e outras organizações se reuniram para a criação e o endosso de um documento com o objetivo de apoiar o uso de relatórios estruturados nos exames de imagem em cardiologia, com o objetivo central de melhoria da prática clínica, integração das informações e aumento da qualidade dos serviços prestados. As etapas preliminares para

a implementação dos relatórios estruturados já foram realizadas, como a padronização dos dados referentes às subespecialidades e a utilização de programas de computador para a geração dos relatórios (DOUGLAS et al. 2009). Os relatórios de exames de mama já utilizam o léxico do *Breast Imaging Report and Data System* (BI-RADS). O BI-RADS foi o primeiro sistema desenvolvido para a prática do diagnóstico por imagem, criado com a intenção de resolver a falta de padronização e uniformidade nos relatórios de mamografia. Apresenta três componentes importantes, um léxico de descritores, uma estrutura de relatórios recomendados com categorias de avaliação final e recomendações de gestão de acompanhamento, além de uma estrutura para a coleta de dados e auditoria. O léxico BI-RADS pode servir como exemplo de uma terminologia padrão altamente bem sucedido para a criação de léxicos em outras subespecialidades da radiologia (DUNNICK e LANGLOTZ 2007; HALL 2009; BURNSIDE et al. 2009). Outras tentativas de uniformização da linguagem e padronização dos laudos estruturados, mas que, na prática ainda, não são empregadas, incluem o *Structured Reporting System for Magnetic Resonance of the Prostate* (PI-RADS) e o *Thyroid Imaging Reporting and Data System* (TI-RADS) para exames de ressonância magnética da próstata e ultrassonografia da tireóide (ROTHKE et al. 2013; RUSS et al. 2013).

Apesar de os estudos citados demonstrarem que a aplicação de tais sistemas na prática médica possibilita mudanças positivas na qualidade da assistência médica, contudo, em certas circunstâncias, a aplicação de um modelo fechado de conhecimento pode causar dificuldade de uso para o

especialista e mesmo problemas de programação de soluções tecnológicas para o desenvolvedor. A dificuldade da transposição de uma terminologia fechada para a prática decorre, fundamentalmente, do distanciamento entre a perspectiva teórica do uso da terminologia desenvolvida (como usar, para que usar e quem deve ou pode usar) e as condições reais de uso local onde será aplicada (o que envolve a cultura institucional e os formalismos próprios de cada indivíduo médico). Dessa forma, os benefícios almejados de um sistema terminológico acabam não sendo alcançados totalmente, já que existe uma lacuna entre o modelo proposto e a prática da medicina cotidiana (RECTOR 1999; SERAPIÃO et al. 2013).

### **2.3 LÉXICOS MÉDICOS**

Uma classificação é um sistema ordenado de conceitos dentro de um domínio, com princípio de sistematização e agrupamentos implícitos ou explícitos. Os sistemas de classificação, em medicina, possuem finalidades estatísticas e epidemiológicas, pesquisa clínica, indexação de documentos e prestação de contas. Esses sistemas de codificação ajudam a lidar com a enorme variabilidade de expressões e termos médicos reduzindo a ambiguidade e relacionando os termos sinônimos. Com a utilização desses sistemas, os conceitos podem ser representados por códigos, que são independentes da linguagem natural. Porém esses léxicos de indexação têm diversas desvantagens, geralmente relacionadas a falta de expressividade dos termos, pois os conceitos semelhantes são agrupados em um termo

único (HOLANDA et al. 2004). Outra desvantagem é que esses sistemas lexicais e vocabulários modelados para a comunicação, em grande parte, se encontram apenas disponíveis em língua inglesa (SERAPIÃO et al. 2013).

### **2.3.1 Classificação Internacional de Doenças (CID)**

A CID trata-se de um léxico médico criado, em 1893, inicialmente, para a classificação de causas de mortalidade, seu âmbito foi, progressivamente, ampliado para incluir uma codificação de causas de morbidade e a indexação de prontuário médico. Esta ferramenta estatística é essencial para a comparabilidade nacional e internacional na área de saúde pública. Ela é usada para monitorar a incidência e prevalência de doenças e outros problemas de saúde, bem como para classificar as doenças e outros problemas de saúde registrados, incluindo certidões de óbito e registros de saúde. Além de permitir o armazenamento e recuperação de informações de diagnóstico para fins clínicos, epidemiológicos e de qualidade, esses registros também fornecem a base para a compilação de mortalidade nacional e estatísticas de morbidade por países membros da Organização Mundial de Saúde (OMS). A capacidade de trocar dados comparáveis entre países e regiões para a comparação de uma população com outra e possibilitar o estudo de doenças entre longos períodos é um dos pontos fortes da CID. A OMS tem sido responsável pela organização, coordenação e execução das atividades relacionadas com a CID desde 1948 (ano da sexta revisão da CID), a revisão é realizada a cada 10 anos, no propósito de acompanhar os avanços da medicina.

A décima revisão da CID, CID-10, é a versão atualmente em uso e foi aprovada pelo Quadragésima Terceira Assembleia Mundial da Saúde, em maio de 1990, e entrou em uso em países membros da OMS a partir de 1994. A CID-10 recorre a um esquema de codificação alfanumérica que consiste em uma letra seguida de três números que compõem quatro caracteres, proporcionando a classificação das doenças por especialidades médicas em subcategorias. A 11ª revisão da CID já começou e deverá ser concluída até 2015 (BRÄMER 1988; World Health Organization-WHO 2010).

O capítulo II da CID 10 sempre foi reservado para a oncologia, contendo a topografia da neoplasia e o seu comportamento, se benigno ou maligno. A partir de uma necessidade dos patologistas e oncologistas de subcategorizar as neoplasias de acordo com a morfologia, em 1976, a OMS criou a classificação internacional de doenças em oncologia (CID-O), que é empregada mundialmente e já foi traduzida para 11 línguas (OMS 2000).

### **2.3.2 SNOMED-CT**

Formado em 1993, o *Systematized Nomenclature of Medicine–Clinical Terms, from the College of American Pathologists, Northfield, IL-SNOMED-CT* (2006) é uma das mais completas nomenclaturas criadas para indexar o conjunto de registros médicos, atualmente, totalizando, aproximadamente, 984.000 termos, organizados segundo tipos semânticos e hierárquicos. É um sistema de classificação abrangente, de nomenclatura multiaxial, criado para a indexação do registro clínico completo, incluindo sinais e sintomas, diagnósticos e procedimentos, e que pretende realizar a integração de toda

a informação do registro médico eletrônico numa única estrutura de dados. Pode ser usada para codificar, recuperar e analisar os dados clínicos. O SNOMED CT resultou da fusão de SNOMED, terminologia de referência (SNOMED RT), desenvolvido pelo Colégio Americano de Patologistas (CAP), e a versão termos clínicos 3 (CTV3), desenvolvido pelo Serviço Nacional de Saúde do Reino Unido. São 19 os eixos hierárquicos com várias subclassificações, assim, um conceito é classificado de acordo com a classe semântica a que pertence. Esta classificação está dividida em conceitos, descrições, relacionamentos, hierarquias e sub-hierarquias, que resultam em, aproximadamente 1,4 milhões de relacionamentos entre termos.

