

**UNIVERSIDADE FEDERAL DO PARANÁ**

**RODRIGO CEZAR MAIA TEIXEIRA**

**PROCESSO DE ONTOLOGIAS APLICADO AO ACERVO FOTOGRÁFICO DA  
COMPANHIA DE SANEAMENTO DO PARANÁ - SANEPAR**

**CURITIBA**

**2010**

**RODRIGO CEZAR MAIA TEIXEIRA**

**PROCESSO DE ONTOLOGIAS APLICADO AO ACERVO FOTOGRÁFICO DA  
COMPANHIA DE SANEAMENTO DO PARANÁ - SANEPAR**

Trabalho de Conclusão de Curso  
apresentado à Disciplina de Projeto de  
Pesquisa em Informação II, do curso  
Gestão da Informação, Setor de Ciências  
Sociais Aplicadas, Universidade Federal  
do Paraná.

Orientador: Prof. Mauro Belli

**CURITIBA**

**2010**

## RESUMO

O presente trabalho realiza estudos exploratórios sobre o conceito de ontologia, verificando a possibilidade de utilização de ontologias como meio para diminuir dificuldades relacionadas com a recuperação da informação. Para isso utilizou-se como objeto de estudo o Acervo Fotográfico da SANEPAR onde criou-se uma amostra que apoiada em um aplicativo de construção de ontologias buscou utilizar ontologias para a indexação e recuperação dos documentos do acervo com o objetivo de melhorar a localização de informação. Caracteriza-se por ter uma abordagem científica qualitativa e caráter descritivo exploratório quanto à forma de pesquisa. A metodologia adotada na execução contemplou as seguintes fases: planejamento, especificação, aquisição de conhecimento, conceitualização, formalização, integração, implementação e avaliação. Como resultado deste trabalho, relata-se que a utilização de ontologias traz benefícios para processos de recuperação de informação além de permitir maior interoperabilidade e compartilhamento de informações entre sistemas distintos. Identifica a possibilidade do emprego de ontologias como forma de disponibilizar informações e conhecimento de forma mais estruturada.

Palavras-Chave: Ontologia. OWL. Protege. Recuperação de Informação. Informação. Conhecimento.

## LISTA DE QUADROS

QUADRO 1 – METODOLOGIAS PARA CONSTRUÇÃO DE ONTOLOGIAS.....	15
QUADRO 2 – LINGUAGENS PARA CONSTRUÇÃO DE ONTOLOGIAS .....	18
QUADRO 3 – FERRAMENTAS PARA CONSTRUÇÃO DE ONTOLOGIAS .....	22
QUADRO 4 – LISTA DE TERMOS .....	30
QUADRO 5 – DESCRIÇÃO FOTOGRÁFICA .....	34

## LISTA DE FIGURAS

FIGURA 1 – REPRESENTAÇÃO DO CICLO DE VIDA DA ONTOLOGIA.....	17
FIGURA 2 – DIAGRAMA DE CLASSES .....	31
FIGURA 3 – CLASSES .....	32
FIGURA 4 – ATRIBUTOS .....	33
FIGURA 5 – ASSINATURA SANEPAR.....	34
FIGURA 6 – VALORES “FOTOGRAFIA” .....	35
FIGURA 7 – VALORES “AUTOR” .....	36
FIGURA 8 – VALORES “CONSERVAÇÃO”.....	36
FIGURA 9 – VALORES “CROMIA” .....	37
FIGURA 10 – VALORES “FORMATO”.....	37
FIGURA 11 – VALORES “MEIO” .....	38
FIGURA 12 – VALORES “PERSONAGEM”.....	38
FIGURA 13 – VALORES “VÍNCULO_PROFSSIONAL” .....	39
FIGURA 14 – FERRAMENTA DE BUSCA.....	39

## LISTA DE ABREVIATURAS E SIGLAS

AT&T – American Telephone and Telegraph  
CODE4 – Conceptually Oriented Description Environment  
DAML – Darpa Agent Markup Language  
DL – Description Logic  
GKB – Editor Generic Knowledge Base Editor  
HTML – Hypertext Markup Language  
IA – Inteligência Artificial  
IKARUS – Intelligent Knowledge Acquisition and Retrieval Universal System  
JOE – Java Ontology Editor  
KIF – Knowledge Interchange Format  
MYSQL – My Structured Query Language  
OC – Organização do Conhecimento  
OIL – Ontology Inference Layer  
OML – Ontology Markup Languages  
OWL – Web Ontology Language  
PHP – Personal Home Page Hypertext Preprocessor  
RDF – Resource Description Framework  
RDFS – Resource Description Framework Schema  
SANEPAR – Companhia de Saneamento do Paraná  
SHOE – Simple HTML Ontology Extensions  
URI – Uniform Resource Identifiers  
VC – Vocabulário Controlado  
W3C – World Wide Web Consortium  
XML – eXtensible Markup Language

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>8</b>
<b>1.1</b>	<b>PROBLEMA DE PESQUISA</b>	<b>9</b>
<b>1.2</b>	<b>JUSTIFICATIVA</b>	<b>9</b>
<b>1.3</b>	<b>OBJETIVOS</b>	<b>10</b>
<b>2</b>	<b>LITERATURA PERTINENTE</b>	<b>12</b>
<b>2.1</b>	<b>ORGANIZAÇÃO DO CONHECIMENTO</b>	<b>12</b>
<b>2.2</b>	<b>ONTOLOGIA</b>	<b>13</b>
<b>2.3</b>	<b>PROCESSOS DE CONSTRUÇÃO DE ONTOLOGIA</b>	<b>15</b>
2.3.1	Methontology	17
<b>2.4</b>	<b>LINGUAGENS DE ONTOLOGIA</b>	<b>18</b>
2.4.1	Linguagem OWL	20
<b>2.5</b>	<b>APLICATIVOS PARA CONSTRUÇÃO DE ONTOLOGIAS</b>	<b>21</b>
2.5.1	Ontowiki	23
2.5.2	Protégé	24
<b>2.6</b>	<b>WEB SEMÂNTICA</b>	<b>25</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>28</b>
<b>4</b>	<b>ANÁLISE DOS RESULTADOS</b>	<b>30</b>
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>41</b>
	<b>REFERÊNCIAS</b>	<b>42</b>
	<b>APÊNDICES</b>	<b>45</b>

## 1 INTRODUÇÃO

Desde a Pré-história até o século XIX, os meios de produção eram baseados essencialmente na manipulação de recursos naturais e na utilização da força física. Os conhecimentos gerados nas organizações dessa época se desenvolviam muito lentamente. A partir da metade do século XIX, houve, contudo, um grande progresso tecnológico e de invenções científicas. A informação passou a ser usada cada vez mais nos processos operacionais.

O sucesso de uma empresa nos dias atuais decorre da sua capacidade de criação do conhecimento organizacional, difundido na organização e incorporado a produtos, serviços e sistemas. Isso poderá garantir mais facilmente para as empresas inovação tecnológica, novos produtos e aquisição de novos mercados.

Modelos gerenciais, sistemas administrativos informatizados, pesquisas, desenvolvimento e engenharia, patentes e cultura organizacional integram a estrutura interna (os relacionamentos internos) que, associados às pessoas, constituem o que chamamos de organização. (FIALHO *et al.*, 2006. p.74)

Fialho *et al.*, (2006) citam ainda que “na organização do conhecimento, a informação assume realidade própria, separada dos bens tangíveis que conhecemos, tornando-se peça fundamental no ambiente competitivo”.

Como instrumentos de organização do conhecimento aparecem às ontologias, consideradas uma alternativa para minimizar o problema da quantidade excessiva de informações, uma vez que facilidades de documentação, manutenção e confiabilidade são características importantes das ontologias, assim como as propriedades de compartilhamento e filtragem.

Atualmente as empresas trabalham com um número cada vez maior de informações e encontram grandes dificuldades na organização do conhecimento. É importante saber gerir estas informações para utilizar o conhecimento extraído delas, e armazenar somente o que for necessário, evitando o acúmulo de informações irrelevantes e o desperdício em seu tempo de recuperação.

Este trabalho tem como proposta realizar estudos exploratórios sobre o conceito de ontologia, verificando a utilização de ontologias como meio para diminuir as dificuldades relacionadas com a recuperação da informação.

## 1.1 PROBLEMA DE PESQUISA

Na Sociedade competitiva em que vivemos o conhecimento transformou-se no principal fator de produção, no elemento fundamental para a diferenciação e destaque dentro e fora das instituições. O conhecimento passa a ser a matéria-prima fundamental, e o capital intelectual passa a ser um dos mais importantes recursos das pessoas e organizações, seu diferencial estratégico.

O aspecto problemático da cultura de nosso tempo está relacionado à questão informacional, é o assim chamado fenômeno da explosão informacional, a grande quantidade de informações produzidas e disponibilizadas por diferentes atividades sociais, dificultando sua identificação, recuperação, acesso e utilização.

Mesmo após muitos anos de estudos e inovações tecnológicas, ainda existe certo grau de dificuldade na tarefa de localização e recuperação de informação. É com base nesta dificuldade que define-se como problema de pesquisa para este trabalho, verificar a possibilidade de utilizar ontologias com o objetivo de diminuir as dificuldades relacionadas com a recuperação da informação.

## 1.2 JUSTIFICATIVA

O número crescente de fontes de dados disponíveis torna cada vez mais difícil a seleção, aquisição e combinação de dados. Dentre as propostas da literatura, um grande número utiliza ontologias como forma de organizar e caracterizar dados. As ontologias atuam sobre fontes de dados, proporcionando organização e, conseqüentemente, recuperação mais efetiva. Possibilitam uma compreensão comum e compartilhada de um domínio, onde ocorre interação entre pessoas e sistemas. Desempenham um papel importante no intercâmbio de informações ao proporcionar estrutura semântica às fontes de dados. A comunicação entre os agentes envolvidos nos processos (computadores ou pessoas) se torna mais efetiva ao serem reduzidas diferenças conceituais ou terminológicas.

Para apropriação do conhecimento relacionado a estas técnicas, bem como a organização do conhecimento, foi selecionado como ambiente de trabalho, o setor de Patrimônio Histórico da Companhia de Saneamento do Paraná (SANEPAR), o qual possui um vasto acervo de materiais históricos referentes ao saneamento do Estado do Paraná, visando preservar a cultura da organização e a história do saneamento e difundir as informações contidas neste acervo, disponibilizando todo este recurso histórico aos seus “clientes”: funcionários da SANEPAR, estudantes, professores, pesquisadores, habitantes do Estado do Paraná e público em geral, interessado na história do saneamento do Estado e informações sobre o assunto.

Dentro deste material encontra-se o acervo histórico fotográfico da SANEPAR, composto de mais de 2.500 registros iconográficos catalogados sem qualquer padronização. Portanto o acervo será objeto de estudo para análise e aplicação de métodos e ferramentas de ontologias deste projeto, como meio de diminuir as dificuldades relacionadas à organização e recuperação das informações no setor de Patrimônio Histórico.

### **1.3 OBJETIVOS**

Este trabalho tem como objetivo principal o estudo das técnicas de organização do conhecimento (OC) com intuito de avaliar a utilização de ontologias como ferramenta para redução das dificuldades relacionadas à organização e a recuperação das informações. Para aplicação deste estudo foi selecionado o acervo fotográfico da SANEPAR, criado e gerenciado pelo setor de Patrimônio Histórico da instituição.

Este trabalho tem como objetivos específicos:

- a) compreender o que é uma ontologia;
- b) identificar linguagens e ferramentas de construção e manipulação de ontologias;
- c) fazer um levantamento das necessidades dos usuários em relação á organização e recuperação das informações presentes no objeto de estudo;

- d) elaborar uma ferramenta tecnológica com base nas necessidades informacionais levantadas, apoiada em tecnologias de construção de ontologia;
- e) avaliar os resultados da aplicação dos métodos e ferramentas de ontologia na organização e recuperação das informações presentes no acervo fotográfico da SANEPAR

## 2 LITERATURA PERTINENTE

A revisão bibliográfica se constitui de uma abordagem sobre a organização do conhecimento, passando pelas ontologias, seu processo de construção, linguagens e ferramentas como instrumento de OC.

### 2.1 ORGANIZAÇÃO DO CONHECIMENTO

Organização do conhecimento é definida por Choo (1998) em um primeiro nível como aquela que possui informações e conhecimentos que a tornam bem informada e capaz de percepção e discernimento. Num nível mais profundo, a organização do conhecimento possui informações que lhe conferem uma vantagem, permitindo-lhe agir com inteligência, criatividade e ocasionalmente, com esperteza.

A organização do conhecimento está preparada para sustentar seu crescimento e se desenvolver em um ambiente dinâmico pois com a competência e experiência de seus membros está em constante aprendizado e inovação, sendo capaz de se preparar para a adaptação com antecedência.

No coração da organização do conhecimento está a administração dos processos de informação, que constituem a base para criar significado, construir conhecimento e tomar decisões.

Choo (1998) cita ainda que a organização do conhecimento, cria condições necessárias para o crescimento sustentável no ambiente dinâmico que hoje envolve todas as empresas. Isso se dá em três etapas, ou em três “arenas estratégicas”. Primeiro, a empresa interpreta a informação sobre o ambiente, dá significado ao que acontece à organização e ao que ela faz. Depois, cria novos conhecimentos, combinando a experiência de seus membros, para aprender e inovar. E, enfim, processa e analisa a informação, escolhendo e empreendendo cursos de ação apropriados.

