



**Universidade de Brasília**  
**Faculdade de Economia, Administração, Contabilidade e Ciência da**  
**Informação e Documentação**  
**Departamento de Ciência da Informação e Documentação**

**Ontologias: Indexação e Recuperação de Fotografias**  
**Baseadas na Técnica Fotográfica e no Conteúdo da**  
**Imagem**

**Alex Sandro Santos Miranda**

Brasília  
2007

**Alex Sandro Santos Miranda**

# **Ontologias: Indexação e Recuperação de Fotografias Baseadas na Técnica Fotográfica e no Conteúdo da Imagem**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação do Departamento de Ciência da Informação e Documentação da Universidade de Brasília como exigência parcial para a obtenção do título de Mestre em Ciência da Informação.

Orientadora: Prof.a Dr.a Miriam Paula Manini

Co-orientador: Prof. Dr. Mamede Lima-Marques

Brasília

2007

# Agradecimentos

A Deus, por tudo...; aos meus pais, e, em especial, à minha mãe, pela ajuda, compreensão e incontáveis orações destinadas a mim. Aos meus irmãos, Marcos, Mara e Aparecida, pelo carinho, compreensão e ajuda constante. Ao meu primo Jaenes Miranda, pela ajuda, incentivo e apoio. À minha avó, Alice Bispo (*in memoriam*), pelo carinho, incentivo e pela pessoa maravilhosa que era. Ao meu tio João Alves, pelo apoio e orientação. Ao meu primo Fredson Bispo, pela amizade e companheirismo. À minha madrinha, Idalice Bispo pela ajuda e apoio. À Prof.a Dr.a Miriam Paula Manini, por propiciar a oportunidade de realizar este trabalho, pela ajuda e constante incentivo. Ao Prof. Dr. Mamede Lima-Marques, pela ajuda e preciosas dicas. À Rita de Cássia (*in memoriam*) por ter sido uma pessoa especial na minha vida. Ao meu amigo Jânio Teixeira pela amizade e companheirismo. À minha amiga Larissa Costa, pela amizade e apoio. Ao meu chefe de trabalho Gabriel Biermann por ter permitido alguns dias de ausência do trabalho (com reposição) para realizar atividades do mestrado. Finalmente, a todos aqueles que direta ou indiretamente contribuíram para a elaboração deste trabalho.

# Resumo

Este trabalho aborda as técnicas de representação da imagem por conceito com o objetivo de fundamentar a construção de uma ontologia de descrição de fotografias através da utilização de padrões da *Web Semântica*, como as linguagens OWL e RDF *Schema*. A ontologia de descrição consiste numa representação ontológica do conteúdo visual da imagem, além de definir um esquema de descrição de imagens. Para validar a ontologia de descrição, foi desenvolvido, como parte da metodologia desta pesquisa, um sistema de descrição e recuperação de imagens que usa a ontologia internamente para descrever e recuperar fotografias. A idéia deste trabalho é enriquecer a descrição de imagens com a associação de conceitos e instâncias extraídos da ontologia. A ontologia de descrição foi construída levando em consideração a informação presente na imagem e dados oriundos da técnica fotográfica usada no momento da produção da fotografia. O resultado da descrição da imagem usando o sistema protótipo de descrição é um metadado em RDF (*Resource Description Framework*) que descreve o conteúdo visual da imagem, somado com a descrição da técnica fotográfica empregada na produção da fotografia.

**Palavras-chave:** Ontologia, Indexação de Imagens.

# Abstract

The work reported in this dissertation studied the techniques of image representation by concept. The main aim was to set the foundations needed to build up an ontology for describing photographs by means of semantic web patterns, such as OWL and RDF-Schema. The describing ontology developed in the study consists of an ontological representation of the visual context of an image, plus the definition of an image description schema. In order to validate the describing ontology, an image (photo) description and retrieval system, which uses the ontology internally, has been developed. The idea that underlined the approach adopted was that if the description of images can be associated with concepts and instances extracted from the ontology, the work is significantly enriched. Accordingly, the describing ontology has been built taking into account the information found in the image itself, and data taken from the photograph technique used at the time the photo was produced. The resultant image description through the use of a prototype description system is a RDF (Resource Description Framework) metadata that describes the visual context of the image, added up to the photograph technique employed in the photo production.

**Keywords:** Ontology, Image Indexing.

# Sumário

<b>Agradecimentos</b>	p. i
<b>Resumo</b>	p. ii
<b>Abstract</b>	p. iii
<b>Lista de Figuras</b>	p. vii
<b>Lista de Quadros</b>	p. x
<b>Lista de Abreviaturas</b>	p. xi
<b>1 Introdução</b>	p. 12
<b>2 Requisitos da Pesquisa</b>	p. 15
2.1 Objetivos . . . . .	p. 15
2.1.1 Objetivo Geral . . . . .	p. 15
2.1.2 Objetivos Específicos . . . . .	p. 15
2.2 Justificativa . . . . .	p. 15
2.3 Metodologia . . . . .	p. 16
<b>3 Técnicas de Descrição e Recuperação</b>	p. 19
3.1 Descrição e Recuperação de Imagens com Base em Texto . . . . .	p. 19
3.2 Descrição e Recuperação de Imagens com Base em Atributos . . . . .	p. 22
3.3 Descrição e Recuperação de Imagens com Base em Ontologias . . . . .	p. 25

---

<b>4</b>	<b>Representação da Imagem</b>	p. 27
4.1	Representação do Conteúdo Visual . . . . .	p. 28
4.2	Representação do Conteúdo Não Visual . . . . .	p. 41
4.3	Técnica e Expressão Fotográfica . . . . .	p. 43
4.3.1	Efeitos Especiais . . . . .	p. 47
4.3.2	Ótica . . . . .	p. 48
4.3.3	Tempo de Exposição . . . . .	p. 51
4.3.4	Luminosidade . . . . .	p. 54
4.3.5	Enquadramento . . . . .	p. 57
4.3.6	Posição da Câmera . . . . .	p. 60
4.3.7	Composição . . . . .	p. 63
4.3.8	Profundidade de Campo . . . . .	p. 65
<b>5</b>	<b>Ontologias</b>	p. 67
5.1	Ontologias e Usos . . . . .	p. 68
5.2	Tipos de Ontologias . . . . .	p. 71
5.3	Metodologias de Construção de Ontologia . . . . .	p. 74
5.3.1	Identificação da finalidade e do escopo da ontologia . . . . .	p. 75
5.3.2	Construção da ontologia . . . . .	p. 75
5.3.2.1	Definição das classes e hierarquias de classes . . . . .	p. 75
5.3.2.2	Codificação da ontologia . . . . .	p. 77
5.3.2.3	Integração de ontologias existentes . . . . .	p. 78
5.3.3	Avaliação de ontologia . . . . .	p. 78
5.3.4	Documentação . . . . .	p. 79
5.4	Exemplos de Ontologias . . . . .	p. 79
5.4.1	CYC . . . . .	p. 79
5.4.2	<i>WordNet</i> . . . . .	p. 80

---

5.4.3	<i>Open Directory Project</i> . . . . .	p. 80
5.5	Linguagens Formais . . . . .	p. 80
5.5.1	Linguagens de predicado de primeira ordem . . . . .	p. 81
5.5.2	XML, RDF, RDF <i>Schema</i> e OWL . . . . .	p. 82
5.5.3	Editor de ontologias <i>Protégé</i> . . . . .	p. 83
<b>6</b>	<b>A Ontologia de Descrição de Imagens</b>	<b>p. 85</b>
6.1	Finalidade e Escopo da Ontologia . . . . .	p. 85
6.2	Lista de Termos . . . . .	p. 87
6.3	Desenvolvimento da Ontologia . . . . .	p. 90
6.3.1	Ontologia do conteúdo visual . . . . .	p. 90
6.3.1.1	Construção da ontologia Quem . . . . .	p. 91
6.3.1.2	Construção da ontologia Onde . . . . .	p. 96
6.3.1.3	Construção da ontologia Quando . . . . .	p. 99
6.3.1.4	Construção da ontologia Que . . . . .	p. 100
6.3.1.5	Construção da ontologia Sobre . . . . .	p. 102
6.3.1.6	Construção da ontologia Técnica Fotográfica . . . . .	p. 104
6.3.2	Ontologia do conteúdo não visual . . . . .	p. 106
6.4	Teste da Ontologia de Descrição de Imagens . . . . .	p. 113
6.4.1	<i>Splash</i> : módulo de descrição de imagens . . . . .	p. 113
6.4.2	<i>Splash</i> : módulo de recuperação de imagens . . . . .	p. 117
<b>7</b>	<b>Conclusão</b>	<b>p. 120</b>
	<b>Referências</b>	<b>p. 122</b>
	<b>Apêndice A - Esquema de indexação de imagens em RDF Schema</b>	<b>p. 127</b>

## Lista de Figuras

1	Juscelino Kubitschek . . . . .	p. 20
2	Dr. Peter Henry Rolfs. . . . .	p. 34
3	Pirâmide de classificação de informação visual de Jaimes e Chang (2000). . . . .	p. 38
4	Ashley Judd 1 . . . . .	p. 44
5	Ashley Judd 2 . . . . .	p. 45
6	Ashley Judd 3 . . . . .	p. 45
7	Ashley Judd 4 . . . . .	p. 45
8	<i>Panning</i> . . . . .	p. 48
9	<i>Splash</i> . . . . .	p. 48
10	Grande angular . . . . .	p. 49
11	<i>Fish eye</i> . . . . .	p. 50
12	Teleobjetiva . . . . .	p. 50
13	Macro . . . . .	p. 50
14	Instantâneo . . . . .	p. 52
15	Longa exposição . . . . .	p. 53
16	Dupla exposição . . . . .	p. 53
17	Luz noturna . . . . .	p. 55
18	Luz diurna . . . . .	p. 55
19	Contraluz . . . . .	p. 56
20	Plano médio . . . . .	p. 58
21	Plano americano . . . . .	p. 58
22	<i>Close</i> ou Grande plano . . . . .	p. 59