### **2.3.3 ACR Index**

Uma das primeiras tentativas de se criar uma linguagem universal para indexação de laudos foi o sistema *ACR Index*, criado pelo Colégio Americano de Radiologia. A primeira edição foi publicada em 1955 e a terceira edição em 1986. As edições foram impressas como livros de bolso para consulta manual. Este sistema foi desenvolvido, originalmente, para organizar os dados de casos baseados em imagens interessantes coletados pelos radiologistas, trabalho que era, até então, realizado de forma manuscrita. De acordo com essa metodologia, era feita uma identificação anatômica e patológica com códigos em números decimais, sendo que o código numérico da localização anatômica aparece antes e a entidade patológica depois. Por conter apenas poucos termos, o *ACR Index* facilitava

seu uso pelos radiologistas quando os computadores ainda não eram utilizados. Atualmente, como os computadores têm capacidade para gerenciar listas enormes de termos e suas inter-relações, novos métodos de indexação foram criados (YAM et al. 2004).

#### **2.3.4 RadLex**

No começo de 2005, a *Radiological Society of North America* (RSNA) convocou um comitê de especialistas em informática e em subespecialidades radiológicas, para criar uma base de dados que unificasse os termos radiológicos. A criação de um léxico uniforme, adequado e completo é a proposta do *Radlex*. Para evitar duplicação de esforços, a Sociedade de Radiologia Americana, com apoio do Colégio Americano de Patologia, aplicou os termos do SNOMED-CT para iniciar sua indexação. Houve apoio, também, do Colégio Americano de Radiologia, permitindo, assim, que os materiais com base em imagens do *ACR Index* pudessem ser facilmente indexados com os códigos do *RadLex*. O *RadLex* trata-se de uma unificação de termos para suprir a necessidade de organizar a linguagem e possibilitar a recuperação dos dados (LANGLOTZ 2006; WEISS e LANGLOTZ 2008) A sua estrutura está organizada em uma hierarquia de categorias e subcategorias, contendo cerca de 12000 termos.

## 2.4 ONTOLOGIAS

Trata-se de uma técnica de organização de informações para representação formal do conhecimento, com objetivo de compartilhamento e entendimento comum de algum domínio de conhecimento que possa ser comunicado entre pessoas e computadores. Na prática a ontologia define um conjunto de termos que será utilizado para formular consultas (ALMEIDA e BAX 2003).

A ontologia representa, formalmente, o domínio do conhecimento, classificando as informações relevantes do texto, e podem funcionar sobre fontes de dados, proporcionando uma melhor organização e recuperação eficiente das informações (ALMEIDA 2003).

O *Protégé-2000* é uma ferramenta de uso livre que tem a função de um editor de ontologias, isto é, um programa com interface gráfica que permite realizar a construção de uma ontologia pelo usuário (RUBIN et al. 2007).

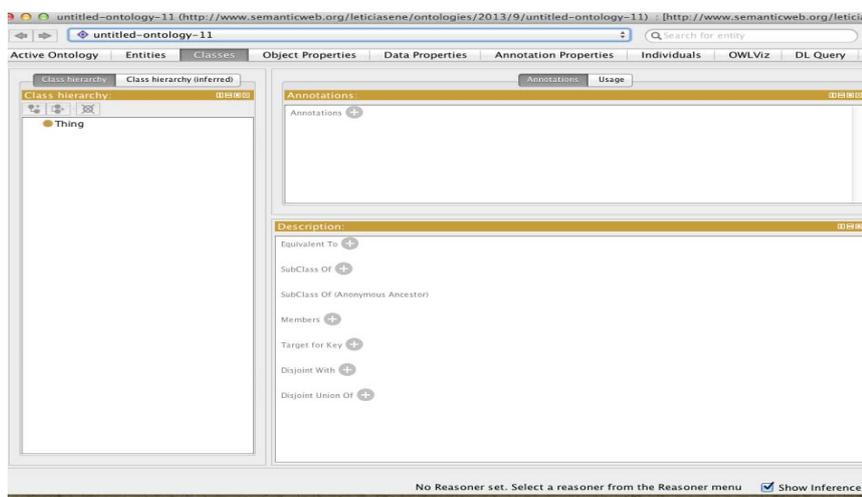


Figura 1 - Modelo de criação de uma ontologia usando *Protege*

Ferramentas de construção de ontologias são essenciais para este processo, já que a construção manual é muito trabalhosa, demanda tempo e pode ser passível de erros de interpretação e de percepção do desenvolvedor. A criação de uma ontologia está relacionada aos métodos de processamento da linguagem natural e a extração automatizada de informações.

### **3 APLICAÇÃO PRÁTICA DO MODELO DE FERRAMENTA DE INDEXAÇÃO**

A escolha da metodologia para este trabalho foi realizada após uma revisão bibliográfica para a busca do melhor modelo aplicável. A metodologia é classificada em quatro aspectos, quanto a natureza da pesquisa, a forma de abordagem do problema, o objetivo e os procedimentos técnicos (GIL 2001; SILVA 2001).

A natureza da pesquisa realizada foi a aplicada, que busca gerar conhecimentos para aplicação prática, dirigidos para a solução de problemas específicos; envolve verdades e interesses locais.

A forma de abordagem do problema foi a qualitativa, já que a pesquisa considera que existe uma relação dinâmica entre o mundo real e o sujeito que não é medida em números. Esta forma de abordagem não requer o uso de métodos ou técnicas estatísticas, o pesquisador analisa os dados indutivamente.

Do ponto de vista dos objetivos, trata-se de uma pesquisa explicativa que visa identificar os fatores que determinam ou contribuem para a ocorrência do problema e formular hipóteses. O trabalho envolve o levantamento bibliográfico, aplicação de questionários e a interação com outros setores que possuem o problema pesquisado.

Os procedimentos teóricos utilizados serão pormenorizados nas etapas a seguir e incluem a pesquisa bibliográfica, o levantamento (por meio

de entrevistas) e a pesquisa-ação, visto que o escopo da pesquisa foi aplicado e demonstrado em ação cooperativa entre o pesquisador e seus colaboradores.

A pesquisa foi realizada no setor de Diagnóstico por Imagem do A. C. Camargo Cancer Center, com a colaboração dos médicos assistentes e da equipe de Tecnologia da Informação do hospital.

### **3.1 ANÁLISE DO SISTEMA ATUAL DE LAUDOS E DESENHO DA FERRAMENTA**

Foi realizada uma reunião primeiramente entre a pesquisadora e o Diretor do Departamento de Marketing da empresa desenvolvedora do RIS, atualmente em uso na instituição, para a apresentação do trabalho e autorização da liberação dos desenvolvedores para a realização do trabalho.

Em seguida foi realizada uma reunião com a equipe médica da anatomia patológica, que serve-se de um *software* para elaboração dos relatórios da mesma empresa, sendo uma versão com muitas características semelhantes à do setor de imagem, e também está em busca de uma forma de criação de um banco de dados para os relatórios de anatomia patológica. Nessa reunião foram discutidas as iniciativas de indexação dos dois departamentos com o intuito de criar um modelo de indexação que pudesse servir para a indexação dos relatórios de ambos os departamentos.

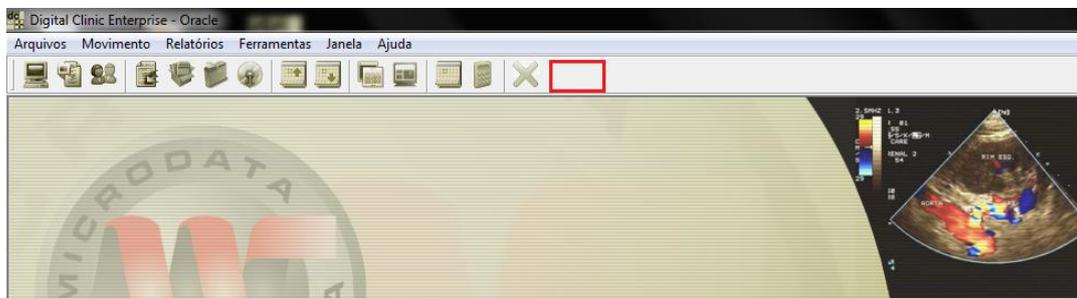
Uma segunda reunião para desenvolvimento da ferramenta foi realizada com a presença da pesquisadora, dois membros da equipe de

tecnologia da informação (TI) do A.C. Camargo Cancer Center e dois desenvolvedores de *software* da empresa Microdata. Nessa reunião, foram expostos pela pesquisadora os objetivos gerais da criação, os requisitos necessários e a funcionalidade da nova ferramenta, além dos prazos para a realização da pesquisa.