Existem diversas técnicas de organização do conhecimento, uma delas o Vocabulário Controlado (VC). A principal função dos vocabulários controlados é obter consistência na descrição de conteúdo de um objeto através da linguagem para facilitar a recuperação. Em 2005, a NISO apresentou a revisão mais recente de

sua norma para Construção, Formatação e Gestão de Vocabulários Controlados Monolíngües (ANSI/NISO Z39.19-2005) com ênfase na organização terminológica, na interoperabilidade e na manutenção (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005).

Um vocabulário controlado, segundo esta norma, pode ser uma lista de termos, uma taxonomia ou até um extenso tesouro com complexa estrutura hierárquica e diversos tipos de relacionamentos entre os termos, mas sempre será necessário um sistema de gestão, pois a linguagem apresenta conceitos emergentes, mudanças terminológicas e obsolescência de termos. Na medida em que um vocabulário controlado é utilizado torna-se vital conduzir testes periódicos e avaliações para garantir o melhor uso possível (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005).

A norma ANSI Z39.19 -1974 sobre Estrutura, Construção e Uso de Tesouros foi revisada em 1980 e em 1993, sendo nesta última equiparada a norma internacional ISO 2788 e a britânica BS 8733, passando a constituir-se nas Diretrizes para Construção, Formatação e Gestão de Tesouros Monolíngües. Porém com as mudanças ocorridas na área da informação foi necessário expandir o escopo da norma para inclusão de vocabulários controlados, taxonomias e outras formas de linguagens documentárias (FAYEN, 2004).

## **2.2 ONTOLOGIA**

Os problemas relativos à representação da informação e do conhecimento são abordados por estudiosos de diversas áreas, e não são triviais, conforme explica Mendes (1991, p. 2): “capturar o conhecimento humano e torná-lo explícito não é tarefa simples”.

No âmbito da Ciência da Informação, os problemas com representação de conteúdo ocorrem à medida que o volume de documentos a processar e a ordenar cresce, de forma que os usuários não mais se contentam com a organização por grandes classes e exigem informações mais precisas (ROBREDO, 1986). Com o grande volume de material disponível atualmente, o problema se torna crítico. Dessa forma, a preocupação, nessa área, está voltada para a recuperação da informação,

padronizando a terminologia usada pelas pessoas para encontrar e classificar a informação. Utilizam-se técnicas variadas para organização da informação e do conhecimento, uma delas denominada ontologia.

Historicamente, o termo ontologia tem origem no grego “ontos”, ser, e “logos”, palavra. O termo original é a palavra aristotélica “categoria”, que pode ser usada para classificar alguma coisa. Ontologia tem um sentido especial em organização da informação, diferente daquele tradicional adotado na Filosofia.

São diversas as definições apresentadas na literatura e existem contradições (ALMEIDA;BAX, 2003).

Ontologia é uma especificação explícita de uma conceitualização. Em outras palavras, “especificação explícita” está relacionada com a definição dos conceitos, propriedades, relações e axiomas da “conceitualização”, que é uma visão abstrata e simplificada do mundo que queremos representar. (GRUBER, 1993. p.199)

Partindo desta definição, McGuinness (2001) argumenta que existem várias interpretações potenciais: os vocabulários controlados apresentam uma lista de termos e seus significados, logo podem ser vistos como uma especificação de uma conceitualização e interpretados como uma ontologia; os tesouros possuem semântica adicional nas suas relações de termos. Eles fornecem informação de relações de termos sinônimos e especificam os termos usados para representar um domínio, o que pode também ser interpretado como uma ontologia; da mesma forma as hierarquias de categorias encontradas nos sites, como o Yahoo, por exemplo representam o conhecimento a partir de hierarquias formadas a partir de conceitos gerais para os específicos.

Uma ontologia define um vocabulário específico usado para descrever certa realidade, mais um conjunto de decisões explícitas fixando de forma rigorosa o significado pretendido para o vocabulário. Uma ontologia envolve, então, um vocabulário de representação que captura os conceitos e relações em algum domínio e um conjunto de axiomas, que restringem a sua interpretação. (DAVIS *et al.*,1992, p.182)

Borst (1997, p. 12), define de maneira simples e completa “uma ontologia é uma especificação formal e explícita de uma conceitualização compartilhada”. Nessa definição, formal significa legível para computadores; explícita diz respeito a conceitos, propriedades, relações, funções, restrições, axiomas que são

explicitamente definidos; compartilhado é o conhecimento consensual; e, conceitualização diz respeito a um modelo abstrato de algum fenômeno do mundo real.

### 2.3 PROCESSOS DE CONSTRUÇÃO DE ONTOLOGIA

A construção de uma ontologia se faz necessária para

compartilhar a estrutura de informação comum entre pessoas ou softwares, para reutilizar o conhecimento do domínio, para tornar explícitos fatos consensuais, para separar um domínio do conhecimento do conhecimento operacional, para analisar um domínio. (Noy e Guinness, 2001, p. 1).

Os métodos para construção de ontologias têm se desenvolvido por iniciativas de especialistas de diversas áreas. Muitas comunidades científicas utilizam ontologias: a de inteligência artificial (no desenvolvimento de sistemas baseados em conhecimento), a que estuda linguagens naturais (para caracterizar o sentido e o significado das palavras), a de bancos de dados (como esquema conceitual e para buscar inter-operabilidade semântica), a de recuperação da informação (abordagem semântica), e a comunidade de ciência da informação (por exemplo, no estudo de metadados, visando a melhoria da eficiência na recuperação da informação).

A construção e manipulação de ontologias têm sido sistematizadas por metodologias. Baseado nos estudos de Corcho, Fernández-López e Gomez-Pérez (1999), apresentam-se várias metodologias para construção de ontologias. O quadro 1 resume os comentários sobre cada metodologia, proporcionando uma visão geral de seu funcionamento.

Metodologia	Descrição
Cyc	Codifica o conhecimento das fontes e, quando já existe conhecimento suficiente na ontologia, novo consenso é obtido por ferramentas de linguagem natural. (Lenat; Guha, 1990).
Unschold	Identifica o propósito, os conceitos e os relacionamentos entre os conceitos, além dos termos utilizados para codificar a

	ontologia e, em seguida, aceita-la. (Uschold; King, 1996).
Grüninger	Método formal que identifica cenários de uso da ontologia, utiliza linguagem natural para determinar o escopo da ontologia, extrai dados sobre os conceitos, propriedades, relações e axiomas, os quais são formalmente definidos em PROLOG (Grüninger; Fox, 1995).
KACTUS	Método recursivo, que consiste em uma proposta inicial para uma base de conhecimento; quando é necessária uma nova base em domínio similar, generaliza-se a primeira base em uma ontologia, adaptada a ambas aplicações; quanto mais aplicações, mais genérica a ontologia. (Bernaras; Laresgoiti; Corera, 1996 citado por Gomez-Perez, (1999).
Methontology	Constrói-se uma ontologia por reengenharia sobre outra ontologia; a construção utiliza o conhecimento do domínio e as atividades principais são: especificação, conceitualização, formalização, implementação e manutenção (Fernández-López <i>et al.</i> , 1999).
Sensus	Abordagem que deriva ontologias a partir de outras; identificam-se termos relevantes para o domínio, os quais são ligados à ontologia mais abrangente (Sensus, com 50.000 conceitos); os termos relevantes são selecionados e um algoritmo monta a estrutura hierárquica do domínio. (Swartout <i>et al.</i> , 1996).
On-to-knowledge	Auxilia a administração de conceitos em organizações; identifica metas para as ferramentas de gestão do conhecimento e utiliza cenários e contribuições dos provedores / clientes de informação da organização. (Staab <i>et al.</i> , 2001).
CO4 (Cooperative Construction of Consensual knowledge bases)	Permite que usuários discutam sobre o conhecimento introduzido em bases de conhecimento, as quais são compartilhadas; essas bases são ontologias, pois deve haver consenso sobre o conhecimento ali representado. Quando uma mudança é proposta, os usuários são notificados e podem aceita-la ou não (Euzenat, 1996).
(KA)	Modela formas de aquisição do conhecimento usando ontologias desenvolvidas em conjunto por pessoas em diferentes locais, mas que utilizam o mesmo padrão; a comunicação e a coordenação são feitas via agentes inteligentes (Kietz; Maedche; Volz, 2000).
Maedche	Uma ontologia genérica é convertida em modelo; especificam-se textos e obtém-se conceitos do domínio a partir das fontes disponíveis; removem-se conceitos genéricos, de forma que apenas os conceitos específicos do domínio permaneçam; a estrutura conceitual da ontologia está estabelecida e obtém-se as relações.

QUADRO 1 – METODOLOGIAS PARA CONSTRUÇÃO DE ONTOLOGIAS

Fonte: Almeida e Bax (2003)

### 2.3.1 Methontology

Neste trabalho iremos utilizar como metodologia para o desenvolvimento de ontologias a Methontology pela avaliação da ontologia ser feita durante o ciclo de vida do desenvolvimento do processo, proporcionando maior eficiência ao final e segundo Fernandez-Lopez (2003) é a mais madura e clara entre as metodologias disponíveis.

Methontology é uma metodologia de desenvolvimento de ontologias idealizada por um grupo de pesquisa em Engenharia de Ontologias da Universidade Politécnica de Madri, sendo fortemente influenciada por metodologias de Engenharia de Software e de Engenharia do Conhecimento (GÓMEZ-PÉREZ *et al.*,2004).

A Methontology propõe um conjunto de estágios através dos quais o processo de desenvolvimento deve ser conduzido. Segundo Fernandez, Gomez-Perez e Juristo (1997), esses estágios, apresentados na FIG 1, são: especificação, conceitualização, formalização, integração, implementação e manutenção.

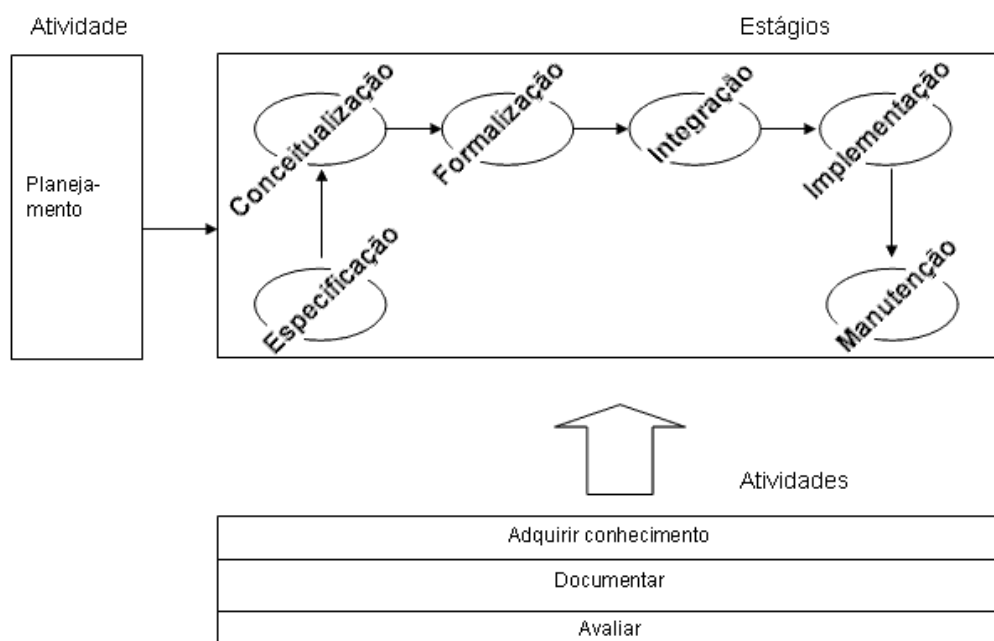


FIGURA 1 – REPRESENTAÇÃO DO CICLO DE VIDA DA ONTOLOGIA

Fonte: Adaptado de Fernandez, Gomez-Perez e Juristo (1997, p.35)

## 2.4 LINGUAGENS DE ONTOLOGIA

Existem várias linguagens que podem ser utilizadas para formalizar as Ontologias. Algumas delas são resumidamente descritas no quadro 2.