---

23	Detalhe ou Plano pormenor . . . . .	p. 59
24	Câmara baixa . . . . .	p. 61
25	Câmara alta . . . . .	p. 61
26	Vista aérea . . . . .	p. 62
27	Regra dos terços . . . . .	p. 63
28	Exemplo da aplicação da regra dos terços . . . . .	p. 64
29	Com profundidade de campo . . . . .	p. 65
30	Sem profundidade de campo . . . . .	p. 66
31	Espectro de ontologias. . . . .	p. 69
32	Exemplo de comunicação entre 4 linguagens. . . . .	p. 70
33	Exemplo de ontologia como interlíngua. . . . .	p. 71
34	Diagrama de classe Quem. . . . .	p. 93
35	Ontologia Quem. . . . .	p. 95
36	Diagrama de classe Onde. . . . .	p. 97
37	Ontologia Onde. . . . .	p. 98
38	Diagrama de classe Quando. . . . .	p. 99
39	Ontologia Quando. . . . .	p. 99
40	Diagrama de classe Que. . . . .	p. 100
41	Ontologia Que. . . . .	p. 101
42	Ontologia Que. . . . .	p. 102
43	Ontologia Que: categoria artes. . . . .	p. 102
44	Diagrama de classe Sobre. . . . .	p. 103
45	Ontologia Sobre. . . . .	p. 104
46	Ontologia Sobre: categoria artes. . . . .	p. 104
47	Diagrama de classe Técnica Fotográfica. . . . .	p. 105
48	Ontologia Técnica Fotográfica. . . . .	p. 106

---

49	Diagrama de classe Mídia. . . . .	p. 108
50	Ontologia de Descrição de Imagens. . . . .	p. 109
51	Ivete Sangalo no carnaval de Salvador. . . . .	p. 110
52	Geração de instância da classe Fotografia no <i>Protégé</i> . . . . .	p. 111
53	Geração de instância da classe Cantora no <i>Protégé</i> . . . . .	p. 111
54	Descrição da fotografia em RDF. . . . .	p. 112
55	Roberto Carlos. . . . .	p. 114
56	<i>Splash</i> : formulário de descrição do conteúdo não visual. . . . .	p. 115
57	<i>Splash</i> : formulário de descrição do conteúdo visual. . . . .	p. 116
58	<i>Splash</i> : módulo de busca de imagens. . . . .	p. 119

## Lista de Quadros

1	Conjunto de atributos do metadado <i>Dublin Core</i> . . . . .	p. 24
2	Categorias para a representação da imagem. . . . .	p. 31
3	Descrição de elementos do conteúdo informacional. . . . .	p. 32
4	Grade de Análise Documentária de Imagens. . . . .	p. 33
5	Exemplo da aplicação da Grade de Análise Documentária de Imagens. . .	p. 35
6	Recursos Técnicos. . . . .	p. 46
7	Tipos de ontologias quanto ao grau de formalidade. . . . .	p. 72
8	Tipos de ontologias quanto à função. . . . .	p. 72
9	Tipos de ontologias quanto à estrutura. . . . .	p. 72
10	Tipos de ontologias quanto ao conteúdo. . . . .	p. 73
11	Lista de Termos. . . . .	p. 89

# Lista de Abreviaturas

1. **CBIR** *Content-Based Image Retrieval*
2. **HTML** *Hipertext Markup Language*
3. **HTTP** *HyperText Transfer Protocol*
4. **JPEG** *Joint Photographic Experts Group*
5. **OWL** *Web Ontology Language*
6. **PNG** *Portable Network Graphics*
7. **RDF** *Resource Description Framework*
8. **UML** *Unified Modeling Language*
9. **URI** *Uniform Resource Identifier*
10. **WWW** *World Wide Web*
11. **W3C** *World Wide Web Consortium*

# 1 Introdução

O foco deste trabalho é o processo de indexação e recuperação de imagens no contexto de sistemas de informação. Abordaremos tanto a indexação quanto a recuperação de imagens fotográficas, visando a melhorar a precisão na recuperação de imagens em sistemas de informação.

Um dos grandes problemas dos sistemas de recuperação disponíveis na Internet é a baixa precisão da pesquisa, ou seja, muitos registros irrelevantes são recuperados no momento da busca; isto acontece, em parte, devido ao paradigma de consulta com base em palavras-chave. Em se tratando de sistemas de recuperação de imagens, o problema agrava-se, pois a documentação de imagens exige técnicas que, na maioria das vezes, são desconhecidas pelos desenvolvedores de sistemas. O problema é que as imagens se tornam virtualmente irrecuperáveis se não existir uma descrição associada a elas. E a maneira como esta descrição é realizada tem um enorme impacto na recuperação de imagens.

Existem atualmente duas abordagens para o problema da indexação de imagens: a primeira é conhecida como indexação com base no conteúdo e a segunda é conhecida como indexação com base em conceitos. Na abordagem que se baseia no conteúdo, a indexação é realizada de modo automático, por algoritmos computacionais, através da extração de características de cor, textura e arranjo de figuras geométricas presentes na imagem. Na indexação com base em conceitos, as imagens são representadas por uma lista de palavras-chave que se referem às informações presentes na imagem. A indexação por conteúdo opera no nível sintático, enquanto a indexação por conceitos opera no nível semântico. Logo, eles não se opõem; são complementares.

Nosso enfoque é a indexação por conceitos. Discutiremos os modelos de indexação de imagens propostos pela Ciência da Informação e veremos em que medida a Ciência da Computação pode contribuir para o aperfeiçoamento dos modelos de indexação de imagens através do desenvolvimento de metadados e linguagem de representação do conhecimento como ontologias. O processo de indexação e recuperação de imagens pode ser simplificado

no seguinte diagrama:

Imagem -> Representação <- Recuperação <- Expressão de busca <- Usuário

Um dos objetivos deste trabalho é analisar este processo e apresentar uma modelagem conceitual para indexação e recuperação de imagens embasada nos avanços das pesquisas sobre estes temas. Seguindo o diagrama, vimos que, inicialmente, a imagem passa pela etapa de representação, ou seja, etapa de seleção de conceitos que irão representar a informação contida na imagem.

Neste momento, a grande dificuldade é: 1) o que indexar? Este problema é fonte inesgotável de pesquisas na Ciência da Informação e em Processamento de Imagens. Ainda não existe um modelo de indexação que represente de modo preciso a informação presente numa imagem, mas existem técnicas que recuperam as informações mais relevantes. Uma vez obtida a lista de indexadores, surge uma nova questão: 2) como representar os indexadores? Existem várias alternativas: palavras-chave, resumo textual, tesouro, ontologias, etc. A outra questão é: 3) como o usuário formulará a sua expressão de busca?

A solução adotada para estas três questões determinará a qualidade do sistema de recuperação. É possível, através do estudo deste processo, verificar a transformação da informação imagética que ocorre desde o momento em que é representada através de indexadores até o momento em que é selecionada por um usuário através de uma expressão de busca.

Pretendemos abordar estes problemas: o que indexar, como indexar e como recuperar no contexto de sistemas de informação; e desenvolver uma ontologia de descrição de imagens levando em consideração as informações intrínsecas da imagem, isto é, que fazem parte do conteúdo da imagem, e as informações extrínsecas, ou seja, informações que não fazem parte do conteúdo visual da imagem, mas que estão associadas à imagem de alguma forma, como, por exemplo, fotógrafo, título, data, etc.

Sobre o problema da indexação (o que indexar) existem métodos e técnicas que são empregados para extração de descritores ou de palavras-chave com base no conteúdo da imagem. Alguns pesquisadores têm explorado a questão da técnica fotográfica como recurso de indexação. Smit (1997, p. 3) categoriza três parâmetros de análise: o que a fotografia mostra (ou seja, seu conteúdo informacional); como a fotografia mostra (a forma adotada para mostrar o conteúdo informacional, ou seja, a expressão fotográfica); onde a fotografia mostra (dedicado à questão do documento fotográfico enquanto objeto físico).

Os dois primeiros itens - o que a fotografia mostra e como a fotografia mostra - estão estreitamente relacionados com a indexação de imagens fotográficas, enquanto o último item não se encaixa neste estudo por não ser de interesse da área de indexação de imagens.

A maior parte dos mecanismos de busca de imagens fotográficas, disponíveis na Internet utilizam técnicas de descrição com base em informações extra-imagéticas, isto é, informações que não estão presentes na imagem, tais como: fotógrafo, título, data, descrição textual, etc. Embora estas informações sejam relevantes, elas não representam a informação presente na imagem. Para melhorar a recuperação de imagens fotográficas há a necessidade de representar o conteúdo informacional (o que a imagem mostra) e a sua expressão fotográfica (como a imagem mostra). A expressão fotográfica é produzida pelo fotógrafo com o uso da técnica fotográfica. Duas imagens sobre o mesmo objeto podem ter uma recepção diferente por causa da técnica fotográfica empregada. A título de exemplo, imagine três fotografias de uma modelo: 1) a fotografia do rosto (*close up*); 2) a fotografia da modelo inteira (plano geral); 3) a fotografia da modelo em preto-e-branco (cromia). Estas três fotografias correspondem ao mesmo conteúdo (modelo) mas com expressões diferentes (enquadramento, cromia, etc.). A partir da seleção de um conjunto de fotografias com mesmo conteúdo, a técnica fotográfica é determinante na escolha da fotografia.

## 2 Requisitos da Pesquisa

### 2.1 Objetivos

#### 2.1.1 Objetivo Geral

Propor um modelo de descrição de imagens com base em ontologias para enriquecer a descrição fotográfica com o objetivo de melhorar a recuperação de imagens através de um sistema de busca.

#### 2.1.2 Objetivos Específicos

- Propor uma ontologia de descrição de imagens;
- Propor um sistema protótipo de descrição e busca de imagens com base na ontologia de descrição.

### 2.2 Justificativa

As imagens são usadas pelos mais diversos tipos de profissionais e para as mais diversas finalidades. Os editores de um jornal ou de uma revista necessitam de imagens específicas para compor um artigo, para representar o conteúdo do texto e até mesmo para sugerir certa situação; os historiadores usam como fonte de pesquisa para comprovação e contextualização de acontecimentos históricos; educadores usam como instrumentos de ensino, onde a representação textual falha ou apresenta lacunas; os publicitários exploram as imagens para atrair clientes e consumidores. Embora exista uma grande aplicabilidade e diversificação de usuários, estes frequentemente encontram dificuldade para recuperar imagens para atender suas necessidades.

Encontrar uma fotografia para uma necessidade particular é tarefa cada vez mais difícil. Ao tentar recuperar uma imagem em um sistema de busca, o usuário depara-se

com o problema da quantidade excessiva de registros irrelevantes e com a dificuldade de selecionar termos adequados e mais precisos para formular a consulta. A baixa precisão dos mecanismos de buscas está relacionada com o fato de serem tradicionalmente voltados para busca com base em palavras-chave, que não tratam adequadamente os termos sinônimos e antônimos, como consequência do processo de indexação ser realizado com base em unidades léxicas, o que torna difícil melhorar a precisão da recuperação.