Durante a reunião, foram estabelecidos os critérios e funcionalidades da nova ferramenta. Foi proposta a criação de um ícone na tela de digitação dos relatórios destinado a inclusão dos termos indexadores. Por se tratar de uma ferramenta de teste, os termos relacionados aos relatórios servirão somente para consulta interna por usuários e pesquisadores, estas informações não constarão no relatório médico. A inclusão da indexação poderá conter mais de um termo por laudo, porque o processo de indexação será realizado em correspondência com os achados radiológicos do exame de imagem e o diagnóstico patológico do paciente. Não será modificado formato do relatório médico, que permanecerá em forma de texto livre. A inclusão do diagnóstico poderá ser realizada em qualquer momento da elaboração do relatório, antes da finalização do mesmo. O léxico utilizado será o CID-10, para possibilitar, no futuro, uma possível integração com os relatórios da patologia, já que o capítulo de neoplasia deste léxico permite uma boa correspondência com os achados da patologia.

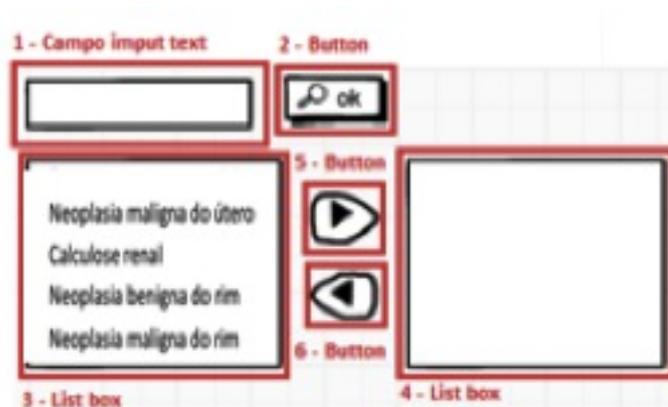
Foi levantada, ainda, a necessidade de criar uma ferramenta de busca destinada a recuperação dos termos vinculados aos relatórios. Como interface para o usuário foi proposta a criação de um ícone da tela de

entrada do sistema, que levaria a uma página destinada à busca de relatórios indexados, conforme modelo ilustrado abaixo, em vermelho.



**Figura 2** - Tela do sistema onde será incluído o ícone de busca

Uma interface de busca dentro deste ícone foi proposta, conforme figura a seguir, com um sistema de filtro manipulável, por meio do qual o usuário pudesse selecionar os critérios de busca dentre os disponibilizados no sistema.



**Figura 3** - Modelo da Ferramenta de Busca

### 3.2 MODELOS DE DESENVOLVIMENTO DE SOFTWARE

Os modelos de desenvolvimento de *software* começaram a surgir na década de 60, no intento de organizar a criação de sistemas, pois a medida que novas tecnologias foram criadas, tornou-se necessário um processo disciplinado de criação.

O modelo incremental é um modelo evolucionário, em que é criada uma versão inicial do produto, que é implementada, testada e, a partir dos comentários dos usuários, é modificada até que seja criada a versão final. Após uma conversa do usuário com o desenvolvedor do *software*, em que os requisitos do produto são apresentados é criado o “núcleo do produto”. O núcleo do produto, segundo PRESSMAN (2005), trata-se do primeiro incremento produzido, que contempla os requisitos necessários, mas algumas características deste produto, conhecidas ou não, ainda não foram elaboradas. É então realizada uma fase de teste com esta versão inicial, em que as críticas e sugestões dos usuários serão consideradas para a melhoria do produto. Até que seja implantada a versão final, que atinja as expectativas do usuário. Esta forma de trabalho pode ser útil quando a equipe de desenvolvimento disponível é insuficiente para assumir o projeto completo.

Uma das vantagens desse modelo de criação é que o usuário não precisa esperar até que o produto final seja entregue para ser utilizado, com isso, os riscos de um fracasso total são minimizados, pois embora sejam encontrados alguns problemas durante a fase de criação, pode-se tentar

solucioná-los. Além disso, as funções prioritárias são as primeiras a serem entregues (PRESSMAN 2005; SOMMERVILLE 2011)

### **3.3 ESTUDO DE VIABILIDADE DE INCLUSÃO DE INTEGRAÇÃO**

Nesta etapa, foram selecionados médicos assistentes, membros da equipe de radiologia e diagnóstico por imagem do A.C. Camargo Cancer Center, que atuam nos setores de tomografia computadorizada e ressonância magnética e que possuem vínculo com o programa de pós-graduação desta instituição. Os critérios acima utilizados visam implementação, uma vez que os estes médicos poderiam atuar como multiplicadores da ferramenta. Os médicos assistentes e pesquisadores trabalham na revisão e auxílio dos médicos residentes, podendo reforçar a necessidade de preenchimento dos dados para obtenção de um banco de dados que irá auxiliar nos próximos trabalhos a serem desenvolvidos. A amostra escolhida foi de forma não probabilística, de acordo com o julgamento da pesquisadora de que os usuários escolhidos seriam fontes de informações mais precisas e mais acessíveis para a realização do teste da ferramenta. Assim, apesar de uma amostra probabilística ser tecnicamente superior, neste caso, se fossem selecionados médicos não interessados em realizar os testes estes não seriam viáveis. Foram então selecionados dois médicos do setor de tomografia computadorizada e dois médicos do setor de ressonância magnética para iniciarem o processo de indexação com o incremento inicial. A equipe selecionada para realizar as indexações iniciais

foi reduzida, para que o projeto de pesquisa não impactasse na rotina de trabalho, evitando atrasos nos relatórios que têm prazos definidos para entrega.

Após a criação e disponibilização pelos desenvolvedores do incremento inicial da ferramenta, contendo o ícone para a inclusão dos relatórios, os médicos selecionados foram acompanhados e assessorados pela pesquisadora para inclusão dos códigos de indexação vinculados ao relatório durante um mês.

### **3.4 VERIFICAÇÃO, VALIDAÇÃO E IMPLANTAÇÃO**

O objetivo desta etapa foi realizar testes para verificação e validação da nova ferramenta, que foram realizados após o período de treinamento de inclusão dos termos realizado na etapa anterior.

A validação foi realizada por meio de entrevistas, que foram organizadas de forma estruturada, realizadas pela pesquisadora possibilitando uma integração das informações e proporcionando a estes usuários um meio de expor suas críticas e sugestões em relação ao incremento inicial. Nestas entrevistas as seguintes questões foram respondidas pelos usuários:

- Foram encontradas dificuldades para indexação dos relatórios?
- Se sim, quais as dificuldades? Não encontrou o código esperado; O código não se adequou ao achado de exame; Falha no treinamento; Outras.

- Os termos disponíveis para a indexação foram adequados?
- O tempo gasto no preenchimento causou atraso na rotina de trabalho?
- Já utiliza ferramenta semelhante em outro serviço?
- Tem alguma crítica ou sugestão para a ferramenta apresentada?