Linguagens	Descrição
Cycl	Linguagem formal cuja sintaxe é derivada lógica de primeira ordem. Expressa conhecimento, através de um vocabulário de termos (constantes semânticas, variáveis, números, seqüências de caracteres etc) os quais são combinados em expressões, sentenças e, finalmente, em bases de conhecimento. (Lenat; Guha, 1990)
Flogic (Frame Logic)	Integra frames e lógica de primeira ordem. Trata de uma forma declarativa os aspectos estruturais das linguagens baseadas em frames e orientadas a objeto (identificação de objetos, herança, tipos polimórficos, métodos de consulta, encapsulamento etc). Permite a representação de conceitos, taxonomias, relações binárias, funções, instâncias, axiomas e regras. (Kifer; Lausen; Wu, 1990)
LOOM	Descendente da família KL-ONE (Knowledge Language One) é baseado em lógica descritiva e regras de produção. Permite a representação de conceitos, taxonomias, relações binárias, funções, axiomas e regras de produção. (Brill, 1993)
CARIN	Trata-se de uma combinação da Datalog (linguagem baseada em regras) e lógica descritiva ALN. Uma ontologia CARIN é construída por dois componentes terminológicos: um conjunto de conceitos com declarações de inclusão e um conjunto de regras que usam os conceitos. (Levy; Rousset, 1996)
GRAIL	É uma linguagem que especifica uma ontologia do domínio médico (GALEN)
Ontolíngua	Combina paradigmas das linguagens baseadas em frames e lógica de primeira ordem. Permite a representação de conceitos, taxonomias de conceitos, relações binárias, funções, axiomas, instâncias e procedimentos. Sua alta expressividade causa problemas na construção de mecanismos de inferência. (Chaudhri <i>et al.</i> , 1998)
OCML	Permite a especificação de funções, relações e classes, instâncias e regras. Utilizada em aplicações de gerenciamento do conhecimento, desenvolvimento de ontologias, comércio eletrônico e sistemas baseados em conhecimento. Aplicada em medicina, ciências sociais, memória corporativa, engenharia, portais da Web etc. (Domingue; Motta; Corcho, 1999; Chaudhri; KARP; Thomere, 1999)
OML (Ontology Markup Languages)	É uma linguagem de especificação de ontologias baseada em lógica descritiva e grafos conceituais. Permite a representação

	de conceitos organizados em taxonomias, relações e axiomas em lógica de primeira ordem. (Kent, 1999)
RDF(Resource Description Framework) / RDFS (RDF Schema)	Desenvolvidos pelo W3 Consortium, tem por objetivo a representação de conhecimento através da idéia de redes semânticas. São linguagens não muito expressivas, permitindo apenas a representação de conceitos, taxonomias de conceitos e relações binárias. (Lassila; Swick, 1999)
NKRL	Desenvolvida para descrever modelos semânticos para documentos multimídia. Baseada em frames, permite definir conceitos organizados em taxonomias e relacionamentos entre eles. Possui uma sub-linguagem para a construção de argumentos complexos. (Bertino; Barbara; Zarri, 1999)
SHOE (Simple HTML Ontology Extensions)	Utiliza extensões ao HTML, adicionando marcações para inserir dados semânticos em páginas Web. As marcações podem ser: para a construção de ontologias e para anotações em documentos da Web. Permite representar conceitos, taxonomias, relações, instâncias e regras de dedução. (Heflin; Hendler, 2000)
XOL	Trata-se de um pequeno subconjunto de primitivas do protocolo OKBC. É uma linguagem restrita, na qual apenas conceitos, taxonomias e relações binárias podem ser especificadas. Não possui mecanismos de inferência e foi projetada para o intercâmbio de ontologias no domínio da biomédica. (Karp, 1997)
OIL	É o precursor do DAML+OIL e a base para o projeto de linguagem para a Web Semântica. Combina várias primitivas de modelagem das linguagens baseadas em frames com a semântica formal e serviços de inferência da lógica descritiva. Pode verificar classificação e taxonomias de conceitos. (Fensel <i>et al.</i> , 2001)
DAML + OIL	DAML + OIL é uma linguagem de marcação semântica para a Web que apresenta extensões a linguagens como o DAML (DARPA Agent Markup Language), RDF e RDFS, através de primitivas de modelagem baseadas em linguagens lógicas. Permite representar conceitos, taxonomias, relações binárias e instâncias. (Horrocks <i>et al.</i> , 2001)

QUADRO 2 — LINGUAGENS PARA CONSTRUÇÃO DE ONTOLOGIAS

Fonte: Almeida e Bax (2003)

Complementando o quadro 2 existe uma linguagem para construção de ontologias denominada OWL, a qual será abordada na sessão 2.4.1.

### 2.4.1 Linguagem OWL

A linguagem no qual será mantido foco neste trabalho é a OWL devido ao fato de ser uma recomendação do W3C e porque segundo McGuinness e Harmelen (2004), OWL está relacionada com o contexto da Web e Web Semântica.

De acordo com McGuinness e Harmelen (2004), a OWL foi projetada para ser usada com aplicações que necessitam de processamento de conteúdos informacionais ao invés de apenas apresentar informações para o homem e para satisfazer a necessidade de uma Linguagem Ontológica para Web. OWL proporciona maior interoperabilidade de máquina do que conteúdos web construídos com XML, RDF, e RDF Schema (RDF-S) por fornecer vocabulário adicional junto com uma semântica formal. OWL possui três expressivas sub-linguagens: OWL Lite, OWL DL e OWL Full e faz parte de um conjunto de recomendações W3C relacionadas com a Web Semântica.

Neste trabalho optou-se por utilizar a linguagem OWL pois dá suporte aos usuários que desejam o máximo de expressividade sem perder a completude computacional e a capacidade de decisão dos sistemas de raciocínio. Dentro da linguagem OWL foi selecionada a versão Lite por se tratar de uma ontologia de classificação hierárquica e restrições simples. Por exemplo, embora suporte restrições de cardinalidade, ela só permite valores de cardinalidade 0 ou 1. É mais simples fornecer ferramentas que suportem OWL Lite que seus parentes mais expressivos, e ela também permite um caminho de migração mais rápido de tesouros e outras taxonomias. OWL Lite também tem uma menor complexidade formal que OWL DL..

Resource Description Framework (RDF) é também uma recomendação do W3C e segundo Klyne e Carroll (2004) é uma estrutura capaz de representar as informações na web que tem uma sintaxe abstrata que reflete um modelo de dados baseado em grafo com uma semântica formal e rigorosa definição de notação provendo a base para deduções em dados RDF.

Segundo Lima e Schwabe (2004), RDF tem a finalidade de proporcionar um modo de afirmar coisas sobre recursos da web ou do mundo real, restringindo a descrição dos recursos a triplas (sujeito, predicado, objeto). O sujeito é sobre quem

será feito à afirmação, predicado é a propriedade do recurso que está sendo descrito e objeto é o valor que será atribuído à propriedade do recurso.

Ainda citando Lima e Schwabe (2004), para tornar as afirmações sobre os recursos processáveis por máquina são necessários identificadores para cada elemento da afirmação de modo a impedir qualquer possibilidade de confusão com outros similares. Segundo estes autores, um documento RDF é uma lista de descrições onde cada descrição corresponde a um recurso e contém uma lista de propriedades.

Segundo Bray *et al.*, (2006), em um documento XML existe a possibilidade de existir o uso de marcadores (vocabulários) iguais, mas que possuem estruturas diferentes e quando ocorre uma situação como esta, dá-se o nome de colisão. Para evitá-la, utiliza-se um recurso chamado namespace que é um marcador XML composto de um nome e um endereço (URI) que define a localização e um apelido para este namespace dentro do documento onde a declaração namespace foi utilizada.

## **2.5 APLICATIVOS PARA CONSTRUÇÃO DE ONTOLOGIAS**

A formalização de ontologias para um conjunto de agentes não é uma tarefa trivial, já que significa tornar explícito algo que normalmente é implícito. Os aplicativos de edição podem simplificar consideravelmente o processo de construção de ontologias, desde o início ou a partir de outras já existentes. Geralmente, estes aplicativos incluem documentação, importação e exportação de ontologias existentes (de diferentes formatos), visualização gráfica, bibliotecas e mecanismos de inferência.

Em função da diversidade de linguagens de construção de ontologias existentes, muitos aplicativos também foram propostos. Praticamente todas as linguagens possuem ao menos um aplicativo para apoiar a construção de ontologias. Entre eles estão:

Ferramentas	Descrição
CODE4 (Conceptually Oriented Description Environment)	Aplicativo de propósito geral. Possui diferentes modos de herança e inferência, interface gráfica de fácil uso, modo de hipertexto para navegação e utilitários para leitura de documentos e gerenciamento léxico. (Skuce, 1995)
VOID	Ambiente para navegação, edição e gerenciamento de ontologias. Através de simulações, possibilita o estudo de questões teóricas como: organização de bibliotecas de ontologias e tradução entre diferentes formalismos. (Schreiber; Terpstra; Sisyphus, 1995)
IKARUS (Intelligent Knowledge Acquisition and Retrieval Universal System)	Explora as capacidades cooperativas do ambiente Web. Utiliza uma representação hierárquica gráfica que permite herança múltipla. As declarações que contém a informação são representadas como predicados, com sintaxe e semântica definidos, ou como fragmentos sem estrutura. (Skuce, 1996)
Ontolíngua	Conjunto de serviços que possibilitam a construção de ontologias compartilhadas entre grupos. Permite acesso a uma biblioteca de ontologias, tradutores para linguagens e um editor para criar e navegar pela ontologia. Editores remotos podem editar ontologias usando protocolos. (Farquhar; Fikes; Rice, 1996)
Ontosaurus	Consiste de dois módulos: um servidor de ontologias, que usa o LOOM para representação do conhecimento, e um servidor de navegação por ontologias, que cria páginas HTML dinamicamente e apresenta a hierarquia da ontologia. (Swartout <i>et al.</i> , 1996)
GKB-Editor (Generic Knowledge Base Editor)	Aplicativo para navegação e edição de ontologias através de sistemas de representação baseados em frame's. Oferece interface gráfica, onde os usuários podem editar diretamente a base de conhecimento e selecionar a parte que é de seu interesse. (Paley; Karp, 1997)
JOE (Java Ontology Editor)	Aplicativo para construção e visualização de ontologias. Proporciona gerenciamento do conhecimento em ambientes abertos, heterogêneos e com diversos usuários. As ontologias são visualizadas como um diagrama entidade-relacionamento, como o gerenciador de arquivos do MS Windows ou como uma estrutura em árvore. (Mahalingam; Huhns, 1997)
OilEd	É um editor de ontologias de código aberto; permite construir ontologias utilizando a linguagem OIL. Não é um ambiente completo para desenvolvimento de ontologias. Verificação da consistência e classificação automática da ontologia podem ser executadas pela ferramenta FaCT (Horrocks; Sattler; Tobies, 1999)
OntoEdit	É um ambiente gráfico para edição de ontologias; permite inspeção, navegação, codificação e alteração de ontologias. O modelo conceitual é armazenado usando um modelo de

	ontologia que pode ser mapeado em diferentes linguagens de representação. As ontologias são armazenadas em bancos relacionais e podem ser implementadas em XML, Flogic, RDF(S) e DAML+OIL. (Maedche <i>et al.</i> ,2000)
Protégé	É um ambiente interativo para projeto de ontologias, de código aberto, que oferece uma interface gráfica para edição de ontologias e uma arquitetura para a criação de ferramentas baseadas em conhecimento. A arquitetura é modulada e permite a inserção de novos recursos. (Noy; Ferguson; Musen, 2000)
WebODE	Ambiente para engenharia ontológica que fornece suporte à maioria das atividades de desenvolvimento de ontologias. Fornece uma API (application programming interface) que facilita a integração com outros sistemas, importa e exporta ontologias de linguagens de marcação. (Arpírez <i>et al.</i> ,2001)
WebOnto	Possibilita a navegação, criação e edição de ontologias, representadas na linguagem de modelagem OCML. Permite o gerenciamento de ontologias por interface gráfica, inspeção de elementos, verificação da consistência da herança e trabalho cooperativo. Possui uma biblioteca com mais de cem ontologias. (Domingue, 2001)

QUADRO 3 – FERRAMENTAS PARA CONSTRUÇÃO DE ONTOLOGIAS

Fonte: Almeida e Bax (2003)

### 2.5.1 Ontowiki

OntoWiki é um aplicativo de apoio a engenharia do conhecimento para ambiente web. Chama-se OntoWiki porque possui características dos sistemas Wiki existentes. No entanto, o projeto é independente e completa as tecnologias convencionais Wiki. OntoWiki tem como abordagem simplificar a apresentação e construção de bases de conhecimento (AUER; DIETZOLD; RIECHERT, 2006).

OntoWiki é uma aplicação web de código aberto escrita em PHP com a finalidade de criar e manter bases de conhecimento.

Sua arquitetura apresenta a possibilidade de construir ontologias utilizando diversos aplicativos como Protege ou Powl (Semantic Web Application Development Platform) e depois importar para o OntoWiki.

Uma das principais funções do OntoWiki é de promover a interação social para desenvolver bases de conhecimento. Isto facilita o intercâmbio de metadados

sobre as bases de conhecimento e promove a colaboração em cenários onde a comunicação pessoal é difícil (AUER; DIETZOLD; RIECHERT, 2006)

OntoWiki foi implementado como uma alternativa de interface do usuário para o editor Powl, uma plataforma de desenvolvimento de aplicações da Web Semântica.

### 2.5.2 Protégé

O projeto Protégé, iniciou em 1987 com Mark Musen criando uma metaferramenta para sistemas baseados em conhecimentos na área médica na Universidade de Stanford. A ferramenta original era uma aplicação que tinha como intenção a construção de ferramentas de aquisição de conhecimento para alguns programas especializados em planejamento médico.

Dentre os vários aplicativos de desenvolvimento de ontologias existentes, optou-se por adotar neste trabalho o Protégé, pelas seguintes razões: é compatível com os padrões da Web Semântica como a linguagem OWL; está disponível para inúmeras plataformas como Windows, UNIX, Mac, Solaris, etc.; além disso, é uma ferramenta open source. Protege é um ambiente de desenvolvimento de ontologias com funcionalidade para edição de classes, instâncias e propriedades. A versão do □Protege, utilizado neste trabalho é a 3.4, uma versão mais antiga da ferramenta porem com uma disponibilidade maior de material de apoio.