A indexação de imagens é, muitas vezes, realizada levando em conta somente a legenda da fotografia, que, de acordo com as metodologias de documentação é apenas uma pequena parte da informação que pode estar associada à imagem. Se a indexação é realizada com base numa informação que não retrata a informação contida na imagem, isso certamente ocasionará problemas na recuperação da informação.

Uma ontologia de descrição que represente o conteúdo da imagem, a expressão fotográfica e a informação não visual somada a uma alternativa de recuperação que facilite a recuperação de imagens com base no conteúdo da imagem pode tornar os sistemas de recuperação de imagens fotográficas mais eficientes e capazes de atender necessidades específicas dos usuários de fotografias.

Ontologias podem ser consideradas uma alternativa para minimizar o problema da quantidade excessiva de registros irrelevantes, uma vez que a recuperação se baseia numa estrutura de relacionamentos semânticos e não mais na unidade léxica, como no mecanismo de busca com base em palavras-chave. As hierarquias de conceitos da ontologia podem ser utilizadas no momento da busca para auxiliar o usuário a selecionar e a combinar conceitos para recuperar a informação. Adicionalmente, ontologias pode ser empregadas para enriquecer a descrição de imagens por meio da associação de conceitos e instâncias ao metadado da imagem.

Este trabalho se justifica devido aos problemas de recuperação de imagens mencionados anteriormente e à falta de tratamento adequado na documentação de imagens no contexto de sistemas de informação; além disso, leva em consideração a questão da técnica fotográfica como recurso de descrição, que ainda não foi empregada no contexto de sistemas de informação e pode contribuir muito para melhorar a recuperação de imagens.

## 2.3 Metodologia

A metodologia desta pesquisa pode ser classificada quanto aos procedimentos usados para atingir os objetivos, como pesquisa bibliográfica, visto que recorreremos à literatura

da área de indexação de imagens para abordar as técnicas de representação da imagem que serão empregadas na construção de um esquema de descrição de fotografias com base em ontologias.

A ontologia de descrição de imagens será desenvolvida com base em algumas técnicas de descrição de imagens por conceitos que geralmente estabelecem um conjunto de atributos para descrever a informação presente na imagem. Na literatura sobre descrição de imagens por conceitos, podemos encontrar várias técnicas de descrição de imagens que estabelecem métodos de transposição da informação iconográfica para informação verbal, através de palavras-chave. Iniciaremos a partir de uma abordagem destas técnicas de representação da imagem por conceito para embasar a nossa proposta; em seguida definiremos quais técnicas pretendemos representar na ontologia de descrição.

Basicamente, empregaremos os seguintes procedimentos para desenvolver a ontologia de descrição de imagens:

- Definição das técnicas de representação de imagens que iremos implementar na ontologia de descrição, com base em um estudo anterior sobre a representação da imagem;
- Definição do escopo da ontologia;
- Classificação da ontologia quanto ao grau de formalidade e ao uso;
- Elaboração de uma lista de termos com a definição dos atributos da imagem que serão representados no esquema de descrição;
- Realização de uma modelagem conceitual da ontologia de descrição;
- Desenvolvimento da ontologia de descrição no ambiente de desenvolvimento *Protege*<sup>1</sup>;
- Construção de um sistema protótipo de descrição e busca de imagens que usa a ontologia de descrição de imagens internamente para descrever e recuperar fotografias;
- Abordagem da descrição de imagens obtida através do protótipo de descrição com o uso da ontologia de descrição de imagens.

Inicialmente, fizemos uma revisão de literatura sobre as técnicas de representação de imagens por conceitos que fundamentará a construção da ontologia. Depois deste estudo,

---

<sup>1</sup><http://protege.stanford.edu>

definimos quais técnicas representamos na ontologia de descrição, justificando a escolha. Depois, aplicamos alguns procedimentos de desenvolvimento de ontologias que geralmente consistem em definir inicialmente a finalidade e o escopo da ontologia. O próximo passo consistiu em classificar a ontologia empregando algum critério de classificação como, por exemplo, quanto ao grau de formalidade, uso, etc. Em seguida, realizamos a modelagem da ontologia de descrição usando a linguagem UML (*Unified Modeling Language*) que consistirá em representar as técnicas de descrição por meio de classes, propriedades e relações com o uso de diagramas da linguagem UML. Esta modelagem forneceu um nível de especificação da ontologia de modo que podéssemos implementá-la num ambiente de desenvolvimento como o *Protégé*. A implementação da ontologia consiste em definir a linguagem de representação formal e construir a ontologia com base na modelagem, através da criação de classes, instâncias, propriedades, relações e organizar o conhecimento numa hierarquia taxonômica. A partir da ontologia implementada numa linguagem formal, como RDF *Schema*, iniciamos a construção de um sistema protótipo de descrição e busca de imagens que usará a ontologia internamente para descrever e recuperar imagens. A construção deste sistema tem como objetivo validar o esquema de descrição de imagens e apresentar uma proposta de recuperação com base na ontologia de descrição.

Espera-se, ao final deste trabalho, que a ontologia de descrição contribua para a produção de descrição de imagens de forma representativa do seu conteúdo, através de sistema de indexação que a empregue internamente e, conseqüentemente, contribua para melhorar a precisão de sistemas de recuperação de fotografias.

## 3 Técnicas de Descrição e Recuperação

Este capítulo discute algumas técnicas de descrição e recuperação de imagens no contexto de sistemas informatizados. Basicamente, dispomos de três alternativas de descrição: com base em texto, atributo e ontologias. Similarmente, a recuperação basear-se-á em texto, atributo ou ontologias. No processo de recuperação, a técnica de descrição empregada determinará a qualidade do mecanismo de recuperação, ou seja, se a técnica de descrição for baseada em texto, conseqüentemente a recuperação se baseará em texto, logo, o mecanismo de recuperação apresentará os problemas inerentes a esta abordagem. Inicialmente, descrevemos a descrição com base em texto; em seguida, discutimos algumas técnicas de recuperação com base em texto, ressaltando o problema da baixa precisão desta abordagem. Depois, descrevemos a técnica de descrição e recuperação com base em atributos. Finalmente, apresentamos a técnica de descrição e recuperação com base em ontologia, ressaltando as vantagens que esta técnica pode trazer para a descrição e a recuperação de imagens.

### 3.1 Descrição e Recuperação de Imagens com Base em Texto

A descrição de imagens com base em texto é muito simples; consiste em representar a imagem com o uso de uma descrição textual, como na Figura 1. Shatford (1986) e Smit (1997) defendem que esta descrição deve ser realizada observando o que a imagem mostra. Uma vez obtida a descrição, ela é associada à imagem no repositório de imagens. O problema da descrição com base em texto é que, conseqüentemente, a recuperação será baseada em texto. Geralmente, um sistema de recuperação com base em texto possui uma interface de pesquisa onde o usuário digita algumas palavras-chave como parâmetro de pesquisa. Depois, o sistema recupera os documentos do repositório com base no casamento de padrão e na freqüência dos termos de pesquisa contidos na descrição do documento (imagem). Quando o número de descrições textuais aumenta, os problemas

clássicos da recuperação com base em texto começam a surgir, como, por exemplo, a grande quantidade de registros irrelevantes recuperados, ou seja, a baixa precisão. Existem algumas técnicas que tornam a recuperação com base em texto mais robusta, mas não resolvem o problema. Dentre elas citemos: modelo booleano, modelo vetorial, *stemming* e o modelo probabilístico.



Figura 1: Juscelino Kubitschek

Juscelino Kubitschek de Oliveira (Diamantina, 12 de setembro de 1902 – Resende, 22 de agosto de 1976) foi um médico, militar e político brasileiro. Conhecido como JK (lê-se jota-cá), foi presidente do Brasil entre 1956 e 1961. Foi casado com Sarah Kubitschek, pai de Márcia Kubitschek e pai adotivo de Maria Estela Kubitschek.

Fonte: [http://pt.wikipedia.org/wiki/Juscelino\\_Kubitschek](http://pt.wikipedia.org/wiki/Juscelino_Kubitschek), 15/04/2007

No modelo booleano um documento é representado por um conjunto de termos de indexação que podem ser definidos de forma manual por profissionais especializados, ou automaticamente, através de algum tipo de algoritmo computacional. A recuperação dos documentos é realizada através de expressões booleanas que empregam os operadores lógicos **AND**, **OR** e **NOT**. Por exemplo, seja **X** o conjunto de documentos indexados pelo termo **P** e seja **Y** o conjunto de documentos indexados pelo termo **Q**. A expressão de busca **P AND Q** recuperará os documentos que foram indexados com termo **P** e com o termo **Q**. Equivale à interseção dos conjuntos de documentos indexados por **X** e **Y**. A expressão de busca **P OR Q** recuperará os documentos que foram indexados por **P** ou por **Q**, ou seja, por pelo menos um termo de indexação. Esta operação equivale à união dos documentos indexados por **P** ou **Q**. A expressão **NOT P** recuperará os documentos que não foram indexados com termo **P**.

Segundo Ferneda (2003), o modelo booleano, apesar de bem formalizado, possui limitações que diminuem sua atratividade. Algumas dessas limitações são:

- sem um treinamento apropriado, o usuário leigo será capaz de formular somente buscas simples. Para buscas que exijam expressões mais complexas será necessário um conhecimento sólido de lógica booleana;
- existe pouco controle sobre a quantidade de documentos resultantes de uma busca. O usuário é incapaz de prever quantos registros satisfarão à restrição lógica de uma determinada expressão booleana, sendo necessário realizar sucessivas reformulações antes que seja recuperado um volume aceitável de documentos;
- o resultado de uma busca booleana caracteriza-se por uma simples partição do corpus em dois subconjuntos: os documentos que atendem à expressão de busca e aqueles que não atendem. Presume-se que todos os documentos recuperados são de igual utilidade para o usuário. Não há qualquer mecanismo pelo qual os documentos possam ser ordenados;
- não existe uma forma de atribuir importância relativa aos diferentes termos na expressão booleana. Assume-se implicitamente que todos os termos têm o mesmo peso.

O modelo vetorial é outra técnica de recuperação com base em texto, mais sofisticada e que pode melhorar a precisão da busca. No modelo vetorial tanto os documentos quanto a expressão de busca são representados por vetores de pesos de termos de indexação. Cada elemento do vetor representa o peso ou a relevância do termo de indexação. Os elementos do vetor são normalizados para atingir valores entre 0 e 1. Quanto mais próximo de 1, maior é a relevância do termo de indexação para a descrição do documento. A recuperação no modelo vetorial é realizada pela similaridade entre o vetor de busca com os vetores dos documentos representados. São recuperados aqueles documentos que possuam um grau de similaridade superior ao grau previamente estabelecido em relação ao vetor de busca. O cálculo de similaridade entre o vetor de busca e um vetor de um documento representado é dado pelo co-seno do ângulo formado por estes dois vetores.