A entrevista foi composta de questões formuladas com intuito de avaliar a usabilidade do sistema de indexação. A normatização *International Organization for Standardization ISO 9241, Ergonomic requirements for office work with Visual Display Terminals*, define o conceito de usabilidade, que consiste na eficiência, eficácia e satisfação do usuário (ISO 1998). Esse conceito norteou a elaboração desta entrevista sem, contudo, realizar uma avaliação com estes critérios complexos da normativa, já que o escopo da pesquisa foi direcionado para a aplicabilidade da ferramenta e não para o processo operacional de desenvolvimento do *software*.

A verificação foi realizada por meio testes, realizados pela pesquisadora, simulando buscas no banco de dados, para a avaliação da capacidade de armazenamento de forma organizada e recuperação dos dados de forma correspondente e com agilidade. O objetivo dessa tarefa foi localizar erros, inconsistências ou ambiguidades a partir do conteúdo inserido. O MV 2000, é um HIS da empresa Microdata que apresenta alguns pontos de integração com o Digital Clinic e já é utilizado no A.C. Camargo Cancer Center. Os testes foram realizados nesse HIS, que apresenta uma integração com o RIS, já que o incremento para a recuperação das

informações armazenadas ainda não estava pronto até esta data.

Antes da implantação foram solicitadas modificações pelos usuários à pesquisadora, dentre elas a inclusão na lista o tipo de exame que o paciente está realizando, se investigação de doença, estadiamento de neoplasia, controle pós-tratamento e re-estadiamento. De acordo com os entrevistados estas informações também serão importantes para conhecermos melhor o paciente que realiza o exame no departamento de imagem. Estas sugestões foram encaminhadas aos desenvolvedores da ferramenta, mas, até a finalização deste trabalho, ainda não tinha sido entregue a versão final.

Após a liberação da versão final da ferramenta pela equipe de tecnologia da informação desenvolvedora do Digital Clinic, propõe-se o envio de e-mails para todos os médicos dos setores de tomografia computadorizada e ressonância magnética informando a respeito da nova ferramenta, destacando a importância da sua utilização e com instruções iniciais para a utilização. Propõe-se ainda a realização de um treinamento no momento de elaboração do relatório pelo médico para divulgação da nova ferramenta e incentivo para que seja utilizada pela equipe da radiologia.

## 4 RESULTADOS

A revisão da literatura destacou a importância dos relatórios radiológicos, bem como a mudança nos meios de elaboração, a partir do desenvolvimento de novas tecnologias, além da tendência de padronização destes. Incluiu, ainda, as iniciativas de unificação da linguagem médica, destacando-se a terminologia radiológica. Foram revistas também as iniciativas de automatização do processo de criação de banco de dados.

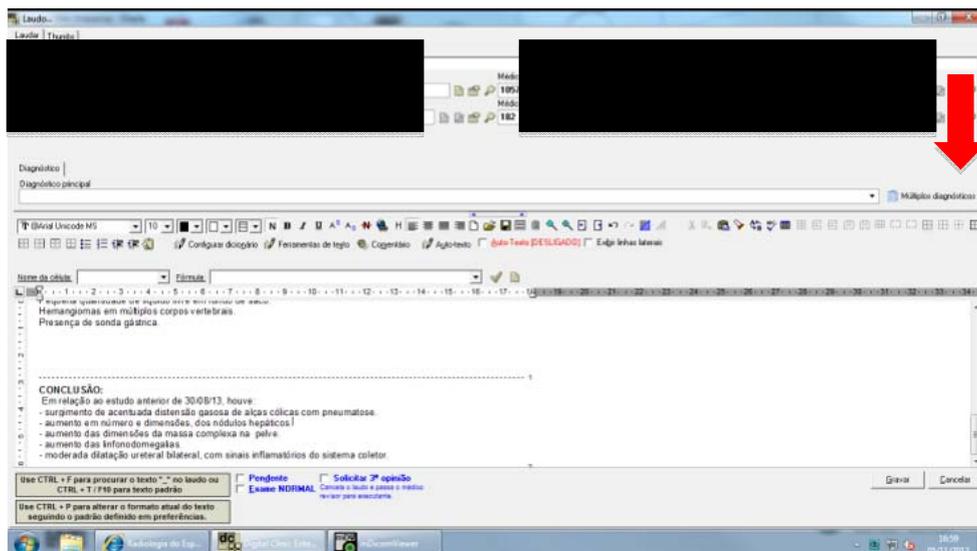
Nessa etapa, foram revisadas também as iniciativas de criação de banco de dados e sua metodologia, o que permitiu a definição dos parâmetros seguidos nas etapas seguintes desta pesquisa.

Para desenvolvimento da ferramenta de indexação dos relatórios radiológicos, foi estabelecido que o processo de indexação seria realizado de forma manual pelo usuário, ou seja, não seriam utilizados métodos de extração automatizados da informação, já que a linha de trabalho não é direcionada para os processos de computação. O léxico médico empregado foi o CID-10, que, apesar de não ser o léxico mais completo para termos radiológicos, trata-se de um léxico mundialmente difundido, com tradução oficializada para a língua portuguesa e com possibilidade de uma futura integração com o setor de anatomia patológica, mesmo que sejam escolhidas versões diferentes do CID, CID-10 ou CID-O, já que estes permitem a integração uma vez que a parte inicial dos códigos é a mesma.

A lista do CID-10 é composta por 14198 termos, para facilitar a busca destes termos pelos médicos na indexação dos relatórios a lista foi editada pela pesquisadora. Nessa etapa, foram excluídos os códigos sem finalidade, ou pouco utilizados na prática dos relatórios radiológicos, também foi acrescentada na frente do diagnóstico a topografia da lesão. A lista modificada, composta de 288 termos, foi encaminhada para o setor de TI para inclusão no sistema de laudos.

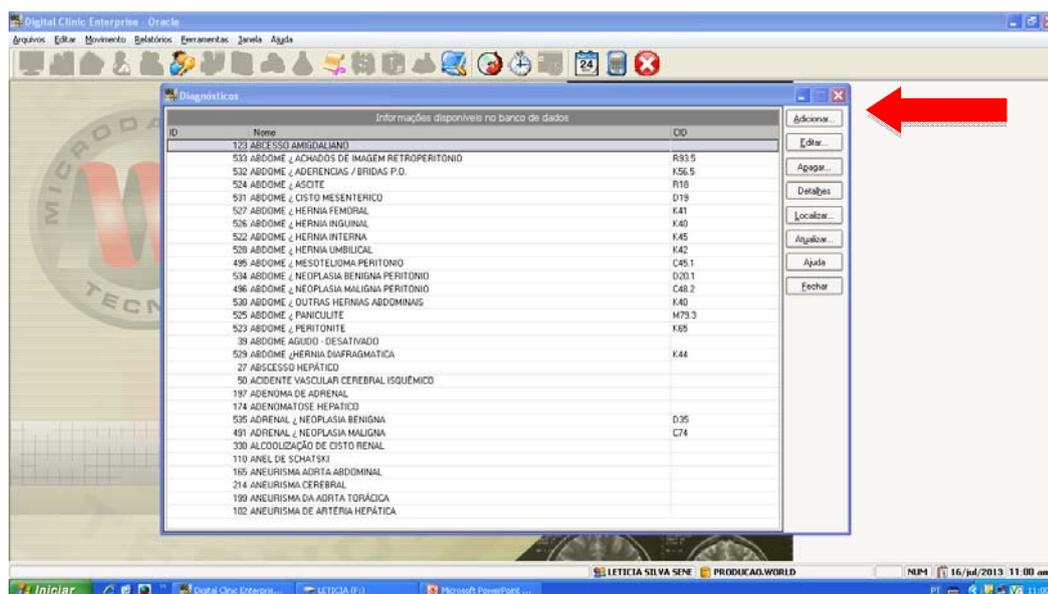
Considerando que o modelo de processo proposto foi o incremental, após a análise do sistema atual de laudos e definição dos requisitos necessários para a ferramenta, a versão inicial da ferramenta foi entregue pelos desenvolvedores. Esta versão inicial tinha por objetivo atender aos requisitos iniciais e principais, ou seja, à criação de um ícone para a inclusão dos diagnósticos e uma ferramenta de busca.