O Protégé foi desenvolvido utilizando a linguagem de programação Java e, posteriormente, optou por abrir seu código. Surgiu então uma arquitetura integrável a diversas aplicações, via componentes que podem ser conectados ao sistema. Como consequência desta decisão e de sua difusão, componentes de vários matizes, elaborados por grupos de pesquisa de usuários, puderam ser adicionados ao sistema, sem necessitar o redesenvolvimento. Foram aproveitados, por exemplo, o Jambalaya, um utilitário com animação e vários outros recursos em visualização de dados, e o Ontoviz, um componente que faz com que o gerador de gráficos Graphviz da AT&T produza gráficos com instâncias, heranças e outros tipos de relacionamento.(Freitas, 2004)

As formas de contribuição são:

- a) backends, permitem usuários armazenar e importar bases de conhecimento em vários formatos;
- b) slot widgets, são usados para mostrar e editar valores de slot ou suas combinações em modos de domínio específico e tarefa específica;
- c) tab plug-ins, são aplicativos baseados no conhecimento usualmente fortemente ligado a bases de conhecimento do Protégé.

A arquitetura do Protégé consiste na interface do usuário definida pelos plugins e slots que fazem a interface frente à base de conhecimento definidas pelo Protégé, representada pelo núcleo do Protege, que faz o intermédio com a base de armazenamento persistente podem ser em formato de banco de dados ou arquivos.

## **2.6 WEB SEMÂNTICA**

A organização da imensa vastidão de conteúdo disponível atualmente na Internet, de uma forma simples, eficiente e focada em nossas necessidades, se tornou um problema. Dessa necessidade surgiu a Web Semântica em 2001, quando Tim Berners-Lee, James Hendler e Ora Lassila publicaram um artigo na revista Scientific American, intitulado: “Web Semântica: um novo formato de conteúdo para a Web que tem significado para computadores vai iniciar uma revolução de novas possibilidades.” Segundo Lee, a Web Semântica estrutura o conteúdo significativo das páginas Web, criando um ambiente onde agentes de software percorrem página por página para executarem tarefas solicitadas pelos usuários. O desenvolvimento da Web Semântica envolve duas importantes tecnologias: XML e RDF.

Sua principal aplicação se refere à capacidade de os sistemas computacionais interpretarem o conteúdo disponível nos sites da Internet e conseguir entender de forma diferenciada uma página em que a palavra bala é um doce ou é um projétil de armas, por exemplo. Ou seja, o conteúdo é interpretado de acordo com seu contexto (Clicko, 2008).

A Web Semântica representa a evolução da Web atual. Enquanto a Web tradicional foi desenvolvida para ser compreendida somente pelos usuários, a Web Semântica está sendo projetada para ser entendida também pelas máquinas. Este entendimento se dá na forma de agentes computacionais, que são capazes de

operar eficientemente sobre as informações, podendo entender seus significados. Portanto, auxiliando os usuários em operações na Web (Dziekaniak *et al.*, 2004).

A Web Semântica tem como objetivo incorporar semântica às informações. Com isso, não somente os usuários entenderão as informações como também as máquinas. Ela pretende fornecer estruturas e dar significado semântico ao conteúdo das páginas Web, criando assim um ambiente onde agentes de software possam trabalhar em conjunto com o usuário (Dziekaniak *et al.*, 2004).

Segundo o mesmo autor, “a Web Semântica é sobre a colocação de arquivos de dados na Web. Não é apenas uma Web de documentos, mas também de dados. A tecnologia de dados da Web Semântica terá muitas aplicações, todas interconectadas. Pela primeira vez haverá um formato comum de dados para todos os aplicativos, permitindo que os bancos de dados e as páginas da Web troquem arquivos”.

Segundo Berners-Lee *et al.*, (2001), grande parte do conteúdo da Web atual é projetada para o entendimento humano apenas, não para programas de computador manipulá-los significativamente, pois as máquinas ainda não possuem um método seguro para processar a semântica. De acordo com ele, o objetivo da Web Semântica é trazer estrutura para o conteúdo das páginas Web, aumentando-lhes o significado, criando um ambiente onde agentes de software podem fluir de página a página, cumprindo sofisticadas tarefas para os seus usuários.

Ainda de acordo com Berners-Lee *et al.*, (2001), e citado em Ramalho (2006), o primeiro passo para o desenvolvimento da Web Semântica é a inclusão de dados em um formato que os sistemas computacionais possam naturalmente compreender de forma direta ou indireta, tendo como principal objetivo possibilitar um melhor aproveitamento das potencialidades do ambiente Web, onde por meio do uso intensivo de linguagens computacionais e instrumentos de metadados espera-se obter o acesso automatizado às informações de maneira mais precisa, utilizando-se para isso processamentos semânticos de dados e heurísticas automáticas.

Assim, pode-se verificar que a Web Semântica é uma tentativa inversa de solução, comparando-se com as tradicionalmente desenvolvidas, onde a idéia é pensar nas máquinas para que estas possam servir aos humanos de maneira mais eficiente (Ramalho, 2006).

As tecnologias mais importantes neste ambiente são a linguagem de marcação eXtensible Markup Language (XML) e a abordagem de representação de

dados Resource Description Framework (RDF). Estão em aperfeiçoamento também linguagens para representação da semântica de dados entre elas DAML+OIL e OWL explicitando restrições sobre a semântica do mundo real.

### 3 METODOLOGIA

Esta pesquisa foi realizada através de uma abordagem científica qualitativa, visando analisar uma situação, a possibilidade de otimização de busca se utilizando de ontologias. Na pesquisa qualitativa trata-se de “uma metodologia geral para desenvolver teoria que está inserida em dados sistematicamente coletados e analisados. A teoria surge durante a própria pesquisa e isso ocorre através da interação contínua entre a coleta e a análise de dados” (STRAUSS; CORBIN, 1994. p.10).

O trabalho caracteriza-se também como sendo uma pesquisa descritiva exploratória, pelo fato de envolver um levantamento bibliográfico na área de ontologias, aplicação de questionários com os usuários do objeto de estudo na SANEPAR e análise de exemplos que estimulem a compreensão da importância do uso das técnicas de ontologias.

Segundo Gil (2002), uma pesquisa exploratória tem como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito. Pode envolver levantamento bibliográfico, entrevistas com pessoas experientes no problema pesquisado. Geralmente, assume a forma de pesquisa bibliográfica e estudo de caso.

Como já citado no item 2.3, a metodologia aplicada a este projeto foi baseada na metodologia de desenvolvimento de ontologias Methontology. As etapas presentes neste tipo de metodologia são as seguintes:

- a) **Planejamento** – identificação das dos métodos, linguagem e aplicativos utilizados na construção da ontologia, descritos no referencial teórico deste trabalho.
- b) **Especificação** – desenvolvimento da ontologia com intuito de avaliar a possibilidade de se utilizar ontologias como ferramenta para redução das dificuldades relacionadas à organização e a recuperação das informações no Acervo Histórico Fotográfico da SANEPAR, o qual não possui qualquer tipo de organização e classificação que possibilite uma recuperação informacional eficaz.
- c) **Aquisição de Conhecimento** – desenvolvimento de uma entrevista estruturada realizada com dois funcionários e um estagiário do Setor de Patrimônio História da Companhia (APENDICE A). As entrevistas foram conduzidas, nos mês de outubro com objetivo de se avaliar o cenário, identificar cargos e funções,

ferramentas e documentos utilizados, as barreiras encontradas pelos funcionários na indexação e recuperação das informações do acervo e também apontar possíveis soluções para estas barreiras, sugeridas pelos próprios usuários. Cada entrevista teve a duração de cerca de trinta minutos, foi gravada e transcrita em um editor de texto.

- d) **Conceitualização** – construção de um modelo conceitual contendo as classes e subclasses desenvolvidas na ontologia bem como seus atributos.
- e) **Formalização** – transformação do modelo conceitual em um modelo formal utilizando o aplicativo Protege 3.4.
- f) **Integração** – integração com o máximo possível as ontologias existentes à nova ontologia.
- g) **Implementação** – implementação da ontologia através da linguagem OWL Lite tornando-se computável.
- h) **Avaliação** – avaliação da ontologia através da análise dos resultados e posteriormente junto aos usuários do setor de Patrimônio Histórico da SANEPAR
- i) **Documentação** – envolve todo este planejamento (etapas da metodologia) e a especificação dos artefatos, sendo assim a documentação da ontologia será desenvolvida no decorrer de todo o trabalho; essa documentação contribui para se obter uma melhor clareza da ontologia, assim como facilita a sua manutenção, uso e reuso.

A etapa de Manutenção prevista na Methontology não fez parte do escopo deste projeto, porém é uma etapa que pode ser realizada após a implementação do produto caso a empresa apresente interesse em efetivar este serviço. Com base nestas etapas foi criada a ontologia proposta neste trabalho.

#### 4 ANÁLISE DOS DADOS E RESULTADOS

Segundo Noy e McGuinness (2003), em termos práticos, o desenvolvimento de uma ontologia consiste em: definir classes hierarquicamente (classes-subclasses); definir propriedades e descrever os seus valores permitidos; e preencher os valores para propriedades e instâncias. As classes representam os conceitos do domínio e as propriedades são características destes conceitos; os relacionamentos representam as relações que existem entre os conceitos do domínio.

A literatura existente sobre ontologias recomenda enumerar os termos principais do domínio produzindo uma lista de termos (termo e significado) mais relevantes. Com base nisto foi desenvolvida uma lista de termos utilizados no domínio de imagens do acervo utilizando-se de alguns termos e significados já preestabelecidos pelo setor de Patrimônio Histórico da SANEPAR levantados na entrevista realizada e outros criados pelo autor, conforme o quadro 4 abaixo:

Autor	Nome do Fotógrafo
Conservação	Estado de conservação em que se encontra a fotografia
Cromia	Coloração da fotografia
Formato	Formato em que se encontra a fotografia
Meio	Meio em que se encontra a fotografia
Tema	Assunto da fotografia
Título	Título dado a fotografia
Local	Local em que a fotografia foi produzida
Dimensão	Tamanho da fotografia
Data	Data de produção da fotografia
Descrição	Descrição textual das informações presentes na imagem
Personagens	Pessoas que fazem parte da fotografia
Vínculo Profissional	Profissão exercida pelos personagens presentes na foto na época de produção da mesma

QUADRO 4 – LISTA DE TERMOS.

Fonte: O Autor (2010).

Com auxílio desta lista de termos, desenvolveu-se a modelagem da ontologia do acervo fotográfico através de um diagrama de classe, conforme a figura abaixo:

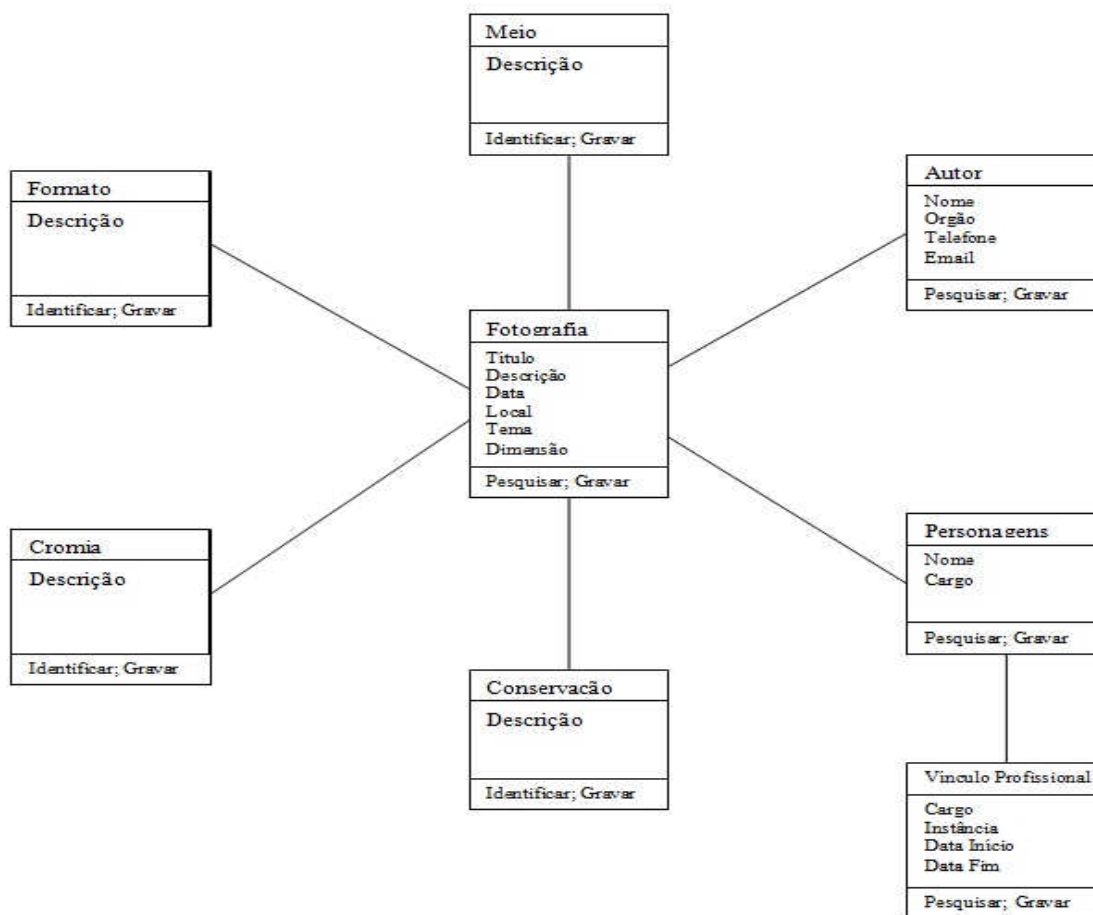


FIGURA 2 – DIAGRAMA DE CLASSES

Fonte: O Autor (2010).