Segundo Styrman (2005, p. 8), o problema é que as técnicas de busca em texto são baseadas em unidades léxicas e não na estrutura semântica das sentenças. O entendimento de uma *string* de texto é um problema sem solução para uma máquina, porque uma correta interpretação de uma frase, como uma consulta, também é impossível para humanos sem o conhecimento do objetivo do usuário ou do contexto específico no qual a consulta deve ser tratada.

Informações sobre as técnicas de *stemming* e recuperação probabilística pode ser encontrada em Baeza-Yates e Ribeiro-Neto (1999).

## 3.2 Descrição e Recuperação de Imagens com Base em Atributos

A descrição de imagens com base em atributos consiste em descrever a imagem a partir de um conjunto de pares de atributo/valor como, por exemplo, fotógrafo (atributo) – o nome do fotógrafo (valor), título – o título, etc. A descrição de imagens com base em atributos geralmente é realizada por metadados. O termo metadado é usado em tecnologia da informação como informação que descreve recursos eletrônicos. Também é usado para se referir à informação processável por computadores. Além disso, é comumente definido como dado sobre dado, ou informação sobre informação. Esta informação pode ser relativamente simples, tais como autor, data de criação, título, etc., ou mais complexa, como a opinião de pessoas sobre um livro. A importância dos metadados reside no fato de que a informação neles contida é pesquisável. Recursos digitais como imagens se tornam virtualmente irrecuperáveis se não houver uma informação associada à imagem para possibilitar a indexação pelos mecanismos de busca. Daí a importância de representar a informação extraída do processo de documentação com a utilização de metadados. Segundo Alves e Souza (2007), os atributos de metadados têm o propósito primário de descrever, identificar e definir um recurso de informação com o objetivo de modelar e filtrar o acesso. Os metadados são importantes na organização, gestão e recuperação da informação digital, principalmente. Nesse sentido, são adotados procedimentos técnicos de catalogação, indexação e categorização dos conteúdos informacionais, o que possibilita a integração de fontes diversificadas e heterogêneas de informação.

Gilliland-Swetland (2000) estabelece cinco tipos de metadados:

- Administrativos - usados no gerenciamento e administração do recurso de informação. Exemplo: copirraites, informação de aquisição, etc.;
- Descritivos - usados para descrever ou identificar recurso de informação. Exemplo: palavras-chave, vocabulário controlado, anotações do usuário, etc.;
- De Preservação - relacionados ao gerenciamento de preservação do recurso de informação. Exemplo: condições físicas do recurso;

- Técnicos - relacionados ao funcionamento do sistema. Por exemplo, formato de arquivo, esquema de compressão de arquivo, etc.;
- De Uso - relacionados aos tipos de uso do recurso de informação. Por exemplo, registro de exposição do recurso (exemplo: exposição de fotografias), registro do uso (exemplo: imagem usada na publicação de um artigo de revista, jornal, etc.) e de usuário, etc.

Um dos metadados mais conhecidos na comunidade da Internet é o *Dublin Core* que é uma recomendação da W3C (*World Wide Web Consortium*), organismo que trata da padronização de protocolos e metadados na Internet. A importância da recomendação da W3C reside no fato de que a definição e a semântica dos atributos do metadado são aceitos mundialmente. O *Dublin Core* visa à descrição de recursos digitais tais como áudio, vídeo, imagens, páginas HTML, etc., para facilitar a descoberta da informação e a interoperação entre os sistemas. A seguir, os 15 elementos do *Dublin Core*.

Poderemos descrever a fotografia da Figura 1 (p. 20), com os atributos do metadado *Dublin Core* da seguinte forma:

- Título - Juscelino Kubitschek.
- Assunto - Juscelino Kubitschek, Presidente.
- Descrição - Juscelino Kubitschek de Oliveira (Diamantina, 12 de setembro de 1902 Resende, 22 de agosto de 1976) foi um médico, militar e político brasileiro. Conhecido como JK (lê-se jota-cá), foi presidente do Brasil entre 1956 e 1961. Foi casado com Sarah Kubitschek, pai de Márcia Kubitschek e pai adotivo de Maria Estela Kubitschek.
- Publicador - Wikipédia.
- Formato - jpg.
- Fonte - [http://pt.wikipedia.org/wiki/Juscelino\\_Kubitschek](http://pt.wikipedia.org/wiki/Juscelino_Kubitschek).
- Tipo - fotografia.
- Formato - jpg.
- Idioma - Português.

Quadro 1: Conjunto de atributos do metadado *Dublin Core*.

Elemento	Descrição
Título	Nome dado ao recurso.
Criador	Entidade originalmente responsável pela criação do conteúdo do recurso.
Assunto	Tema do conteúdo do recurso. Pode ser expresso em palavras-chave e/ou categoria. Recomenda-se o uso de vocabulários controlados.
Descrição	Relato do conteúdo do recurso. Exemplos: texto livre, sumário e resumo.
Publicador	Entidade responsável por tornar o recurso disponível.
Colaborador	Entidade responsável pela contribuição intelectual ao conteúdo do recurso.
Data	Data associada a um evento ou ciclo de vida do recurso.
Tipo	Natureza ou gênero do conteúdo do recurso. Exemplos: texto, imagem, som, dados, <i>software</i> .
Formato	Manifestação física ou digital do recurso. Exemplos: html, pdf, ppt, gif, xls.
Identificador	Referência não-ambígua (localizador) para o recurso dentro de dado contexto.
Fonte	Referência a um recurso do qual o presente é derivado.
Idioma	Língua do conteúdo intelectual do recurso.
Relação	Referência para um recurso relacionado.
Cobertura	Extensão ou escopo do conteúdo do recurso; pode ser temporal e espacial.
Direitos autorais	Informação sobre os direitos assegurados dentro e sobre o recurso.

Fonte: Alves e Souza (2007)

Uma forma de melhorar a descrição e, conseqüentemente, a recuperação, utilizando a técnica de descrição com base em atributos, é empregar um vocabulário controlado como AAT – *Art and Architecture Thesaurus*, ICONCLASS - Sistema de Classificação Iconográfica, para auxiliar a indexação e a recuperação. Adicionalmente, o sistema pode fazer uma associação do campo do metadado com a respectiva categoria do vocabulário e, com isso, fornecer um melhor suporte para indexadores humanos. A recuperação com base em atributos é realizada usando expressões booleanas que podem envolver vários atributos do metadado. Basicamente, a recuperação com base em atributos é um caso particular da recuperação com base em texto, pois, os valores dos atributos do metadado geralmente são baseados em linguagem natural.

### 3.3 Descrição e Recuperação de Imagens com Base em Ontologias

A descrição de imagens usando ontologias permite uma descrição mais completa. A descrição não é somente um conjunto de pares atributo/valor, como na abordagem por atributo, mas também com o uso de descrições envolvendo relações. A descrição de um carro pode incluir a descrição de seus componentes, por exemplo, motor, câmbio, etc. Os componentes são objetos que podem ser novamente descritos usando atributos como cor, tamanho, forma, etc. Segundo Styrman (2005), teoricamente a descrição de componentes pode ir até um nível de detalhe onde não resta mais subcomponentes específicos para descrever. Segundo Schreiber (2001) a descrição de imagens usando ontologia consiste basicamente em associar conceitos e instância à imagem. Com isso, a imagem herda não somente um termo, mas o termo acompanhado de sua descrição através de suas propriedades que o descreve; logo a descrição de imagens torna-se bem mais rica, do que a alternativa de descrição com base em atributos. Esta característica das ontologias é a motivação para a proposta deste trabalho que é o desenvolvimento de uma ontologia de descrição de imagens para, logicamente, descrever e recuperar imagens através de um sistema de indexação e recuperação.

Uma diferença fundamental que existe entre a descrição com base em atributo e a descrição com base em ontologia está na forma como são fornecidos os valores para os atributos do metadado. Na descrição com base em atributos, os valores são geralmente fornecidos em linguagem natural, enquanto que na descrição com base em ontologias os valores são obtidos da árvore de categorias da ontologia. O indexador inicia o processo de descrição a partir da seleção de uma categoria de alto nível e segue expandindo a hierarquia de conceitos até encontrar um conceito que melhor descreva a imagem; se não encontrá-lo, o indexador pode criar uma instância do conceito. Tipicamente, a criação de instância envolve o fornecimento de valores para as propriedades do conceito. A árvore de categorias pode fornecer informação intuitivamente clara para guiar o processo de descrição e recuperação. A descrição de imagens pode beneficiar-se da especificação formal da ontologia, dos mecanismos de restrição de valores e de cardinalidades que um atributo pode conter; desta forma, estabelece controles no processo de descrição de modo que seja processável por computadores.

A recuperação pode ser realizada usando a mesma árvore de categorias usada pelo indexador no momento da descrição da imagem. A combinação da técnica de recuperação com base em facetas com a recuperação com base em ontologias num sistema de recu-

peração de imagens possibilita a apresentação da ontologia na forma de hierarquias de conceitos de modo que permita ao usuário elaborar a expressão de pesquisa a partir da combinação de conceitos de várias categorias. Desta forma, o usuário pode elaborar sua expressão de busca dentro de um contexto fornecido pela hierarquias de conceitos. Além disso, as hierarquias de conceitos resolvem o problema da ambigüidade de termos, que é um problema recorrente em recuperação da informação.

## 4 Representação da Imagem

Atualmente existem dois paradigmas que abordam os problemas de representação de imagens: um com base em conceito e outro com base no conteúdo CBIR – *Content-Based Image Retrieval*. Na representação por conceito, as imagens são descritas por um conjunto de termos de indexação (conceitos). Um termo de indexação é geralmente uma palavra que representa um conceito ou significado presente na imagem. A maior contribuição vem da Ciência da Informação, com a utilização de modelos de descrição que permitem extrair as informações mais importantes contidas numa imagem. A representação por conceito é também conhecida como indexação manual, pois existe o papel do indexador, que é a pessoa que realiza a descrição da imagem, ou melhor, é o profissional que analisa a imagem e, com a utilização de um modelo, seleciona conceitos que irão representá-la. A introdução de uma pessoa no processo de descrição representa uma interferência no processo, pois entra em cena a subjetividade do indexador.