A figura a seguir mostra a interface inicial da ferramenta. Na tela de elaboração de laudos foi incorporado um ícone “múltiplos diagnósticos”, destinado a inclusão do código de indexação. Assim, o médico, com a tela de laudos aberta, clica com o botão da esquerda do *mouse* e é direcionado para uma tela com os léxicos.



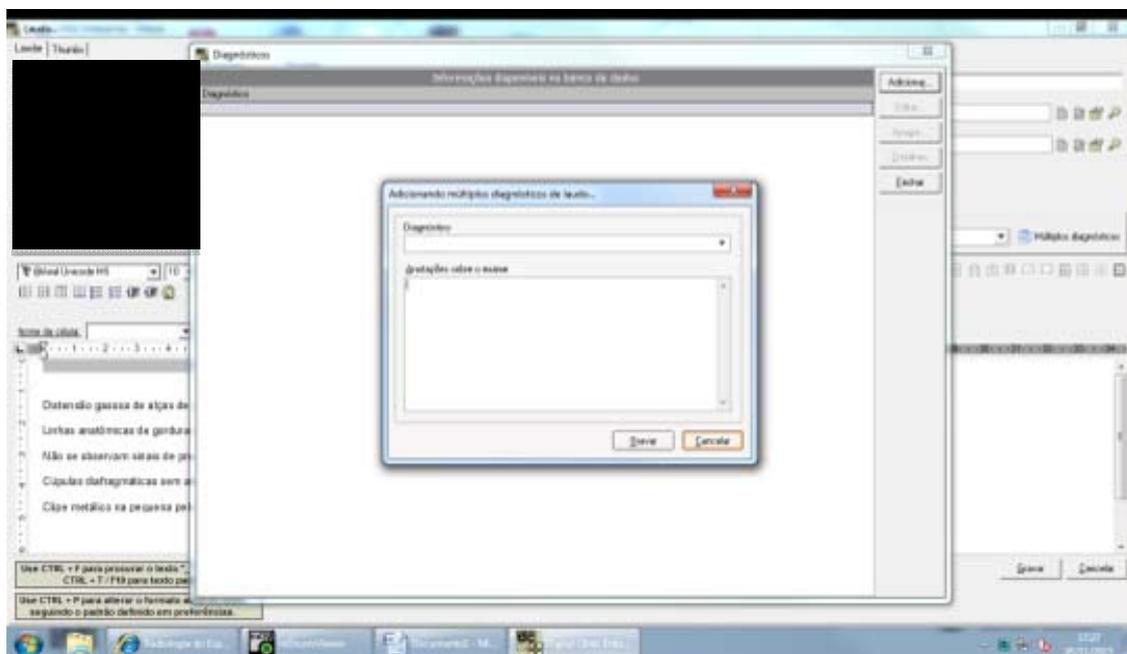
**Figura 4 - Ícone Múltiplos Diagnósticos**

Na Figura 5, é ilustrada a interface por onde são adicionados os diagnósticos dos exames. Por meio desta interface, é possível vincular um exame a um código indexador, sendo possível que cada exame tenha até 3 códigos. A tela demonstra a lista de termos disponíveis e o ícone destinado a inclusão, “Adicionar”, apontado na seta vermelha.



**Figura 5 - Léxico Criado**

A Figura 6 mostra a interface de inclusão de uma observação a respeito do exame. Essa figura ilustra o campo criado para possibilitar a inclusão de informações ao exame que podem ter relevância no momento de recuperação do relatório, mas que não apresentavam, no momento de inclusão, um código correspondente, ou até uma informação a respeito do exame que não foi adicionada ao relatório, por ser uma informação ainda não confirmada, uma opinião pessoal do usuário, ou até mesmo um lembrete para o próprio usuário no momento da busca. Esses diagnósticos não aparecem no relatório final entregue ao paciente e também não ficam acessíveis no prontuário, somente são acessados mediante a ferramenta de busca que será demonstrada na Figura 7.



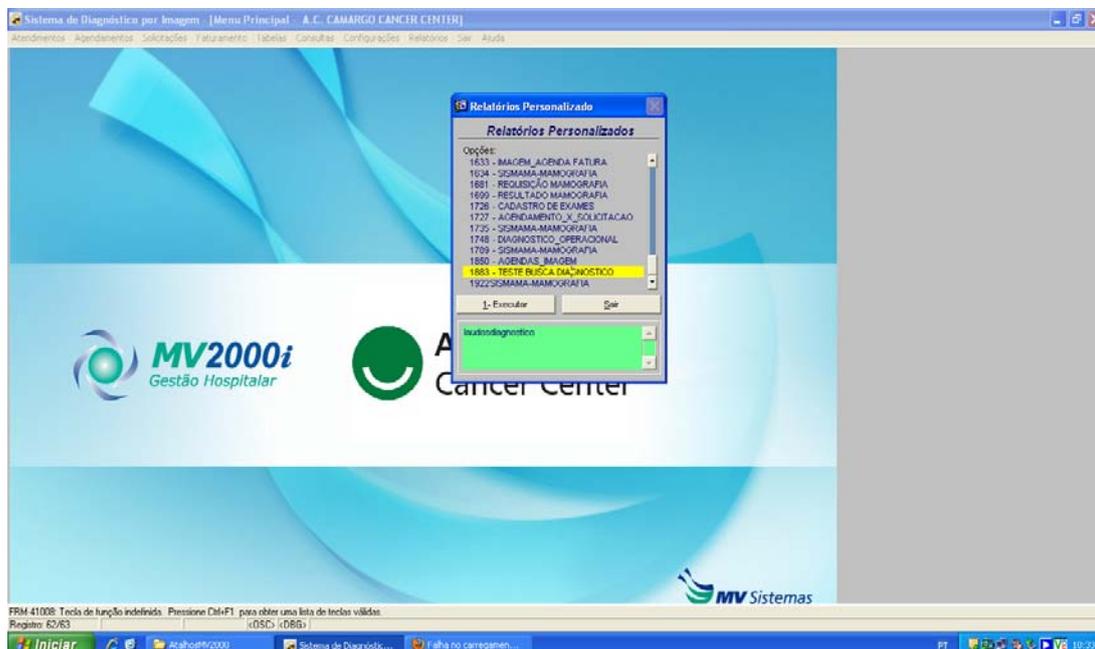
**Figura 6** - Interface para Observações

Durante um mês, os médicos foram acompanhados e assessorados pela pesquisadora para a inclusão dos códigos de indexação vinculados ao relatório. Nessa fase de teste, foram incluídos 163 diagnósticos.