Observa-se no diagrama a criação de Fotografia como classe principal e Meio, Formato, Cromia, Conservação, Personagens e Autor como sub-classes de Fotografia. Foi criada ainda uma sub-classe denominada Vínculo Profissional com a intenção de relacionar as informações dos personagens identificados nas fotografias com o cargo e instancia ocupados na época da criação da imagem.

A partir da especificação deste diagrama deu-se início á transformação do modelo conceitual em um modelo formal utilizando o aplicativo Protege 3.4.

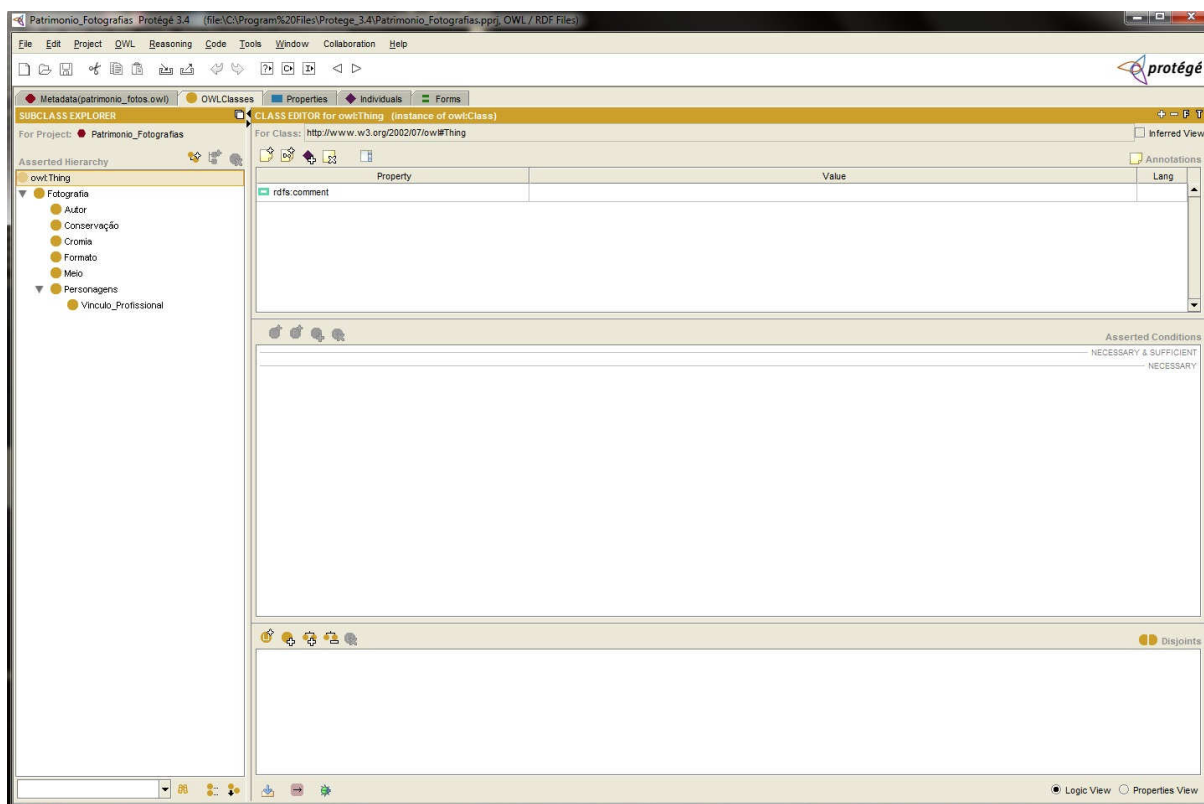


FIGURA 3 – CLASSES

Fonte: O Autor (2010).

A figura 3 apresenta a construção das classes e subclasses da ontologia já na interface do aplicativo Protege 3.4, implementada através da linguagem OWL Lite. Assim como no diagrama de classes, observa-se Fotografia como classe principal e suas subclasses relacionadas hierarquicamente.

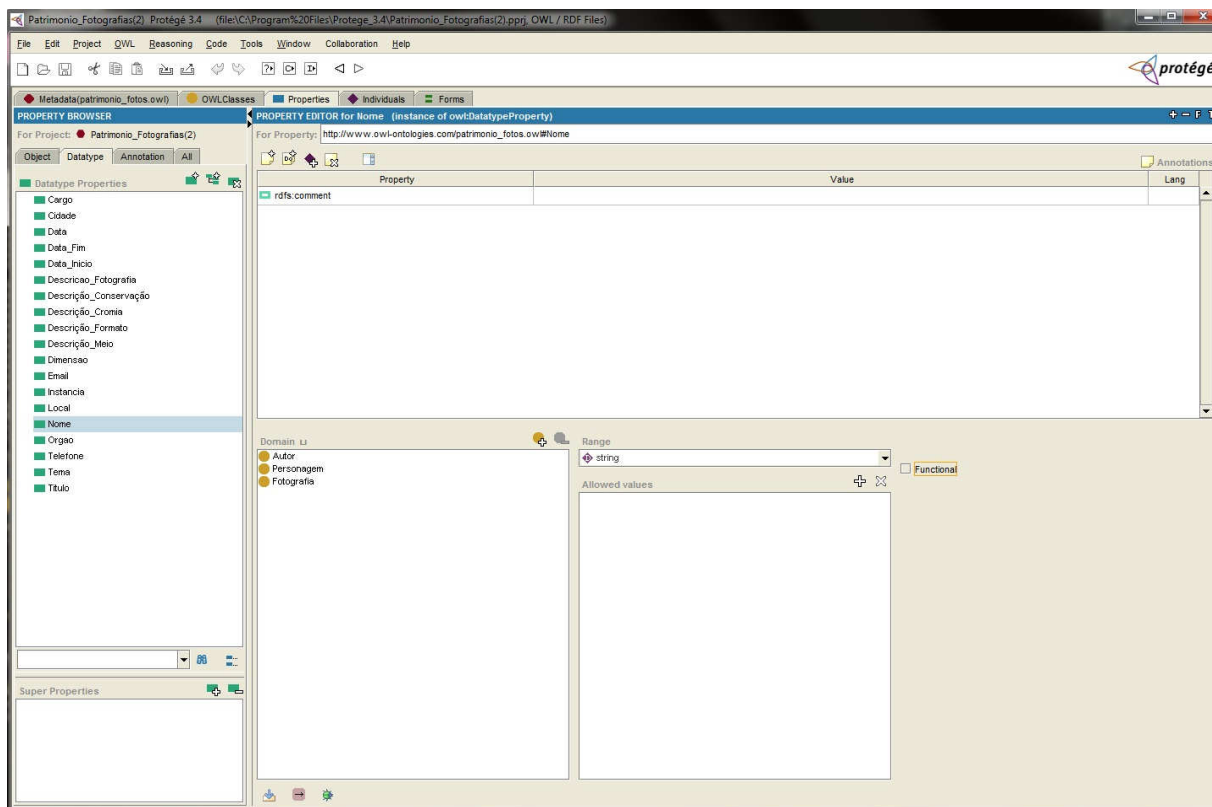


FIGURA 4 – ATRIBUTOS

Fonte: O Autor (2010).

Na figura 4 observa-se a construção dos atributos bem como as classes a que se relacionam e seu tipo de cadastro. O exemplo utilizado na figura demonstra o atributo “Nome” o qual se relaciona com a subclasse “Autor” e “Personagem” que estão relacionados diretamente com “Fotografia” por ser a classe principal. Este atributo será cadastrado com caracteres tipo string, pois permite valores alfanuméricos.

Depois de relacionadas às classes, subclasses e seus devidos atributos, iniciou-se o processo de definição das instâncias, ou seja, o cadastro dos valores na base de conhecimento. Foram cadastradas cinco fotografias presentes no Acervo Fotográfico da SANEPAR, porém como forma de exemplificação deste processo, foi selecionada uma fotografia aleatória dentre as que foram utilizadas para popular a base deste trabalho.



FIGURA 5 – ASSINATURA SANEPAR

Fonte: Acervo Fotográfico da SANEPAR

<b>Fotografia 01</b>	
Tema	Assinaturas
Título	Convênio com farmácia para recebimento de contas de água
Local	Sede Sanepar – Curitiba/PR
Data	17/06/1998
Descrição	Stênio Salles Jacob (Presidente Sanepar), juntamente com Venício Bley Filho (Diretoria Sanepar), assinam convênio com rede de farmácia, para recebimento das contas da Sanepar.
Dimensão	18x12
Cromia	Colorida
Conservação	Ótima
Formato	Impresso
Meio	Papel Fotográfico
Autor	João Henrique
Órgão	Sanepar
Telefone	32057889
Email	jh@sanepar.com

QUADRO 5 – DESCRIÇÃO FOTOGRÁFICA

Fonte: O Autor (2010).

Observam-se nas figuras 6 a 13 abaixo que a aba de instância para cadastro referente à classe “Fotografia” apresentam todos os valores de atributos, já que, como citado acima, se trata da classe principal da ontologia. Já as abas das subclasses apresentam apenas os atributos a elas relacionados. Esta característica de associação de todos os atributos a classe principal ou o relacionamento de atributos apenas com suas respectivas subclasses, possibilita uma recuperação de informações mais precisa, uma vez que se pode buscar uma fotografia específica através de seu título, todas as fotografias de um determinado tema, fotografias com o mesmo estado de conservação ou então somente fotografias de um determinado autor.

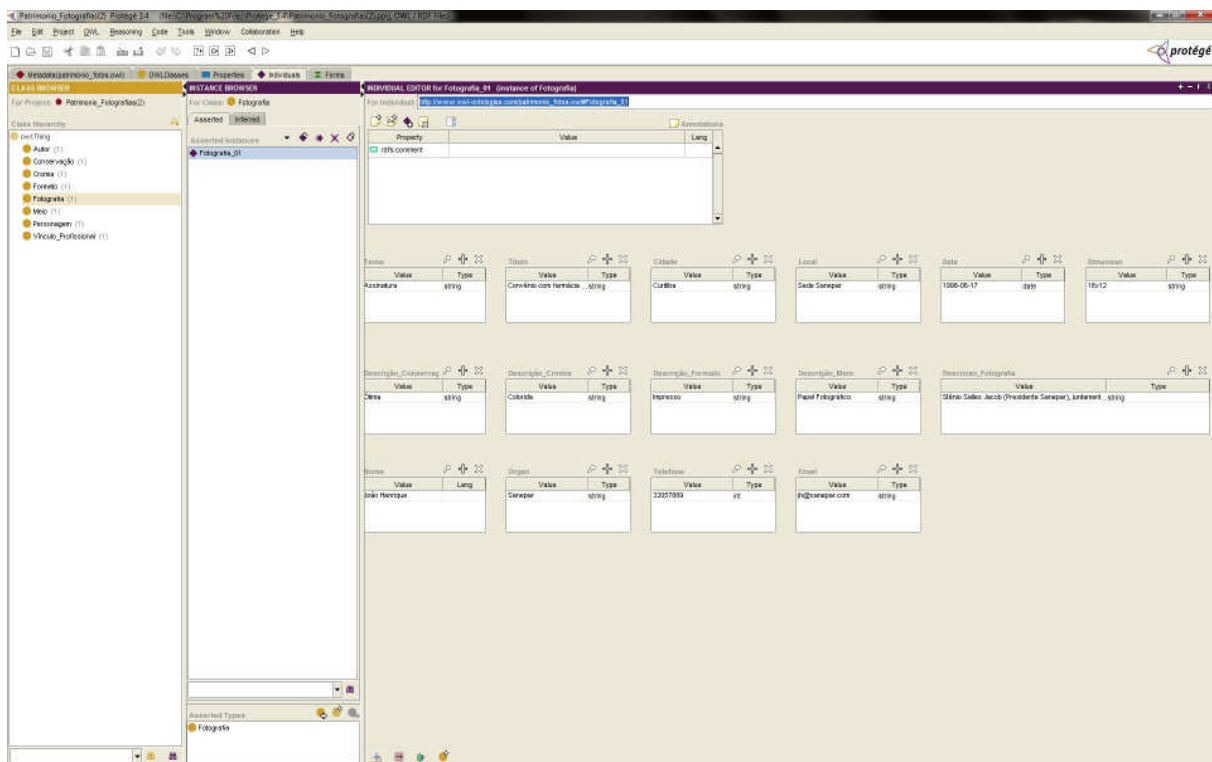


FIGURA 6 – VALORES “FOTOGRAFIA”

Fonte: O Autor (2010).