Na representação por conteúdo, as imagens são descritas com base em características de baixo nível da imagem, como textura, cor e arranjo de figuras geométricas contidas na imagem. A representação com base no conteúdo é também conhecida como indexação automática, uma vez que é realizada por algoritmos computacionais, de modo automático, dispensando a participação de uma pessoa (o indexador) no processo. A grande vantagem deste paradigma é a possibilidade de indexação de grandes quantidades de imagens de modo automático e num tempo bem mais rápido que a indexação manual, o que viabiliza a indexação de imagens na Internet com o uso desta abordagem. Sua grande desvantagem é que a descrição é realizada no nível sintático, o que dificulta a recuperação semântica de imagens. Este paradigma é abordado pela área de Processamento de Imagens, e não é o foco deste trabalho. A nossa proposta é a descrição de imagens por conceitos, logo, a descrição de imagens com base em características de baixo nível que exigem técnicas de Processamento de Imagens para descrevê-la está fora do escopo deste trabalho.

Este capítulo é uma revisão de literatura sobre a representação da imagem por conceito, que podemos subdividir em: representação do conteúdo visual da imagem e re-

presentação do conteúdo não visual. A representação do conteúdo visual está relacionada com as técnicas de transposição da informação presente na imagem para a informação verbal, ou seja, como transpor a informação iconográfica para informação verbal; a representação do conteúdo não visual está relacionada com a descrição da informação extrínseca à imagem, ou seja, informação que não está presente na imagem, mas que de alguma forma está relacionada a ela e que pode ser útil para melhorar a recuperação de imagens. Iniciaremos o nosso estudo a partir das técnicas de representação do conteúdo visual da imagem; em seguida descreveremos algumas categorias de representação do conteúdo não visual. Depois abordaremos como variáveis da técnica fotográfica podem contribuir para a representação da imagem.

## 4.1 Representação do Conteúdo Visual

A representação do conteúdo visual consiste em descrever a informação contida na imagem através de descritores, que são, geralmente, palavras que representam um conceito ou significado presente na imagem; eles são a ponte entre o usuário e a imagem. Representar a informação imagética por meio de descritores é uma tarefa que está longe de ser trivial. Este é um tema que tem demandado muitas pesquisas tanto em Ciência da Informação quanto em Processamento de Imagens.

O processo de representar a informação iconográfica através de descritores sempre resulta em perda da informação iconográfica, pois o que está sendo feito é, na verdade, uma tradução de linguagens – da linguagem icônica para a linguagem verbal – e, como em todo processo de tradução, existem interferências que resultam em perda da informação original. “A descrição de uma imagem nunca é completa” Smit (1989, p. 102), “pois, por mais que se privilegie um detalhamento minucioso na tentativa de dizer verbalmente o que se vê na imagem, sempre haverá algo a se perguntar sobre ela, algo que a pessoa que descreve desconhece, esquece ou que lhe passou despercebido” Manini (2002, p. 18). Isto implica que a indexação é realizada com base numa descrição que não corresponde exatamente à informação imagética, por isso há necessidade de utilizar técnicas de descrição que minimizem a distância entre a informação imagética e a descrição verbal.

A estrutura da informação imagética pode ser vista como um hipertexto, pois numa imagem podemos encontrar qualquer tipo de objeto e estes objetos possuem algum tipo de relacionamento semântico entre si que pode suscitar conceitos abstratos como poder, vitória, fracasso, etc. Quando realizamos a descrição por meio da seleção de conceitos es-

tamos tentando representar esta estrutura hipertextual de significados através de unidades léxicas (descritores) sem qualquer relação semântica entre si.

Portanto, as questões-chave da descrição de imagens são: o que indexar, e como indexar, ou como selecionar os descritores que irão representar a imagem e que darão acesso à mesma. Além disso, vários níveis de descrição são possíveis. Qual o nível de descrição da imagem, levando em consideração os diferentes tipos de usuários e a precisão do mecanismo de busca?

A representação por conceitos, como já dissemos, consiste na representação da imagem por um conjunto de termos (conceitos) selecionados com base naquilo que a imagem mostra. A análise de conteúdo de imagens teve um grande avanço a partir dos níveis de descrição de Erwin Panofsky, um historiador da arte, que estabeleceu três níveis para descrição de imagens artísticas: o pré-iconográfico, o iconográfico e o iconológico (PANOFSKY, 1991, p. 47).

Nível pré-iconográfico - a análise pré-iconográfica consiste numa descrição genérica dos objetos e eventos representados pela imagem. Somente o conhecimento obtido pela experiência prática do mundo é suficiente para a representação da imagem neste nível. Exemplo: o reconhecimento de um homem, de uma mulher, de uma criança, de um carro, etc., numa imagem.

Nível iconográfico - determina o assunto secundário ou convencional da imagem. Trata, em suma, da determinação do significado mítico, abstrato ou simbólico da imagem, sintetizado a partir dos elementos componentes desta, detectados pela análise pré-iconográfica. Ao contrário da análise pré-iconográfica, a análise iconográfica compreende a identificação de objetos específicos e da representação de temas e conceitos abstratos.

Nível iconológico - propõe uma interpretação do significado intrínseco do conteúdo da imagem. A análise iconológica constrói-se a partir das anteriores, mas recebe fortes influências do conhecimento que o analista possui do ambiente cultural, artístico e social dentro do qual a imagem foi gerada. Shatford (1986) argumenta que este nível de significado não pode ser descrito com algum grau de consistência.

A título de exemplo, uma imagem de um atleta correndo de braços abertos e quebrando a linha de chegada numa pista de atletismo pode ser analisada da seguinte forma: no nível pré-iconográfico, o que temos é um atleta correndo, linha de chegada e pista de atletismo, sendo que no nível iconográfico os elementos e ações – atleta correndo de braços abertos, linha de chegada e pista de atletismo – simbolizam vitória. A análise iconológica

contextualizaria a vitória de acordo com a realidade social e cultural do local e da época na qual a imagem foi realizada. Smit (1996) argumenta que a análise iconológica pode ser assimilada à elaboração de um modelo, ou uma teoria, a ser validada, baseada na análise da imagem, mas cujo objetivo a ultrapassa, uma vez que se encontra fora da imagem.

O nível pré-iconográfico é subdividido em significado fatural e expressivo. O significado fatural consiste na enumeração dos objetos e ações visíveis na imagem. No exemplo anterior, o atleta correndo de braços abertos, a pista de atletismo, a fita de chegada, etc., correspondem ao significado fatural, enquanto o significado expressivo é determinado não somente pela identificação, mas principalmente pela composição destes elementos (por exemplo, a composição dos elementos na imagem do exemplo anterior transmite a idéia de vitória).

Shatford (1986) questiona sobre a possibilidade de representar os assuntos da imagem para torná-la acessível e disponível a diferentes tipos de uso e usuário. A autora identifica e classifica os tipos de assunto que uma imagem pode ter com base nos conceitos de descrição de imagens de Panofsky. A partir dos níveis de descrição de imagens de arte propostos por Panofsky, Shatford (1986) expande a análise de imagens para um universo mais abrangente de tipos de imagens (fotografias, desenhos, pinturas, etc.). Para Shatford (1986), a imagem é DE alguma coisa ou SOBRE alguma coisa. A autora emprega estes dois aspectos para analisar os níveis pré-iconográfico e iconográfico esclarecendo que no nível pré-iconográfico o aspecto DE é a descrição genérica de objetos e eventos, enquanto que no nível iconográfico o aspecto DE é a identificação específica destes objetos e eventos; argumenta que a imagem é simultaneamente genérica e específica. Quanto ao aspecto SOBRE, no nível pré-iconográfico corresponde ao significado expressivo, e no nível iconográfico o SOBRE corresponde à identificação do significado simbólico, mítico ou de conceitos abstratos. Segundo Shatford (1986), o DE descreve pessoas, objetos, lugares, ações, que têm manifestação física, enquanto o SOBRE descreve emoções (amor, felicidade) e conceitos (honra, verdade).

As questões QUEM, O QUE, QUANDO e ONDE são empregadas juntamente com os aspectos DE Genérico, DE Específico e SOBRE, numa matriz 4 X 4, para extrair o conteúdo informacional da imagem (o que a imagem mostra). Cada uma das categorias QUEM, O QUE, QUANDO e ONDE pode ser subdividida dentro dos aspectos Genérico e Específico. Abaixo, o esquema de classificação de assuntos de imagens proposto por Shatford e traduzido por Smit.

Quadro 2: Categorias para a representação da imagem.

<b>Categoria</b>	<b>Definição geral</b>	<b>DE Genérico</b>	<b>DE Específico</b>	<b>SOBRE</b>
<i>QUEM</i>	Animado e inanimado, objetos e seres concretos.	Esta imagem é de quem? De que objetos? De que seres?	De quem, especificamente, se trata?	Os seres ou objetos funcionam como símbolos de outros seres ou objetos? Representam a manifestação de uma abstração?
	Exemplo	Ponte	Ponte das Bandeiras	Urbanização
<i>ONDE</i>	Onde está a imagem no espaço?	Tipos de lugares geográficos, arquitetônicos ou cosmográficos.	Nome de lugares geográficos, arquitetônicos ou cosmográficos.	O lugar simboliza um lugar diferente ou mítico? O lugar representa a manifestação de um pensamento abstrato?
	Exemplo	Selva	Amazonas	Paraíso - supõe um contexto que permita esta interpretação
	Exemplo	Perfil de cidade	Paris	Monte Olimpo - como o exemplo anterior
<i>QUANDO</i>	Tempo linear ou cíclico, datas e períodos específicos, tempos recorrentes.	Tempo cíclico.	Tempo linear.	Raramente utilizado, representa o tempo, a manifestação de uma idéia abstrata ou símbolo.
	Exemplo	Verão	1996	Esperança, fertilidade
<i>O QUE</i>	O que os objetos e seres estão fazendo? Ações, eventos, emoções.	Ações, eventos.	Eventos individualmente nomeados.	Que idéias abstratas (ou emoções) estas ações podem simbolizar?
	Exemplo	Jogo de futebol	Copa do Mundo	Esporte

Fonte: Smit (1996, p. 7).

As questões QUEM, ONDE, QUANDO e O QUE são empregadas para extrair o conteúdo da imagem. Representar a imagem empregando o método de Shatford consiste em observar a imagem e responder a estas perguntas nos níveis genérico e específico e, além disso, identificar o SOBRE com base nessas respostas. Abaixo, seguem as definições dos elementos do conteúdo informacional segundo Smit (1997).

Quadro 3: Descrição de elementos do conteúdo informacional.