Após esse mês de teste de inclusão de diagnósticos, foi realizada a validação da ferramenta. A validação foi realizada por meio de uma entrevista individual com cada usuário, com modelo anexado, Anexo 1, destinado aos usuários. Todos os usuários relataram que a principal dificuldade foi lembrar de inserir o diagnóstico, que, mesmo na presença da pesquisadora ao lado, muitas vezes, essa etapa da inclusão não foi realizada. Foi sugerida por dois usuários que o sistema gerasse uma forma de lembrete no final da emissão do relatório para o usuário que ainda não tivesse incluído o relatório no banco de dados. A segunda crítica foi para que a lista fosse mais clara e direta, uma vez que a inclusão da topografia antes do termo criada pela pesquisadora foi vista como um fator de confusão no momento da inclusão. A terceira crítica foi em relação ao tempo gasto para o preenchimento, como se trata de uma etapa nova, esse preenchimento determinou um gasto de tempo a mais na finalização do relatório, porém os usuários não souberam quantificar este tempo.

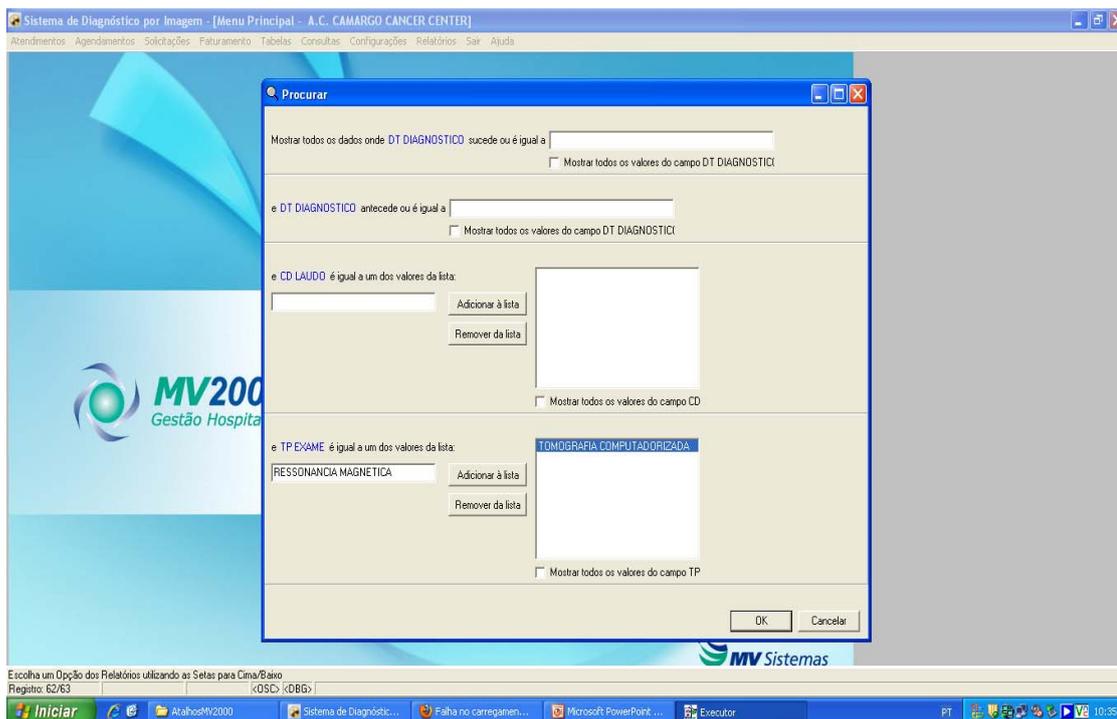
A verificação da ferramenta foi realizada mediante testes simulando buscas no banco de dados. Os testes foram realizados em um *software* auxiliar ao de elaboração dos laudos, já que o incremento da busca ainda não estava pronto até a data de conclusão deste trabalho. Foi criado um ícone “teste busca diagnóstico” dentro da área destinada a geração de relatórios pelo HIS. A partir deste ícone o usuário é direcionado para uma

tela com os filtros da pesquisa.



**Figura 7** - Tela para a Simulação de Busca de Diagnóstico

A figura a seguir exibe os filtros disponíveis para a realização da consulta dos diagnósticos indexados. O primeiro campo e o segundo campo correspondem a um termo do léxico, para busca específica. O quarto campo corresponde ao tipo de exame, se tomografia computadorizada ou ressonância magnética. O terceiro campo não apresenta correspondência prática para a busca e deverá ser retirado do filtro.



**Figura 8 - Filtro de Busca Manipulável**

O resultado da busca está demonstrado em um modelo de simulação na figura abaixo. A interface demonstrada representa a busca de exames que continham diagnóstico, em um período de tempo determinado. Na lista, os seguintes dados são levantados: nome e número de registro do paciente; tipo de exame; médico que emitiu o relatório; diagnósticos incluídos e as observações preenchidas na inclusão.

Na busca por data todos os 163 relatórios catalogados foram recuperados. A partir desta recuperação outro teste foi realizado. Alguns dos termos recuperados pela busca realizada por data foram incluídos no primeiro campo para realizar uma busca direcionada para o termo, que foram recuperados.



## 5 DISCUSSÃO

O A.C. Camargo Cancer Center tornou-se, ao longo do tempo, referência internacional na pesquisa e tratamento do câncer. Deste modo, trabalhos direcionados para dar suporte nas pesquisas tornam-se importantes. O setor de diagnóstico por imagem dispõe de um arquivo digital de grande relevância científica, mas que não permitia a recuperação eficiente deste arquivo. Portanto, é de fundamental importância a criação de uma forma de indexação para organizar a informação e facilitar a busca, auxiliando futuras atividades de ensino e pesquisa.

A recuperação eficiente da informação médica é essencial para os aspectos funcionais de um sistema de saúde. Com o aumento da quantidade de documentos disponibilizados nos RIS o processo de busca destas informações é prejudicado. O processo de busca é possível pela introdução de códigos em um banco de dados.

A integridade e exatidão das informações recuperadas dependem, portanto, do sistema de codificação utilizado. Com o recente crescimento da tecnologia da informação, nos dias atuais uma maior quantidade de dados e com maior detalhamento podem ser codificados e processados (VARDY et al. 1988).

STANFILL et al. (2010) realizaram uma revisão sistemática da literatura analisando estudos que avaliaram todos os tipos de codificação automatizadas e sistemas de classificação para determinar o desempenho

de tais sistemas. Foram incluídos 113 estudos anteriores a março de 2009. Mediante esta revisão, concluiu-se que a codificação clínica automatizada e o desempenho do sistema de classificação é relativa à complexidade da tarefa e ao resultado desejado. Codificações automatizadas e sistemas de classificação em si não são generalizáveis, tampouco os resultados destes estudos. Sendo assim, a codificação automática e os sistemas de classificação são, ainda, uma promessa, e a aplicação da codificação automatizada deve ser considerada no contexto. O desenvolvimento destes sistemas e uma melhor compreensão das tarefas para as quais eles serão utilizados são necessários antes que se possa concluir que a codificação automatizada e os sistemas de classificação possam atender aos padrões de desempenho adequados para uso em processos de codificação clínicos complexos, sendo capazes de aplicar as orientações adequadas para relatar estes dados (STANFILL et al. 2010).

O trabalho realizado se constituiu na criação de um banco de dados de forma manual e em uma ferramenta de busca para a recuperação dos dados inseridos.

Para o desenho da ferramenta, o ponto de partida foi a escolha da linha de trabalho, direcionada para a aplicabilidade da ferramenta de indexação. Assim, o sistema de codificação selecionado foi o manual. Optou-se pela codificação manual dos achados de imagem, já que o sistema de indexação automatizado é direcionado a outra linha de pesquisa.