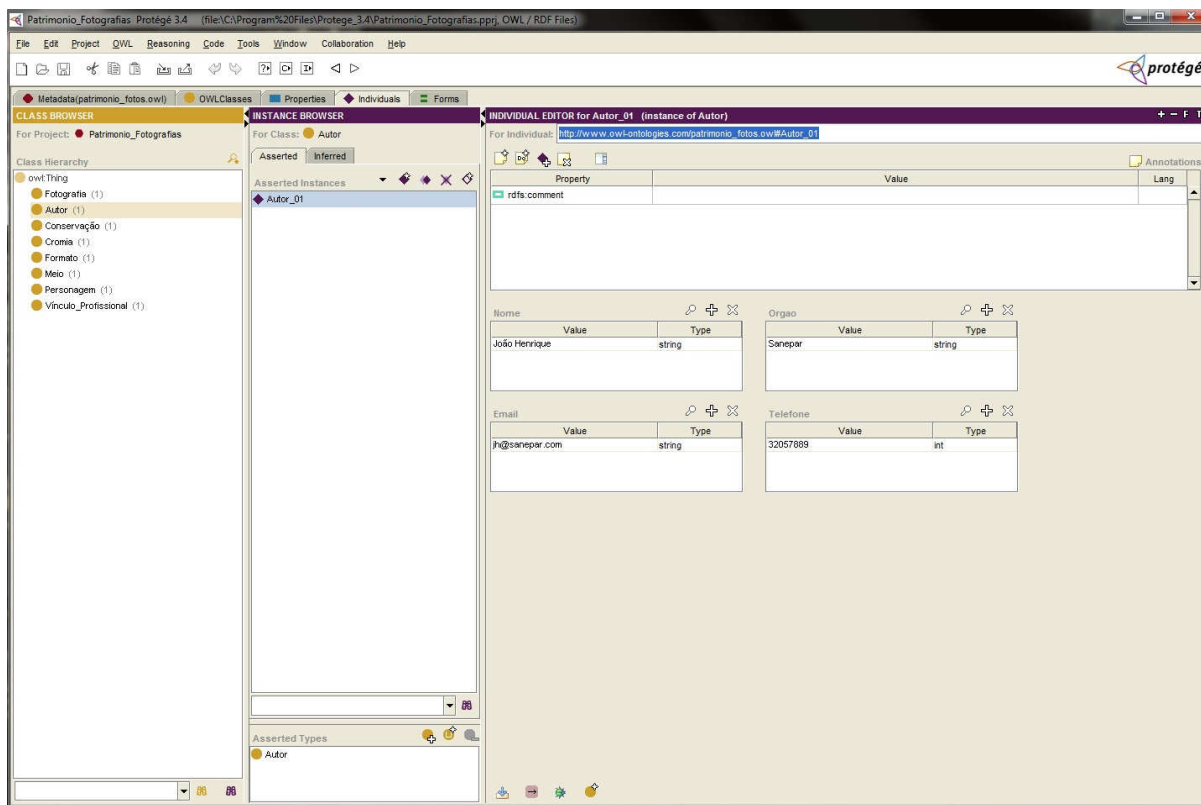


FIGURA 7 – VALORES “AUTOR”

Fonte: O Autor (2010).

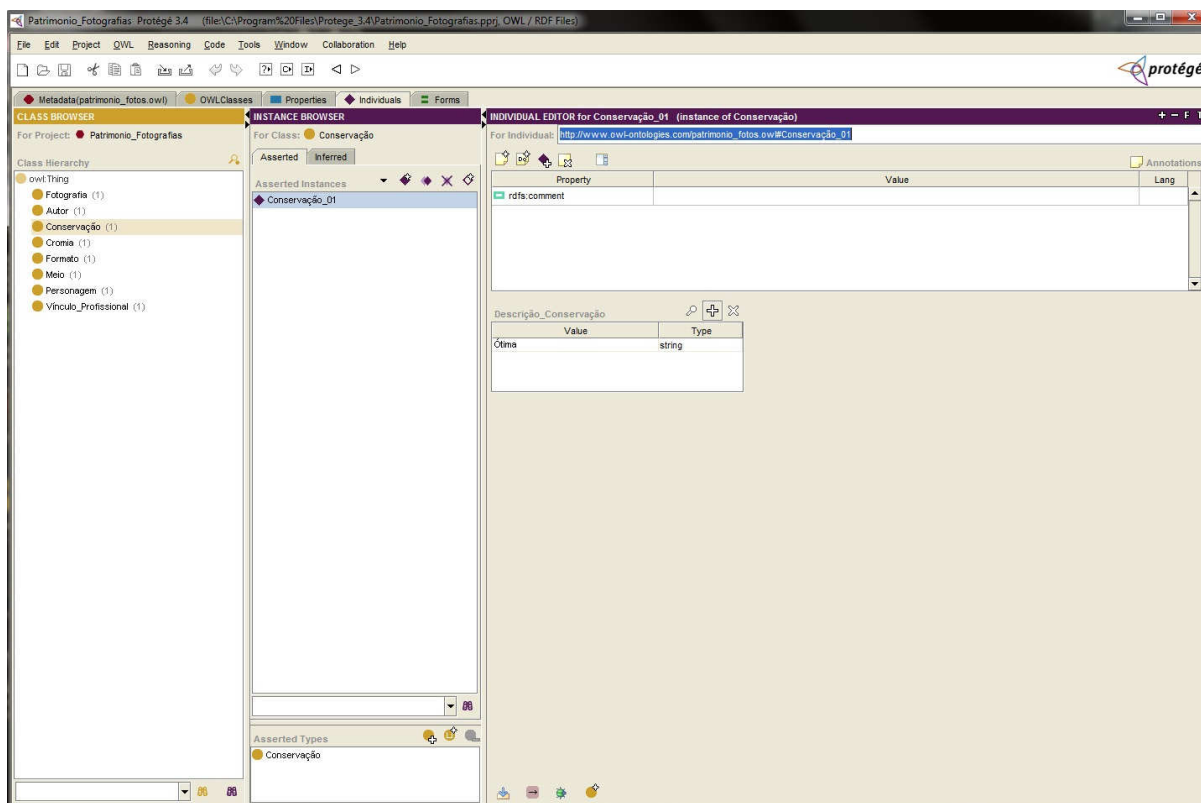


FIGURA 8 – VALORES “CONSERVAÇÃO”

Fonte: O Autor (2010).

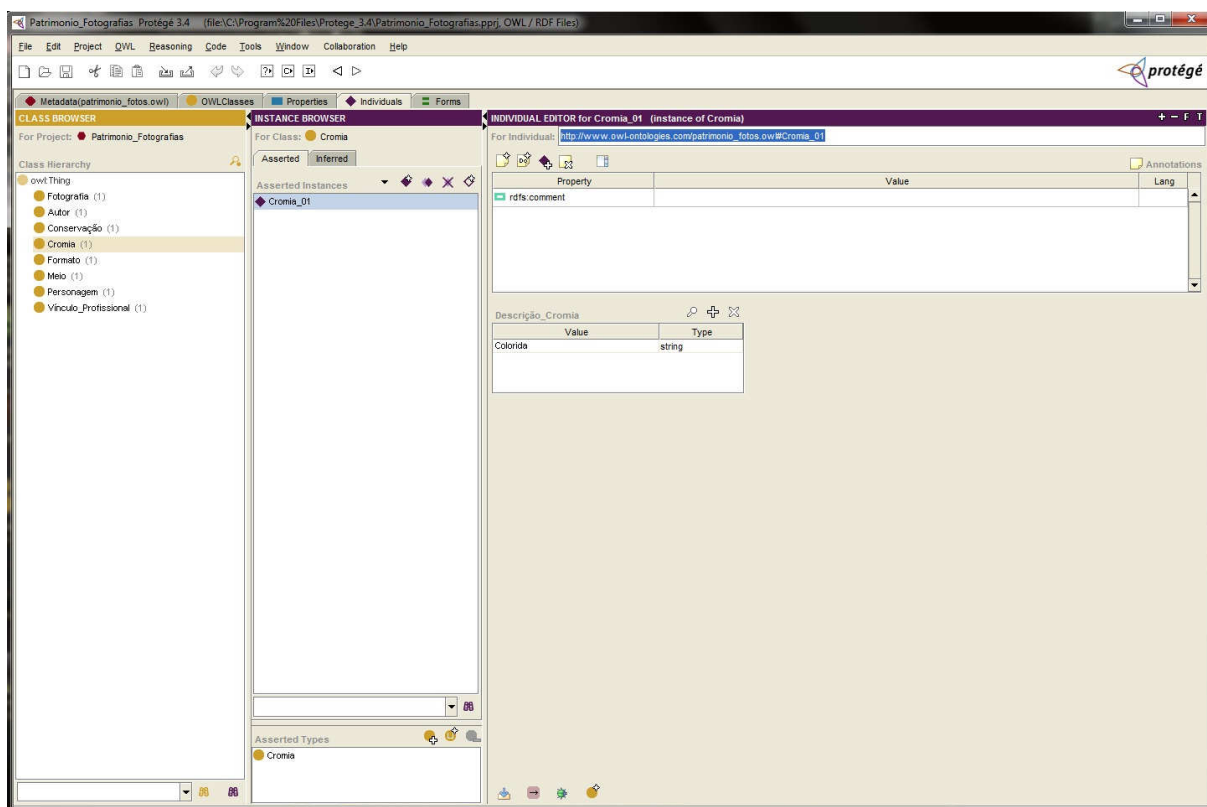


FIGURA 9 – VALORES “CROMIA”

Fonte: O Autor (2010).

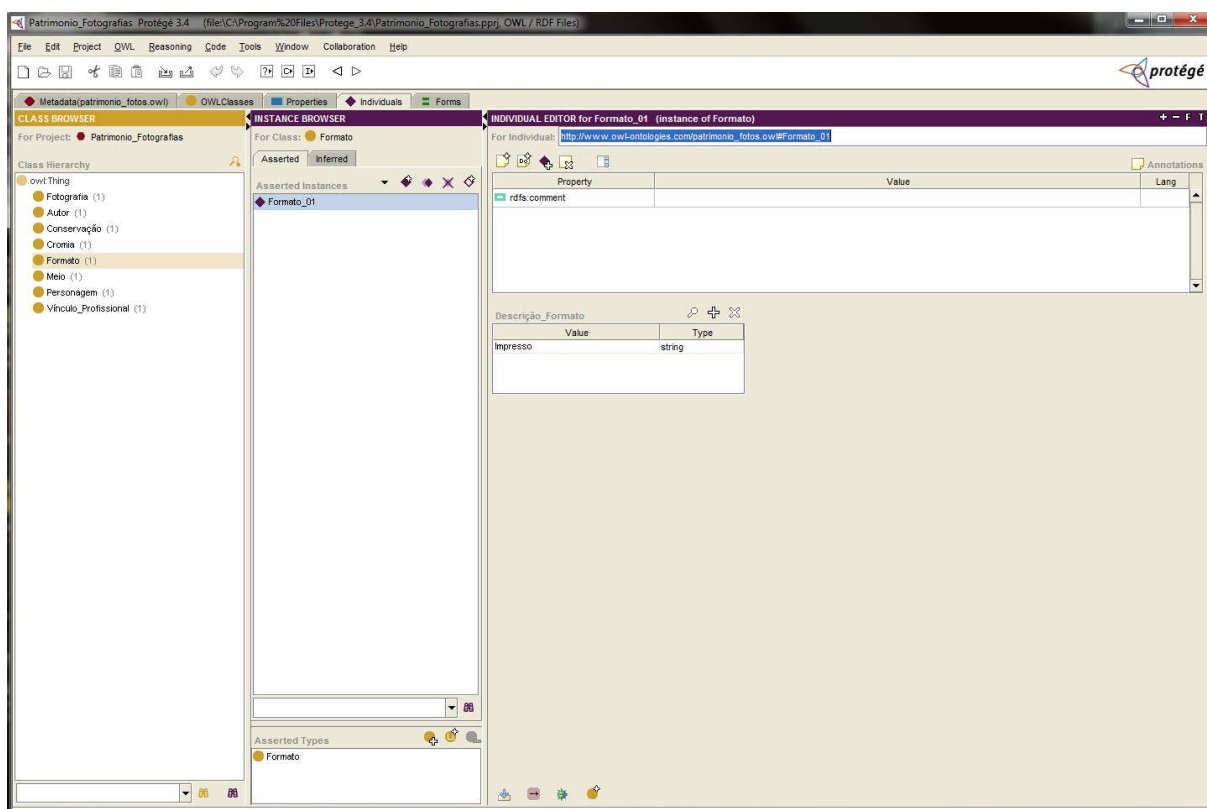


FIGURA 10 – VALORES “FORMATO”

Fonte: O Autor (2010).