QUEM	Identificação do “objeto focado”: seres vivos, artefatos, construções, acidentes naturais, etc.
ONDE	Localização da imagem no espaço: espaço geográfico ou espaço da imagem.
QUANDO	Localização da imagem no tempo: tempo cronológico ou momento da imagem p. ex.: junho de 1997 ou dia de verão.
COMO / O QUE	Descrição de atitudes ou detalhes relacionado ao ‘objeto focado’ quando este é um ser vivo (p. ex.: cavalo correndo, criança trajando roupa do século XVIII).

Fonte: Smit (1997, p. 4).

Sob o ponto de vista das necessidades de informação do usuário, a representação de imagens nos níveis genérico e específico encontra sustentação, pois, como afirma Shatford (1986), o usuário só pode formular suas necessidades informacionais nos termos daquilo que ele já conhece, ou seja, se o usuário só conhece o nível pré-iconográfico (genérico) de uma imagem, então ele não pode formular suas necessidades em termos iconográficos (específicos).

O modelo de Shatford (1986) é genérico o suficiente para ser aplicado a qualquer tipo de imagem, pois o que está sendo analisado para representação é o conteúdo da imagem. No contexto de representação de imagens fotográficas, Smit (1997) argumenta que há necessidade de levar em consideração, na representação da imagem, a expressão fotográfica.

Segundo Smit (1997), a bibliografia da área da informação preconiza o tratamento da fotografia exclusivamente pelo que esta mostra, ou seja, pelo seu conteúdo informacional, desprezando sua dimensão expressiva. Para Smit (1996), imagem = conteúdo informacional + expressão fotográfica. O conteúdo informacional pode ser entendido como o que a imagem mostra, enquanto dimensão expressiva ou expressão fotográfica é, segundo Manini (2002), algo ligado à forma da imagem e que se encontra em justaposição ao seu conteúdo informacional.

Segundo Smit (1996), a bibliografia aborda a fotografia no âmbito da documentação,

de formas muito variadas, mas ressalta, invariavelmente, que o conteúdo informacional da imagem (o que esta mostra) deve ser analisado justapondo-a a outra categoria de variáveis, a saber, os dados oriundos da geração da imagem fotográfica, tais como angulação, enquadramento, tempo de exposição, presença/ausência de cor, luminosidade, etc. Estes dados agrupados na categoria “expressão fotográfica” são freqüentemente associados à noção de “forma”, tornando a distinção entre “forma” e “conteúdo” da imagem muito tênue.

A partir da afirmação de Smit (1996) de que imagem = conteúdo informacional + dimensão expressiva, Manini (2002) faz uma adaptação do modelo de descrição de imagens de Shatford para o contexto de indexação de imagens fotográficas, incorporando a questão da Dimensão Expressiva (Expressão Fotográfica). A grade abaixo é conhecida como grade de análise documentária de imagens.

Quadro 4: Grade de Análise Documentária de Imagens.

	Conteúdo Informacional		Dimensão Expressiva
	DE	Sobre	
Categoria	Genérico	Específico	
Quem/O Que			
Onde			
Quando			
Como			

Fonte: Manini (2002, p. 105).

A coluna da Dimensão Expressiva é descrita com o uso de variáveis de recursos técnicos como enquadramento, luminosidade, profundidade de campo, ótica, posição da câmera fotográfica, direção da luz artificial, cromia, etc., como visto no Quadro 6 (p. 46). Para proceder à descrição da coluna Dimensão Expressiva, o indexador precisa conhecer o vocabulário de técnicas fotográficas e ter, até, alguma noção do que significa tais técnicas. A descrição da coluna da Dimensão Expressiva consiste numa enumeração de variáveis da técnica fotográfica presentes na imagem.

Conceitos relacionados à Dimensão Expressiva aparecem quando realizamos uma análise da descrição do conteúdo informacional com a descrição das variáveis da técnica fotográfica. Por exemplo, quando analisamos a fotografia de um político fotografado com o uso da técnica câmera baixa (ângulo de tomada de baixo pra cima), percebemos que a posição da câmera produz a idéia de um político poderoso e imponente. O conceito de político poderoso e imponente é a Dimensão Expressiva da fotografia, que surgiu da justaposição do conteúdo informacional com a técnica fotográfica.

Uma informação importante para ser levada em consideração ao se descrever uma fo-

tografia usando a técnica fotográfica é que toda imagem tem composição, enquadramento, luminosidade, profundidade de campo, uso de objetivas, tempo de exposição e posição da câmera; logo, todos estes recursos precisam ser analisados para tentar identificar quais as variáveis de cada recurso estão sendo empregadas. É importante que todos os recursos sejam analisados para se obter uma descrição mais precisa e, conseqüentemente, melhorar a eficiência do sistema de recuperação de imagens.

É necessário fazer as seguintes perguntas – entre outras – quando se estiver indexando a Dimensão Expressiva: Qual é a composição? Qual é o Enquadramento? Qual é a Posição da Câmera? Qual é o tempo de Exposição? Qual a Objetiva que está sendo usada? Qual é a corromia? É uma fotomontagem? Para cada questão aparecerá como resposta uma variável técnica.

Na Grade de Análise Documentária de Imagens, Quadro 4 (p. 33), Manini (2002) associa O QUE com QUEM, enquanto Smit (1997), associa O QUE com COMO. Para Manini (2002), a questão COMO está mais relacionada a ações (determinada por verbos), enquanto O QUE está mais para substantivos; por isso a autora associa a questão O QUE com QUEM. Segundo Shatford (1986), a questão QUE representa Que ação os objetos ou seres estão fazendo? Que emoções? Que eventos? Que atividades? Neste trabalho adotaremos a questão QUE, mais associada a questão COMO e de acordo com o modelo de Shatford (1986).

Vamos analisar a fotografia abaixo utilizando a Grade 4 (p. 33) com a alteração anterior (O Que/Como)



Figura 2: Dr. Peter Henry Rolfs.

Dr. Peter Henry Rolfs, o primeiro diretor da ESAV. Época 1932.  
Fonte: <http://www.cpd.ufv.br/museuhistorico>, 23/04/2005.

Quadro 5: Exemplo da aplicação da Grade de Análise Documentária de Imagens.

Categoria	Conteúdo Informacional		Dimensão Expressiva
	Genérico	Específico	
Quem	Homem	Peter Henry Rolfs	Retrato, Pose, Preto-e-Branco, Câmera lateral, Plano Médio, Objetiva Normal
Onde			
Quando		1932	
O Que/Como		Sentado, fumando, cachimbo, olhando para fotografia de uma mulher	

Palavras-chave: **Homem, Peter Henry Rolfs, Diretor da ESAV, História, ESAV, Sentado, Fumando, Cachimbo.** Termos relacionados à Dimensão Expressiva: **Retrato, Pose, Preto-e-Branco, Câmera Lateral, Plano Médio, Objetiva Normal.**

Além dos descritores relacionados ao conteúdo da imagem vários descritores da técnica fotográfica foram obtidos, complementando a descrição da imagem com dados associados à noção de forma.

Layne (1994) argumenta que a indexação de imagens deve fornecer acesso às imagens com base nos seus atributos e que, além disso, deve fornecer acesso não somente a imagens individuais, mas a grupos de imagens e esclarece que este agrupamento deve ser baseado nos atributos da imagem ou nos atributos do que é representado na imagem. A autora apresenta algumas razões para este argumento: primeiro, em algumas disciplinas, como história da arte, a justaposição e comparação de imagens que compartilham uma ou mais características são essenciais para o processo de pesquisa; segundo, um pesquisador de imagens pode não ser capaz de verbalizar todo o critério de busca da imagem desejada e, portanto, necessitará observar grupos de imagens que fornecerão critérios que podem ser verbalizados a fim de identificar precisamente a imagem desejada; e, terceiro, o pesquisador pode ter critérios muito específicos que podem ser mais eficientemente identificados por uma observação de um grupo de imagens do que descrições textuais detalhadas ou

palavras-chave.

Jørgensen (1996) realizou uma pesquisa sobre descrição de imagens com 82 participantes, com diferentes níveis de conhecimento e experiência, com o objetivo de identificar classes de atributos de imagens. Foi solicitado aos participantes realizarem a descrição do conteúdo de várias imagens em linguagem natural. A autora diferenciou três tipos de atributos: Perceptivo, Interpretativo e Reativo. Os Atributos Perceptivos são aqueles identificados em resposta direta ao estímulo visual, como objetos (pessoa), cor (vermelho), etc. Os Atributos Interpretativos são aqueles que requerem uma interpretação dos atributos perceptivos; Esta categoria inclui termos que designam idéias abstratas como vitória, poder, etc. Atributos Reativos descrevem reações e atitudes das pessoas na imagem. Segundo Jørgensen (1996), para determinar as classes de atributos, cada ocorrência do termo na descrição era codificado, ou seja, a presença do termo na descrição era indicativo do foco de atenção num determinado momento. Os atributos derivados da análise dos dados desta pesquisa foram classificados e definidos da seguinte forma:

- Objeto Literal (Perceptivo) - Esta classe contém itens que são classificados como objetos percebidos na imagem;
- Pessoas (Perceptivo) - a presença de forma humana;
- Qualidade das Pessoas (Interpretativo) - Interpretação do estado mental ou emocional, ou ocupação das pessoas mostradas na imagem;
- Informação Histórica da Arte (Interpretativo) - Informação que está relacionada com o contexto de produção da representação, como artista, estilo, meio e tipo;
- Cor (Perceptivo) - Inclui a identificação de cores e termos relacionados com vários aspectos da cor como, brilho, matiz, etc;
- Local (Perceptivo) - Inclui atributos relacionados com a localização geral e específica dos componentes na imagem;
- Elementos Visuais (Perceptivo) - Consiste na identificação de orientação, forma, componente visual (linhas, detalhes, luminosidade) ou textura de objetos;
- Descrição (Perceptivo) - Consiste em adjetivos descritivos e palavras relacionadas com a dimensão ou quantidade dos objetos;
- Conceitos Abstratos (Interpretativo) - Consiste em termos abstratos, temáticos ou simbólicos;

- Conteúdo Histórico (Interpretativo) - Atributos relacionados a um evento específico que está sendo mostrado;
- Reação de Pessoas - Reação de pessoas na imagem;
- Relacionamentos externos - Comparação de atributos dentro de uma imagem ou entre imagens ou referências à entidade externa.

O resultado da pesquisa de Jörgensen (1996) mostrou que as pessoas descrevem imagens em função de objetos e pessoas, devido ao alto índice de termos associados a estes dois itens.