Depois de definida a forma de indexação, foi então estabelecido o léxico a ser empregado. Os estudos mostraram que o uso dos termos do

RadLex são fundamentais para a manutenção de terminologia atualizada diante dos constantes avanços tecnológicos na área. Esta estratégia pretende minimizar a obsolescência do banco didático, mas esbarra em dois importantes aspectos complicadores para a sua utilização neste trabalho. O primeiro aspecto é que os termos estão em língua inglesa, assim, o tempo necessário para a tradução além, de extrapolar o tempo do projeto, é praticamente inviável sem um órgão centralizador no Brasil e a adoção de um termo apropriado em português exigiria consenso entre especialistas, já existem iniciativas para a tradução ainda não concluídas. O Segundo aspecto é que os termos do RadLex são constantemente atualizados. Nos últimos meses, foram observadas diversas alterações significativas na árvore de termos disponível, sendo assim, seria necessária a atualização constante dos termos utilizados (PINHEIRO et al. 2009).

Assim, o léxico escolhido para a codificação foi o CID 10, que, apesar de não conter toda a terminologia radiológica, é um léxico mundialmente empregado na área médica, com tradução para a língua portuguesa, possui ainda um capítulo exclusivo para oncologia, que é o foco de atendimento do hospital, e possibilita uma integração de dados com a anatomia patológica. O léxico CID-10 possui, ainda, códigos para classificar o motivo de realização do exame, proporcionando, também, a inclusão desta categoria.

A criação da ferramenta foi de acordo com o modelo de processo de *software* incremental, realizado em etapas, cada incremento entregue foi testado de forma antecipada e modificado, antes que o produto final fosse implantado. Este modelo é útil quando o desenvolvedor não possui mão de

obra suficiente para entrega do produto final no prazo estipulado, como no projeto em questão, já que a empresa desenvolvedora precisa realizar esse trabalho em paralelo com outros projetos do próprio hospital que apresentam maior impacto financeiro.

Os resultados obtidos da análise qualitativa das entrevistas evidenciaram que todos os usuários encontraram dificuldades para a utilização da ferramenta. As dificuldades encontradas pelos usuários estão relacionadas a interface para a busca dos termos para a indexação, de acordo com os relatos, esta busca demandou muito tempo dos usuários para encontrar o termo que se adequasse ao relatório.

O sistema mostrou-se compatível com o banco de dados testados. Não foram encontrados problemas na recuperação dos dados armazenados. Na forma como foi concebido, o modelo criado cumpre o papel de um banco de dados que pode ser utilizado como alicerce para o ambiente de ensino e pesquisa em radiologia e diagnóstico por imagem. As etapas preliminares para a implementação da nova ferramenta já foram completadas. Ao longo do projeto, foram vislumbradas novas possibilidades de melhorias que não puderam ser desenvolvidas, o que corrobora os problemas no desenvolvimento de software apontados por PRESSMAN (2005), no quesito prazo e orçamento estipulado. Entre as possibilidades, destacam-se melhorias que simplifiquem o uso da ferramenta, a criação de uma forma de lembrete para os usuários incluírem os dados, a utilização de um léxico mais voltado para a radiologia e a inclusão do ícone de busca no menu inicial, para que a busca não necessite ser realizada em outro programa.

## 6 CONCLUSÕES

Apesar de já ter sido estabelecida a importância da recuperação de informações e a necessidade de criação de técnicas de busca na prática radiológica para minimizar a perda de informações importantes, a literatura direcionada para o médico radiologista é ainda escassa. Assim, a proposta de criação de um modelo de ferramenta para a indexação dos relatórios radiológicos encontrou dificuldades, principalmente porque, grande parte da literatura disponível é direcionada aos profissionais da área de ciência da computação, enfatizando mais o processo de desenvolvimento do que a aplicabilidade do *software*.

De qualquer forma, as referências utilizadas foram importantes no direcionamento da metodologia da pesquisa proporcionando uma mudança no cenário da recuperação de informações. Apesar de a busca pelo achado do exame já ser uma demanda antiga dos radiologistas, antes do início do trabalho, a recuperação dos relatórios pelo RIS era realizada apenas com as informações demográficas associadas ao relatório, como a identificação do paciente, a data do exame, o profissional solicitante ou o executor do exame. Nesse sentido, o modelo desenvolvido tornou possível a indexação de relatórios radiológicos e a recuperação fidedigna dos relatórios armazenados a partir do termo inserido. Todavia, muito embora o modelo tenha sido bem sucedido em suas etapas iniciais, durante a fase de implantação algumas complicações foram percebidas levando à conclusão

de que melhorias ainda precisam ser alcançadas para que o processo de indexação seja parte da rotina dos médicos radiologistas do Departamento de Diagnóstico por Imagem do A.C. Camargo Cancer Center. As principais complicações apontadas pelo trabalho estão relacionadas à implantação de uma etapa a mais na rotina de liberação de relatórios. Os usuários selecionados para os testes iniciais relataram dificuldade ao assimilar o processo de indexação como parte da rotina de trabalho e identificaram o tempo despendido com esse novo processo um fato que atuou como barreira para a implementação da ferramenta.

Por fim, o trabalho reafirma a importância do tema e demonstra a viabilidade da criação de um processo de indexação de relatórios radiológicos. Sem exaurir as possibilidades e considerando os limites da pesquisa realizada, conclui-se que o processo de criação do banco de dados é possível, mesmo com pouca tecnologia adicional envolvida, o que serve de estímulo para novas pesquisas relacionadas ao tema.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

A.C. Camargo Cancer Center. **Nossos números**. 2012. Disponível em: <URL:<http://www.accamargo.org.br/nossos-numeros/>> [2014 jan 12]

Almeida MB. Roteiro para a construção de uma ontologia bibliográfica através de ferramentas automatizadas. **Perspect Ciênc Inf** 2003; 8:164-79.

Almeida M, Bax M. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Rev Ciência Infor** 2003; 32:7-20.

Azevedo-Marques PM, Caritá EC, Benedicto AA, Sanches PR. Integração RIS / PACS no Hospital das Clínicas de Ribeirão Preto: uma solução baseada em "Web". **Radiol Bras** 2005; 38:37-43.

Boehme JM 2nd, Choplin RH. Systems integration: requirements for a fully functioning electronic radiology department. **Radiographics** 1992; 12:789-94.

Brämer GR. International statistical classification of diseases and related health problems. **World Health Stat Q** 1988; 41:32-6.

Burnside ES, Sickles EA, Bassett LW, et al. The ACR BI-RADS experience: learning from history. **J Am Coll Radiol** 2009; 12:851-60.

Do BH, Wu A, Biswal S, Kamaya A, Rubin DL. RADTF: A semantic search-enabled, natural language processor-generated radiology teaching file. **Radiographics** 2010; 30:2039-48.

Douglas PS, Hende RC, Cummings J, et al. ACCF/ACR/AHA/ASE/ASNC/HRS/NASCI/RSNA/SAIP/SCAI/SCCT/SCMR 2008 health policy statement on structured reporting in cardiovascular imaging. **J Am Coll Cardiol** 2009; 53:76-90.

Dunnick NR, Langlotz CP. The radiology report of the future: a summary of the 2007 Intersociety Conference. **J Am Coll Radiol** 2007; 015: 626-29.

Duque CG. **Uma abordagem ontológica para a indexação de documentos eletrônicos.** 2006. Disponível em: <URL:[http://www.filologia.org.br/ileel/artigos/artigo\\_295.pdf](http://www.filologia.org.br/ileel/artigos/artigo_295.pdf)> [2014 jan 15]

Garfield E. A retrospective and prospective view of information retrieval and artificial intelligence in the 21st Century. **J Am Soc Inf Sci Technol** 2001; 52:18-21.

Gil AC. **Como elaborar projetos de pesquisa.** São Paulo: Atlas; 2001.

Hall MF. The radiology report of the future. **Radiology** 2009; 251:313-16.