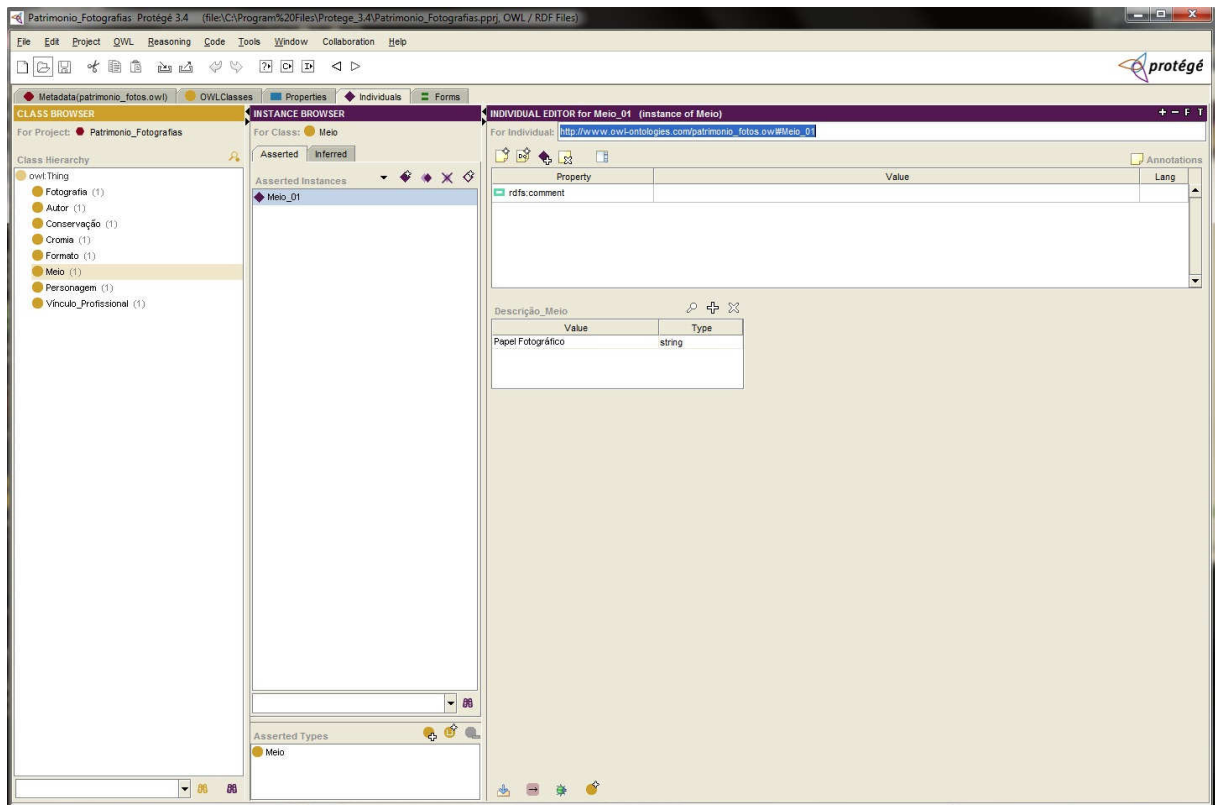


FIGURA 11 – VALORES “MEIO”

Fonte: O Autor (2010).

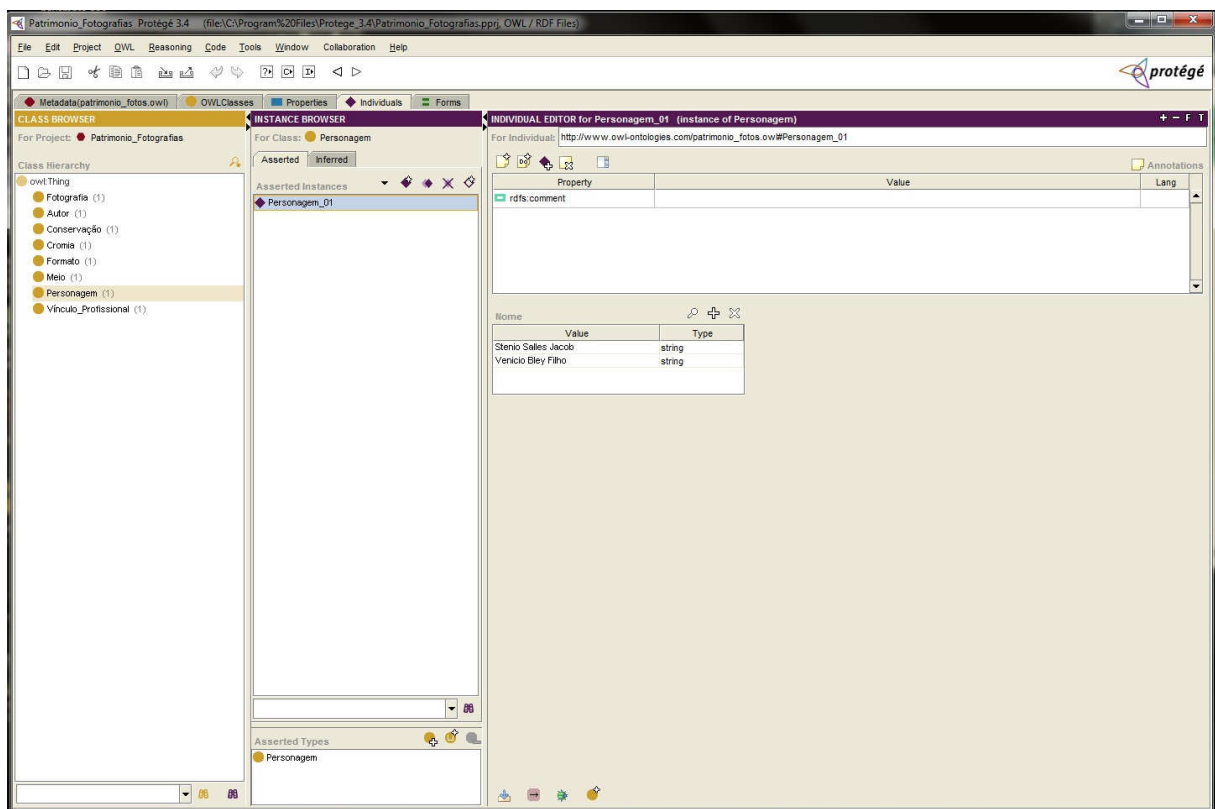


FIGURA 12 – VALORES “PERSONAGEM”

Fonte: O Autor (2010).

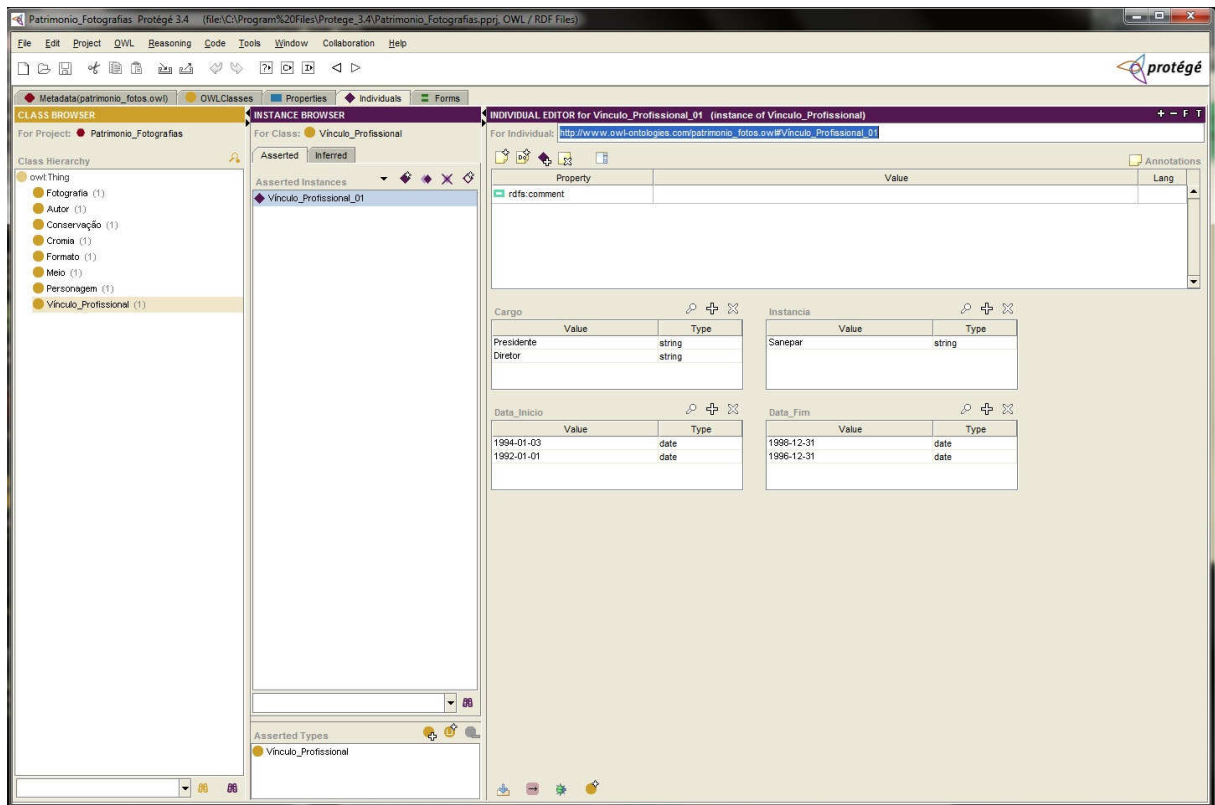


FIGURA 13 – VALORES “VÍNCULO\_PROFISSIONAL”

Fonte: O Autor (2010).

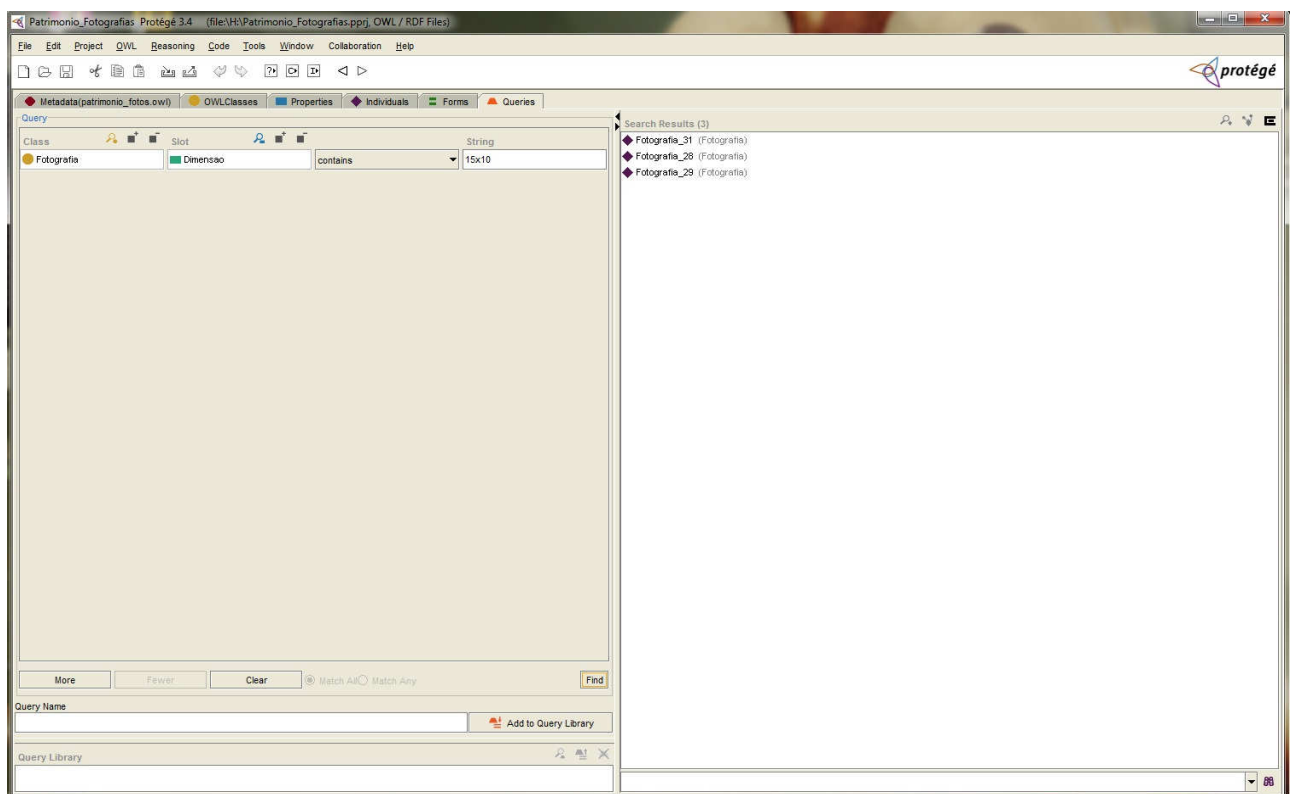


FIGURA 14 – FERRAMENTA DE BUSCA

Fonte: O Autor (2010).

Como ferramenta de busca o Protégé possui uma aba denominada “Queries” onde é possível recuperar os valores cadastrados através da sua classe, slot e por características semânticas. No exemplo da figura acima recuperou-se informações presentes na classe “Fotografia”, no slot “Dimensão”, contendo os valores 15x10.

Após a construção desta estrutura foi gerado pelo aplicativo Protégé um código OWL (APENDICE C) como forma de exportar a ontologia, garantindo um produto que possibilite interoperabilidade entre ambientes. A informação gerada pela OWL pode ser facilmente trocada entre diferentes tipos de computadores usando diferentes sistemas operacionais e linguagens de programação.

## 5 CONSIDERAÇÕES FINAIS

Este trabalho apresentou como proposta avaliar a possibilidade de utilização de ontologias como forma de diminuir as dificuldades relacionadas com a recuperação de informação.

Ao término deste trabalho, verificou-se que a aplicação de ontologias como meio para melhorar processos de recuperação da informação é possível e se justifica porque as ontologias possibilitam a elaboração de classes e suas relações permitindo que o usuário estabeleça o significado e contexto desejado e fornecendo meios para a construção de filtros mais eficientes.

Validou-se a utilização de ontologias como forma de disponibilizar informações e conhecimento de forma mais estruturada e de melhorar a recuperação, processamento e compartilhamento destas informações através de sua aplicação no Acervo Fotográfico da SANEPAR. O setor de Patrimônio Histórico pode se utilizar do portal da SANEPAR para disponibilizar as informações estruturadas sobre seus acervos, geradas pela ontologia, facilitando a recuperação e o compartilhamento destas informações com usuários internos e externos.