Jaimes e Chang (2000) propõem uma estrutura de classificação de informação visual em níveis, Figura 3 (p. 38), que tenta unir os avanços das pesquisas em representação de imagens por conteúdo e por conceito. O modelo possui dez níveis de descrição, sendo que os quatro primeiros (tipo/técnica, distribuição global, estrutura local e composição global) se referem à descrição de imagens com base em atributos sintáticos, que podem ser extraídos automaticamente através de algoritmos computacionais; enquanto os níveis restantes (objetos genéricos, cena genérica, objeto específico, cena específica, objetos abstratos e cenas abstratas) se referem ao conteúdo semântico da imagem, cuja principal contribuição vem da Ciência da Informação.

A forma de pirâmide do modelo conceitual transmite a idéia de quanto conhecimento é necessário para descrever cada nível. O conhecimento exigido aumenta do topo para a base da pirâmide. Os níveis sintáticos ficam localizados na parte superior, uma vez que o conhecimento exigido para os descrever é mínimo. Quanto aos níveis semânticos, a quantidade de informação exigida para indexar cresce do nível correspondente ao objeto genérico em direção à base da pirâmide, ou seja, a cena abstrata é o nível que requer mais conhecimento, uma vez que é necessária uma interpretação dos objetos genéricos e específicos presentes na imagem.

Segundo Jörgensen et al. (2001), a pirâmide pode ser usada para organizar o conteúdo visual para recuperação, guiar o processo de indexação e classificar as descrições obtidas de forma manual e automática. Os dez níveis de descrição da pirâmide de classificação de informação visual estão agrupados dentro de duas categorias: sintático/perceptível e semântico/conceitual. Segundo Jaimes e Chang (2000), imagens são representações multidimensionais da informação. No nível mais básico elas simplesmente causam uma resposta sob a ação da luz. No nível mais complexo, entretanto, imagens representam idéias abstratas que dependem muito do conhecimento e da experiência de cada indivíduo.

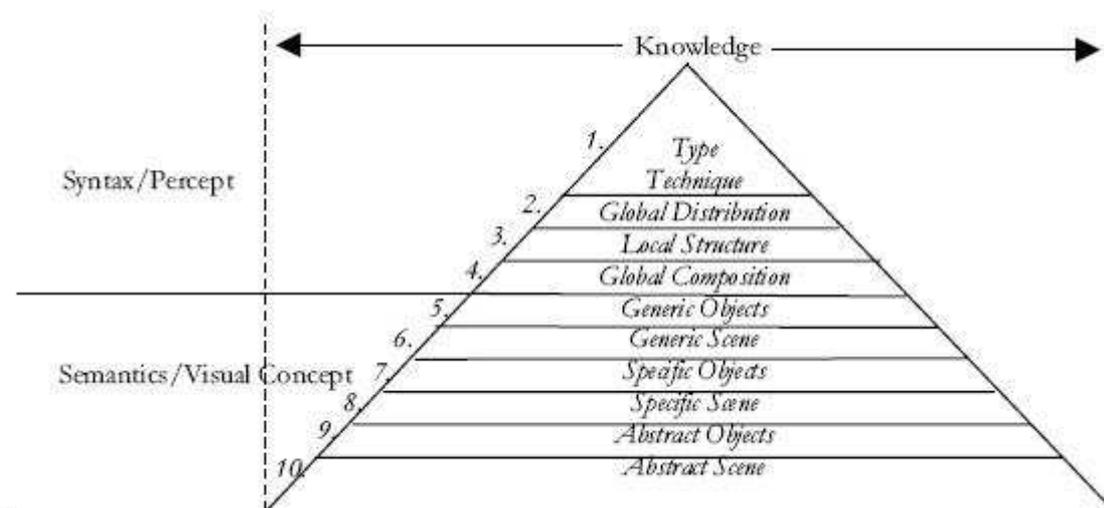


Figura 3: Pirâmide de classificação de informação visual de Jaimes e Chang (2000).  
 1º) *Type/Technique* – Tipo/Técnica 2º) *Global Distribution* – Distribuição Global 3º) *Local Structure* – Estrutura Local 4º) – *Global Composition* – Composição Global 5º) *Generic Objects* - Objetos Genéricos 6º) *Generic Scene* – Cena Genérica 7º) *Specific Objects* – Objetos Específicos 8º) *Specific Scene* – Cena Específica 9º) *Abstract Objects* – Objetos Abstratos 10º) *Abstract Scene* – Cena Abstrata.

Jaimes e Chang (2000) fazem uma distinção muito adequada entre percepção e conceito e entre sintático e semântico. A percepção refere-se àquilo que os nossos sentidos percebem em um sistema visual sob a luz. Estes padrões de luz produzem a percepção de diferentes elementos, tais como textura e cor. Nenhum processo de interpretação é necessário quando estamos nos referindo ao nível perceptível – nenhum conhecimento é requerido. O nível conceitual, por outro lado, refere-se a uma idéia abstrata ou genérica de alguma coisa e depende do conhecimento e da experiência de cada pessoa; além disso, tende a ser muito subjetivo.

De modo similar ao nível perceptível, que não requer qualquer interpretação, o nível sintático refere-se ao modo como os elementos visuais são arranjados numa imagem, sem levar em consideração a relação semântica entre tais arranjos. Já o nível semântico lida com os significados destes elementos e de seus arranjos.

Abaixo faremos uma descrição sucinta de cada nível da pirâmide. Vale lembrar que a estrutura de classificação também está sendo usada para descrição de áudio e vídeo, além das imagens estáticas.

Os níveis da pirâmide:

- tipo/técnica - é usado para descrever características gerais da imagem ou seqüência de vídeo. Descrição do tipo de imagem ou seqüência de vídeo ou técnica usada

para o produzir são muito gerais, mas de grande importância para a classificação de imagens. Por exemplo: pintura, fotografia, desenhos, etc.;

- distribuição global - destinado para descrição de características globais da imagem ou seqüência de vídeo. É descrito como o uso de características perceptíveis de baixo nível como cor global (cor dominante, histograma de cor, etc.), textura global (granularidade global, direcionalidade, contrastes, etc.), etc. Componentes individuais da imagem não são processados neste nível, ou seja, nenhuma forma é atribuída a estes elementos no sentido em que as medidas são tomadas globalmente. Atributos globais de baixo nível têm sido usados em vários sistemas de recuperação de conteúdo (por exemplo: QBIC WebSEEK, VIRAGE, etc.) e para organizar o conteúdo de uma base de dados para navegação;
- estrutura local - ao contrário da distribuição global, que não oferece qualquer informação sobre partes locais da imagem, a estrutura local está focalizada na extração de características de componentes da imagem. No nível mais básico, estes componentes resultam do processamento de baixo nível e incluem elementos como ponto, linha, cor, textura, etc. Os atributos obtidos de regiões locais têm uma grande aplicabilidade na indexação e recuperação de imagens médicas. Geralmente, raios-X e imagens microscópicas são tipos de imagens para as quais os detalhes da estrutura local da imagem têm grande importância para auxiliar no diagnóstico médico;
- composição global - este nível está interessado no arranjo específico dos elementos dado pela estrutura local. Mas o foco está na composição global, que consiste em analisar a imagem como um todo, usando os elementos básicos descritos anteriormente (linhas, círculos, quadrados, etc.). Neste nível não há conhecimento de objetos específicos. A imagem é representada por um conjunto de linhas, círculos, quadrados, etc.;
- objetos genéricos - consiste em representar objetos e seres vivos de um modo genérico. Apenas a experiência e o conhecimento cotidiano são suficientes para a representação deste nível, que corresponde ao nível de descrição pré-iconográfico de Panofsky (1991) e ao nível DE Genérico de Shatford (1986);
- cena genérica - diferente do nível objetos genéricos, que representa objetos individuais presentes na imagem, a cena genérica consiste na representação da imagem a partir da combinação de todos os objetos e de seus arranjos. Exemplos de cena genérica são: cidade, paisagem, praia, estilo de vida, etc. A informação necessária

para a indexação deste nível é mínima, bastando somente o conhecimento e a experiência de mundo. Não é necessária a identificação específica da rua, cidade, praia, etc.; apenas que a imagem se trata de uma praia, de uma cidade, etc.;

- objetos específicos - ao contrário do nível objetos genéricos, em que os objetos são representados de modo genérico, o nível objetos específicos requer a identificação específica dos objetos que estão presentes na imagem. Isto significa que este nível requer mais conhecimento do indexador. Corresponde ao nível de descrição iconográfico de Panofsky (1991) e ao nível DE Específico de Shatford (1986);
- cena específica - ao contrário da cena genérica em que a identificação da cena é representada de modo geral como praia, cidade, etc., o nível *cena específica* representa a cena de modo específico. Não basta saber que se trata de uma praia ou rua; é necessário identificá-la unicamente como Praia de Copacabana, Rua Castro Alves, etc.;
- objetos abstratos - este nível corresponde ao nível SOBRE de Shatford (1986). Logo, este nível exige uma interpretação do que os objetos representam na imagem. É o mais difícil de ser indexado, pois é muito subjetivo e depende da experiência e do conhecimento (repertório) de cada pessoa;
- cena abstrata - consiste na identificação do que a imagem representa como um todo. É também muito subjetivo. Exemplos deste nível incluem felicidade, tristeza, paraíso, etc.

Os quatros primeiros níveis recuperam os aspectos perceptíveis e sintáticos da imagem. A principal vantagem é que nenhum conhecimento é necessário para realizar a indexação; tudo é feito de modo automático. Entretanto, o usuário final não expressa suas necessidades de imagens em termos de atributos sintáticos como proporção de cor, textura, linhas, quadrados, etc., mas em termos de conceitos e de relações semânticas. Geralmente, eles querem uma imagem DE alguma coisa ou SOBRE alguma coisa. Os primeiros níveis são usados em recuperação de imagens por conteúdo. É importante ressaltar que a indexação e a recuperação de imagens por conteúdo apresenta bons resultados em alguns domínios de aplicação, como, por exemplo, na área de indexação e recuperação de imagens médicas, onde a característica de recuperação do sistema é a recuperação com base em similaridades de imagens. Os níveis semânticos da pirâmide de classificação de imagens têm grande influência da representação de imagem de Shatford (1986), mas a

classificação não contempla os aspectos COMO, ONDE e QUANDO, o que pode levar a uma representação pobre do conteúdo informacional da imagem.