Holanda AJ, Bailão LA, Pisa IT, Ruiz EES. **Aplicação de padrões em um sistema de gerenciamento de dados para uma UTI neonatal.** 2004. Disponível em: <URL:<http://www.sbis.org.br/cbis9/arquivos/695.pdf>> [2014 jjan 8]

[ISO] International Organization for Standardization. **Ergonomic requirements for office work with visual display terminals (VDTs): part 11 - Guidance on usability (ISO 9241-11:1998(E)).** 1998. Available from: <URL:<http://www.it.uu.se/edu/course/homepage/acsd/vt09/ISO9241part11.pdf>> [2014 jan 12]

Langlotz CP. RadLex: A new method for indexing online educational materials. **Radiographics** 2006; 26:1595-97.

Mongkolwat P, Channin D, Kleper V, Rubin DL. An open-source and open-access cancer biomedical informatics grid annotation and image markup template builder. **Radiographics** 2012; 32:1223-32.

[OMS] Organização Mundial de Saúde. **Classificação internacional de doenças para oncologia**. Trad. do Centro Colaborador da OMS para a Classificação de Doenças em Português. 3<sup>rd</sup> ed. São Paulo: EDUSP; 2005.

Pinheiro MG, Canalli HL, Figueiredo LR, Elias Júnior J. Ambiente computacional para ensino de radiologia e diagnóstico por imagem: uma proposta para arquivo didático. **J Health Informatics** 2009; 1:43-52.

Pressman RS. **Software engineering a practitioner's approach**. 2005. Available from: <URL:[http://ce.sharif.edu/courses/84-85/1/ce474/resources/root/Pressman\\_Software%20Engineering.pdf](http://ce.sharif.edu/courses/84-85/1/ce474/resources/root/Pressman_Software%20Engineering.pdf)> [2014 jan 5]

Ramaswamy MR, Patterson DS, Yin L, Goodacre BW. MoSearch: A radiologist-friendly tool for finding-based diagnostic report and image retrieval. **Radiographics** 1996; 16:923-33.

Rector AL. Clinical terminology: why is it so hard? **Methods Inf Med** 1999; 38:239-52.

Reiner BI. Strategies for radiology reporting and communication: part 2: using visual imagery for enhanced and standardized communication. **J Digital Imaging** 2013; 26:838-42.

Rodrigues BC, Crippa G. A recuperação da informação e o conceito de informação: o que é relevante em mediação cultural? **Perspect Ciênc Inf** 2011; 16:45-64.

Rothke M, Blondin D, Schlemmer HP, Franiel T. PI-RADS classification: structured reporting for mri of the prostate. **ROFO** 2013; 185:253-61.

Rubin LD, Noy NF, Musem MA. Protégé: a tool for managing and using terminology in radiology applications. **J Digital Imaging** 2007; 2007; 20 Suppl 1:34-46.

Russ G, Royer B, Bigorgne C, Rouxel A, Bienvenu-Perrard M, Leenhardt L. Prospective evaluation of thyroid imaging reporting and data system on 4550 nodules with and without elastography. **Eur J Endocrinol** 2013; 168:649-55.

Seeley GW, McNeill K, Dallas WJ. Displays in radiology: past, present and future. **Radiographics** 1987; 7:1261-66.

Serapião PRB, Honorio Junior R, Santos MA, Santos LRA, Moraes JCB, Marques PMA. Construção de um índice de informação da prática clínica em radiologia e diagnóstico por imagem baseada em mineração de texto e tesauro. **Radiol Bras** 2013; 46:290-98.

Silva EL. **Metodologia da pesquisa e elaboração da dissertação**. Florianópolis: Laboratório de Ensino à Distância da UFSC; 2001.

[SNOMED CT] Systematized Nomenclature of Medicine Clinical Terms. 2006. Available from: <URL:<http://www.ihtsdo.org/snomed-ct>> [2014 jan 8]

Sommerville I. **Software engineering**. 9<sup>th</sup> ed. 2011. Available from: <URL:[http://sandeepdasari.files.wordpress.com/2013/09/softwareengineering\\_sommerville.pdf](http://sandeepdasari.files.wordpress.com/2013/09/softwareengineering_sommerville.pdf)> [2014 jan 13]

Spyns P. Natural language processing in medicine: an overview. **Methods Inf Med** 1996; 35:285-301.

Stanfill MH, Williams M, Fenton SH, Jenders RA, Hersh WR. A systematic literature review of automated clinical coding and classification systems. **J Am Med Inform Assoc** 2010; 17:646-51.

Stevenson M, Wilks Y. Word-sense desambiguation. In: Ruslan M, editor. **The oxford handbook of computational linguistics**. New York: Oxford University Press; 2003. p.249-54.

Vardy DA, Gill RP, Israeli A. Coding Medical Information: classification versus nomenclature and implications to the Israeli Medical System. **J Med Syst** 1998; 22:203-10.

Varma DR. Managing DICOM images: tips and tricks for the radiologist. **Indian J Radiol Imaging** 2012; 22:4-13.

[WHO] World Health Organization. **International statistical classification of diseases and related health problems**. 10<sup>th</sup> revision. 2010. Available from: <URL:<http://www.who.int/classifications/icd/en>> [2014 jan 13]

Weiss DL, Langlotz CP. Structured reporting: patient care enhancement or productivity nightmare. **Radiology** 2008; 249:739-47.

Yam CS, Kruskal J, Sitek AL. A web-based ACR index for radiological diagnoses. **AJR Am J Roentgenol** 2004; 183:1517-21.

Anexo 1 – Entrevista de Avaliação da Ferramenta Pelos Usuários



Foram encontradas dificuldades para indexação dos relatórios?

Sim  Não

---

Quais dificuldades foram encontradas?

Não encontrou o código esperado

O código encontrado não se adequou ao diagnóstico ou achado de exame

Falha no treinamento

Outras

---

Os códigos disponíveis para a indexação foram adequados?

Sim  Não

O tempo gasto no preenchimento dos diagnósticos causou atraso na rotina de trabalho?

Sim  Não

---

Já utiliza ferramenta semelhante em outro serviço?  Sim  Não

Tem alguma crítica ou sugestão em relação a ferramenta apresentada?

---

Comentários:

---

---

---

---

## Anexo 2 - Carta de Aprovação do Comitê de Ética em Pesquisa

São Paulo, 04 de Junho de 2012.

Ao  
**Dr. Rubens Chojniak**

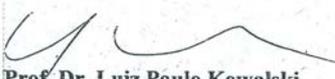
**Ref.: Projeto de Pesquisa nº. C11/12**  
**“Criação e Implantação de um sistema de Indexação de laudos radiológicos em um centro de referência no ensino e pesquisa em oncologia”.**

Os membros do Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antonio Prudente – Hospital do Câncer - A.C. Camargo/SP, em sua última reunião de 29/05/2012, **aprovaram** a realização do estudo em referência, datado de Março de 2012 e tomaram conhecimento dos seguintes documentos:

- Termo de Compromisso do Pesquisador com as Resoluções do Conselho Nacional de Saúde;
- Termo de Compromisso para Utilização de Banco de Dados;
- Orçamento Financeiro Detalhado;
- Declaração de Infraestrutura e Instalações do Departamento de Diagnóstico por Imagem;
- Declaração de Ciência e Comprometimento do Departamento de Diagnóstico por Imagem.

**Informações a respeito do andamento do referido projeto deverão ser encaminhadas à assistente do CEP dentro de 12 meses.**

Atenciosamente,

  
**Prof. Dr. Luiz Paulo Kowalski**  
Coordenador do Comitê de Ética em Pesquisa

1/1