Como possibilidade para desenvolvimento de trabalhos futuros identificou-se:

- a) ampliação da base de conhecimento, cadastrando todos os registros presentes no Acervo Fotográfico da SANEPAR;
- b) possibilidade de agregar outros dados a ontologia como, por exemplo, o Acervo Documental, Mobiliário e Audiovisual da SANEPAR, construindo uma ontologia que relacione e disponibilize todas as informações dos Acervos em uma única base de conhecimento;
- c) criação de um ambiente Web utilizando as tecnologias da Web Semântica como o aplicativo Ontowiki, possibilitando a exportação e disponibilização da ontologia para a Intranet ou até mesmo portal da Companhia.

## REFERENCIAS

ALMEIDA, Maurício Barcellos; Bax, P. Marcello. **Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção.** Ciência da Informação, Brasília, v. 32, n. 3, set./dez. 2003. Disponível em: <<http://www.scielo.br/pdf/ci/v32n3/19019.pdf>>. Acesso em: 17 maio 2010.

AUER, Sören; DIETZOLD, Sebastian; RIECHERT, Thomas. **OntoWiki – A Tool for Social, Semantic Collaboration.** Disponível em: <http://www.informatik.unileipzig.de/~auer/publication/ontowiki.pdf>>. Acesso em: 01 novembro 2010.

BERNERS-LEE; HENDLER, J.; LASSILA, O. **The semantic web.** Scientific American, 2001.

BORST, W.N. **Construction of Engineering Ontologies for Knowledge Sharing and Reuse.** 1997. Tese (Phd). Disponível em: <<http://www.ub.utwente.nl/webdocs/inf/1/t0000004.pdf>>. Acesso em: 03 outubro 2010.

BRAY, Tim; HOLLANDER, Dave; LAYMAN, Andrew; TOBIN, Richard. **Namespaces in XML 1.0 (Second Edition),** 2006. Disponível em: <<http://www.w3.org/TR/RECxml-names/#sec-namespaces>>. Acesso em: 08 dezembro 2010.

CHOO, C. W. **The Knowing Organization: How Organizations Use Information for Construct Meaning, Create Knowledge and Make Decisions.** Nova Iorque: Oxford Press, 1998.

CLICKO, **O início Web 3.0 Web Semântica.** Disponível por WWW em <http://clicko.blogspot.com/2008/01/o-inicio-web-30-web-semantica.html>, 2008.

CORCHO, O. Knowledge modeling in web onto and OCML: a user guide. [ S. l. : s. n.], 1999. Disponível em: <[http://kmi.open.ac.uk/projects/webonto/user\\_guide.2.4.pdf](http://kmi.open.ac.uk/projects/webonto/user_guide.2.4.pdf)>. Acesso em: 09 jun 2002. [ Links ]

DAVIS, H. et al. Towards an integrated environment with open hypermedia systems. In: ACM CONFERENCE ON HYPERTEXT, 1992. Proceedings of... Milan, Italy, dec. 1992.

DZIEKANIAK, G. V., KIRINUS, J. B., **Web Semântica.** Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, n.18, 2º sem. Disponível por WWW em [http://www.encontrosbibli.ufsc.br/Edicao\\_18/2\\_Web\\_Semantica.pdf](http://www.encontrosbibli.ufsc.br/Edicao_18/2_Web_Semantica.pdf), 2004.

FAYEN, E. A new standard for controlled vocabularies. **The Indexer**, v.24, n.2, 2004.

FERNÁNDEZ-LÓPEZ, M. et al. Building a chemical ontology using methodology and the ontology design environment. [S. 1.] : IEEE Intelligent Systems & their Applications, p. 37-46, Jan./Feb. 1999.

FERNANDEZ, M.; GOMEZ-PEREZ, A.; JURISTO, H. Methontology; from ontological art towards ontological engineering (1997). Available from Internet: <<http://citeseer.ist.psu.edu/context/544607/0/>>. Access: 20 July 2005.

FIALHO, Francisco Antônio; MACEDO, Marcelo; SANTOS, Neri; MITIDIERI, Tibério. **Gestão do conhecimento e aprendizagem: as estratégias competitivas da sociedade pós-industrial**. Florianópolis: Visual Books, 2006a.

GIL, Antonio Carlos. Como elaborar projetos de pesquisa. 4. ed. São Paulo: Atlas, 2002.

GÓMEZ-PÉREZ, A. Evaluation of taxonomic knowledge in ontologies and knowledge bases. In: TWELFTH WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT, 12. 1999, Alberta, Canadá. [S. 1 : s. n.], 1999.

KLYNE, Graham; CARROLL, Jeremy J. Resource Description Framework (RDF): Concepts and Abstract Syntax. Disponível em: <<http://www.w3.org/TR/2004/RECrdf-concepts-20040210/>>. Acesso em: 01 novembro 2010.

LIMA, Fernanda; SCHWABE, Daniel. **Introdução à Web Semântica – MiniTutorial do evento WWS2004 I Workshop de Web Semântica - SBBB/SBES2004 - Brasília**. Disponível em: <<http://ftp.informatik.rwth-aachen.de/Publications/CEURWS/Vol-123/WWS2004MiniTutorial.pdf>>. Acesso em: 01 novembro 2010.

McGuinness, Deborah L.; Harmelen, Frank van. **OWL Web Ontology Language Overview**, 2004. Disponível em: <<http://www.w3.org/TR/owl-features/>>. Acesso em: 26 agosto 2010.

McGuinness, Deborah L.; Welty, Chris; Smith, Michael K. **OWL Web Ontology Language Guide**, 2004. Disponível em: <<http://www.w3.org/TR/owl-guide/>>. Acesso em: 26 setembro 2010.

MENDES, Raquel D. Projeto e desenvolvimento de um sistema especialista para diagnóstico de sarcomas ósseos. S.Paulo: ITA, 1991, Dissertação - mestrado.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **Guidelines for the construction, format, and management of monolingual controlled vocabularies**, 2005. (ANSI/NISO Z39.19-2005.)

NOY, F. N.; GUINNESS, D. L. **Ontology development 101: a guide to create your first ontology**. 2001. Disponível em: <<http://ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.doc>>. Acesso em: 03 Outubro 2010.

RAMALHO, R. A. S. **Web Semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação**. 2006. 120 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006

ROBREDO, J. Documentação de hoje e amanhã. Brasília, 1986.

T. R. Gruber. **Toward principles for the design of ontologies used for knowledge sharing. Presented at the Padua workshop on Formal Ontology**, Maço 1993, Vol. 43, cap 4-5, November 1995, pp. 907-928

Strauss, A. and Corbin, J. (1994). **Grounded Theory methodology: An overview**, In: **Handbook of Qualitative Research** (Denzin, N., K. and Lincoln, Y.,S., Eds.). Sage Publications, London, 1-18.

## APÊNDICES

## APENDICE A – Roteiros de entrevista

UNIVERSIDADE FEDERAL DO PARANÁ  
SETOR DE CIÊNCIAS SOCIAIS APLICADAS  
CURSO DE GESTÃO DA INFORMAÇÃO  
SIN030 – PESQUISA EM INFORMAÇÃO II  
ORIENTADOR – MAURO BELLI  
ALUNO: RODRIGO CEZAR M. TEIXEIRA

ROTEIRO DE ENTREVISTA: O PROCESSO DE ONTOLOGIAS APLICADO AO  
ACERVO FOTOGRÁFICO DA COMPANHIA DE SANEAMENTO DO PARANÁ –  
SANEPAR

1) COMO SE DA O PROCESSO DE IDENTIFICAÇÃO DOS ATRIBUTOS DE UMA  
FOTOGRAFIA?

---

---

---

---

---

2) QUAIS SÃO AS CARACTERÍSTICAS DAS FOTOGRAFIAS UTILIZADAS NA  
INDEXAÇÃO?

---

---

---

---

---

3) QUAL A FERRAMENTA UTILIZADA PARA O ARMAZENAMENTO DAS  
INFORMAÇÕES FOTOGRÁFICAS?

---

---

---

---

---

4) DE QUE FORMA SÃO CLASSIFICADAS E ORGANIZADAS AS INFORMAÇÕES  
FOTOGRÁFICAS?

---

---

---

---

---

5) DE QUE FORMA ESTAS INFORMAÇÕES FOTOGRÁFICAS SÃO  
RECUPERADAS?

---

---

---

---

---

6) QUAIS AS DIFICULDADES ENCONTRADAS NA RECUPERAÇÃO DESTAS INFORMAÇÕES?

---

---

---

---

---

7) ALGUMA SUGESTÃO PARA A MELHORIA DO ARMAZENAMENTO/ RECUPERAÇÃO DAS INFORMAÇÕES?

---

---

---

---

---

## **APÊNDICE B – Análise dos roteiros de entrevista**

Diante das informações obtidas na realização das entrevistas com os funcionários do setor chegou-se a seguinte visão do ambiente:

Após ser realizado o processo de identificação dos atributos de uma fotografia a mesma é indexada no software BrOffice Calc, um programa similar ao Excel, no qual são utilizadas planilhas para a inserção de dados. Este software mostra-se precário e pouco estruturado para a função de armazenagem e recuperação de fotos apresentando diversas barreiras no processo de gestão informacional das fotografias do setor. O BrOffice Calc não apresenta qualquer tipo de relacionamento entre informações impossibilitando um resultado de busca preciso.

A sugestão para melhoria no processo informacional apontada pelos funcionários do setor foi unânime. A aquisição de uma base de dados voltada à gestão de todo o acervo do Patrimônio Histórico possibilitando o armazenamento de um vasto acervo documental de maneira organizada e estruturada, visando uma recuperação eficiente das informações e conseqüentemente a melhoria nos serviços do setor.

## APÊNDICE C – Ontologia gerada na linguagem OWL

```

<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns="http://www.owl-ontologies.com/patrimonio_fotos.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://www.owl-ontologies.com/patrimonio_fotos.owl">
  <owl:Ontology rdf:about=""/>
  <owl:Class rdf:ID="Cromia"/>
  <owl:Class rdf:ID="Fotografia"/>
  <owl:Class rdf:ID="Formato"/>
  <owl:Class rdf:ID="Conservação"/>
  <owl:Class rdf:ID="Personagem"/>
  <owl:Class rdf:ID="Autor"/>
  <owl:Class rdf:ID="Vinculo_Profissional"/>
  <owl:Class rdf:ID="Meio"/>
  <owl:DatatypeProperty rdf:ID="Data">
    <rdfs:domain rdf:resource="#Fotografia"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="Descrição_Meio">
    <rdfs:domain>
      <owl:Class>
        <owl:unionOf rdf:parseType="Collection">
          <owl:Class rdf:about="#Meio"/>
          <owl:Class rdf:about="#Fotografia"/>
        </owl:unionOf>
      </owl:Class>
    </rdfs:domain>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="Dimensao">
    <rdfs:domain rdf:resource="#Fotografia"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="Cargo">
    <rdfs:domain rdf:resource="#Vinculo_Profissional"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="Data_Fim">
    <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
      ></rdfs:comment>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#date"/>
    <rdfs:domain rdf:resource="#Vinculo_Profissional"/>
  </owl:DatatypeProperty>
  <owl:DatatypeProperty rdf:ID="Telefone">
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int"/>
    <rdfs:domain>
      <owl:Class>

```

```

<Descrição_Cromia rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Colorida</Descrição_Cromia>
<Orgao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Sanepar</Orgao>
<Data rdf:datatype="http://www.w3.org/2001/XMLSchema#date"
>1998-06-17</Data>
<Descrição_Formato rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Impresso</Descrição_Formato>
<Tema rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Assinatura</Tema>
<Local rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Sede Sanepar</Local>
<Descrição_Meio rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Papel Fotográfico</Descrição_Meio>
<Descricao_Fotografia rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Stênio Salles Jacob (Presidente Sanepar), juntamente com Venício Bley Filho (Diretoria Sanepar), assinam convênio com rede
<Dimensao rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>18x12</Dimensao>
<Email rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>jh@sanepar.com</Email>
<Titulo rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Convênio com farmácia para recebimento de contas de água</Titulo>
<Telefone rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
>32057889</Telefone>
<Cidade rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Curitiba</Cidade>
</Fotografia>
<Meio rdf:ID="Meio_01">
<Descricao_Fotografia rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Papel Fotográfico</Descricao_Fotografia>
<Descrição_Meio rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Papel Fotográfico</Descrição_Meio>
</Meio>
<Vinculo_Profissional rdf:ID="Vinculo_Profissional_01">
<Data_Inicio rdf:datatype="http://www.w3.org/2001/XMLSchema#date"
>1994-01-03</Data_Inicio>
<Data_Inicio rdf:datatype="http://www.w3.org/2001/XMLSchema#date"
>1992-01-01</Data_Inicio>
<Cargo rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Presidente</Cargo>
<Cargo rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Diretor</Cargo>
<Instancia rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Sanepar</Instancia>
<Data_Fim rdf:datatype="http://www.w3.org/2001/XMLSchema#date"
>1998-12-31</Data_Fim>
<Data_Fim rdf:datatype="http://www.w3.org/2001/XMLSchema#date"
>1996-12-31</Data_Fim>
</Vinculo_Profissional>
</rdf:RDF>

```