Ao comparar as técnicas de representação de imagens, observa-se a presença dos níveis de descrição pré-iconográfico, iconográfico e iconológico nas várias representações, embora nomeados com termos diferentes. Shatford (1986) usa DE Genérico, DE Específico e SOBRE para representar respectivamente os níveis pré-iconográfico, iconográfico e iconológico, enquanto Jörgensen (1996) usa os termos Perceptivo e Interpretativo. Nota-se também esforços na tentativa de empregar num único modelo de descrição técnicas de representação de imagens por conceito e por conteúdo. Isto é visto claramente na classificação de recursos visuais de Jaimes e Chang (2000). Talvez seja o diferencial da pirâmide de classificação, uma vez que os paradigmas de indexação de imagens por conceito e por conteúdo são complementares. Enquanto a indexação de imagens por conceito opera no nível semântico, a indexação de imagens por conteúdo opera no nível sintático, logo, eles não se opõem.

## 4.2 Representação do Conteúdo Não Visual

O conteúdo não visual são aquelas informações extrínsecas ao conteúdo da imagem, isto é, que não estão presentes na imagem, mas que é importante representá-las na descrição devido à relevância que elas têm para a recuperação. A classificação de Layne (1994) fornece uma visão melhor dos tipos de atributos de imagem, que, segundo a autora, podem ser: Atributos Biográficos, Atributos de Conteúdo, Atributos de Exemplo e Atributos de Relacionamento.

- Atributos Biográficos - estão relacionados com o que podemos chamar de “biografia” da imagem, ou seja, com os dados da criação da imagem. Pode ser subdividido em duas categorias menores. Primeiro, existem aqueles atributos que estão ligados ao “nascimento” da imagem: criadores se eles são fotógrafos ou artistas; o momento e o local de sua criação e o nome ou título dado por seus criadores. Segundo, existem atributos que estão ligados com a “viagem” da imagem através do mundo: onde ela está agora, onde ela esteve, quem é o proprietário, quanto custa e se foi alterada de algum modo;
- Atributos de Conteúdo - é necessário considerar três aspectos de atributos de conteúdo: a imagem pode ser DE ou SOBRE alguma coisa; o conteúdo da imagem é simultaneamente genérico e específico; e, o conteúdo de uma imagem pode

ser classificado em quatro facetas: Tempo, Espaço, Atividades e Eventos, e Objetos, usado para representar tanto objetos inanimados como objetos animados;

- Atributos de Relacionamento - a imagem pode ser relacionada com imagens ou com documentos textuais, ou mesmo com objetos. Um documento textual pode estar relacionado com a imagem de diferentes modos. Por exemplo, uma imagem pode aparecer como ilustração num texto. Ao indexar os atributos de relacionamento pode ser importante não somente identificar a imagem ou documento ou objeto mas indicar a natureza deste relacionamento;
- Atributos de Exemplos - a imagem pode ser de um tipo particular, isto é, ela pode exemplificar ou ser exemplo de alguma coisa. Por exemplo, uma imagem pode ser uma gravura ou uma fotografia ou um pôster que é diferente da imagem que representa.

O conteúdo não visual da imagem com base nesta classificação compreende os atributos biográficos, os de relacionamentos e os de exemplos, excetuando os atributos de conteúdo que são intrínsecos à imagem, isto é, fazem parte do que a imagem mostra, logo faz, parte do conteúdo visual. No contexto de sistemas informatizados de indexação e recuperação de imagens, a classificação de metadados de Gilliland-Swetland (2000) pode complementar os atributos do conteúdo não visual. De acordo com Gilliland-Swetland (2000), os tipos de metadados para descrição de recursos digitais podem ser:

- Administrativos - usados no gerenciamento e administração do recurso de informação. Exemplo: copirraites, informação de aquisição, etc.;
- Descritivos - usados para descrever ou identificar recurso de informação. Exemplo: palavras-chave, vocabulário controlado, anotações do usuário, etc.;
- De Preservação - relacionados ao gerenciamento de preservação do recurso de informação. Exemplo: condições físicas do recurso;
- Técnicos - relacionados ao funcionamento do sistema. Por exemplo, formato de arquivo, esquema de compressão de arquivo, etc.
- De Uso - relacionados aos tipos de uso do recurso de informação. Por exemplo, registro de exposição do recurso (exemplo: exposição de fotografias), registro do uso (exemplo: imagem usada na publicação de um artigo de revista, jornal, etc.) e de usuário, etc.

Os metadados administrativos, técnicos, de preservação e de uso estão relacionados com o conteúdo não visual, enquanto os metadados descritivos estão relacionados ao conteúdo visual, isto é, com a descrição do conteúdo.

### 4.3 Técnica e Expressão Fotográfica

O contexto deste trabalho é a indexação e a recuperação de fotografias com base no conteúdo da imagem e na técnica fotográfica. Vimos anteriormente um modelo de descrição de imagens que emprega como descritores, variáveis da técnica fotográfica para representar a forma como o conteúdo da imagem é mostrado. Este capítulo demonstra a influência da técnica fotográfica na produção de sentido numa imagem fotográfica; além disso, descreve e ilustra algumas técnicas fotográficas que serão usadas na nossa proposta de descrição de imagens. A idéia de propor uma ontologia de descrição de imagens que represente o conteúdo visual da imagem, mais a sua forma a partir de variáveis da técnica fotográfica empregadas como descritores é talvez o diferencial da nossa proposta.

Segundo Souza (2002), barthesianamente, poderíamos considerar que entre os mais relevantes elementos potencialmente conferidores de sentido a uma mensagem fotojornalística se inscrevem o texto, insuflador de sentido à imagem, e os elementos que fazem parte da própria imagem, como a pose, a presença de determinados objetos, o embelezamento da imagem ou dos seus elementos, a trucagem, a utilização de várias imagens, etc. Mas temos ainda a considerar os elementos específicos da linguagem fotográfica, como a relação espaço-tempo, a utilização expressiva da profundidade de campo, da travagem do movimento e do movimento escorrido, etc.

A técnica fotográfica (enquadramento, tempo de exposição, posição da câmera, luminosidade, etc.) é a base para a construção da Dimensão Expressiva (Expressão Fotográfica). Sobre expressão e forma da fotografia, Lacerda (1993, p. 47) argumenta:

[...] a fotografia apresenta dois aspectos: imagem e objeto. Acrescentaríamos ainda um outro, estreitamente relacionado à imagem, e que diz respeito à sua expressão. Essa expressão seria a forma como uma imagem é mostrada, estando ligada a uma linguagem que lhe é própria e que envolve a técnica específica empregada, a angulação, o enquadramento, a luminosidade, o tempo de exposição, entre outros. Essas três dimensões do registro fotográfico – conteúdo, expressão e forma – é que constróem, em última instância, a mensagem que informa (LACERDA, 1993, p. 47).

O fotógrafo, no momento de um clique fotográfico, faz uso da linguagem técnica foto-

gráfica (*close*, ângulo, enquadramento, luminosidade, etc.) para ressaltar ou superestimar aquilo que quer mostrar em uma fotografia numa tentativa de externar suas percepções. É a este artifício que se denomina de dimensão expressiva da imagem. A dimensão expressiva sempre foi desconsiderada no contexto de sistemas de informação; no entanto, ela interfere muito no modo como o conteúdo informacional é mostrado.

O conteúdo informacional é recuperado a partir de perguntas realizadas à imagem: QUEM, QUANDO, COMO/O QUE e ONDE. A partir das respostas a estas questões é possível transpor as informações mais importantes da imagem para a representação verbal. No entanto, o usuário final está interessado tanto no conteúdo informacional quanto na forma como este conteúdo informacional é apresentado, ou seja, na dimensão expressiva da imagem.

A importância de se considerar a dimensão expressiva na Análise Documentária de Imagens está no fato de que o ponto decisivo da escolha de uma fotografia (a partir de um conjunto de imagens recuperadas num sistema de recuperação de informações visuais) pode estar justamente na forma como a mensagem imagética foi construída para transmitir determinado conteúdo informacional (MANINI, 2002, p. 88).

Como exemplo, um usuário faz uma pesquisa num sistema de informação de imagens fotográficas hipotético utilizando a palavra-chave Ashley Judd. O sistema recupera inúmeras fotografias (Figuras 4 a 7) que satisfazem o parâmetro de pesquisa, ou seja, que satisfazem o conteúdo informacional, mas o usuário seleciona somente uma delas, a última fotografia (Figura 7), aquela que lhe chamou mais a atenção pela forma como foi mostrada (*Close*, Retrato). Neste exemplo fica claro que aquilo que presidiu a escolha da fotografia foi algo além do conteúdo informacional: a dimensão expressiva da imagem teve um papel decisivo.



Figura 4: Ashley Judd 1

Fonte: <http://www.ashley-judd.com/images/displayimage.php?album=4&pos=11>



Figura 5: Ashley Judd 2

Fonte: [http://www.celebritywonder.com/picture/Ashley\\_Judd/ashleyjudd\\_035.html](http://www.celebritywonder.com/picture/Ashley_Judd/ashleyjudd_035.html)



Figura 6: Ashley Judd 3

Fonte: <http://www.ashley-judd.com/images/displayimage.php?album=4&pos=13>



Figura 7: Ashley Judd 4

Fonte: [http://www.celebritywonder.com/picture/Ashley\\_Judd/ashleyjudd\\_003.html](http://www.celebritywonder.com/picture/Ashley_Judd/ashleyjudd_003.html)

A Dimensão Expressiva da imagem pode contribuir para melhorar a recuperação de imagens, tornando os sistemas de recuperação mais precisos e eficientes, capazes de atender melhor às necessidades específicas dos usuários finais, através de filtros que empreguem os elementos da dimensão expressiva; é construída a partir da linguagem fotográfica, de recursos técnicos como luminosidade, enquadramento, angulação, *close*, fotomontagem, etc. Manini (2002, p. 91), a partir de Smit (1997), sugere um vocabulário de recursos técnicos e suas variáveis para a análise de imagens a partir de como a fotografia expressa seu conteúdo informacional.

Quadro 6: Recursos Técnicos.

Recursos Técnicos	Variáveis
<b>Cromia</b>	<b>preto e branco, colorida, etc.</b>
<b>Abstração</b>	<b>grafismo, etc.</b>
Efeitos Especiais	fotomontagem, estroboscopia, alto-contraste, <i>splash</i> , <i>panning</i> , trucagem, esfumação, etc
Ótica	utilização de objetivas: <i>fish-eye</i> , lente normal, grande angular, teleobjetiva, <b>objetiva 360º</b> , etc. Utilização de filtros: infravermelho, ultravioleta, etc.
Tempo de Exposição	instantâneo, pose, longa duração, <b>dupla exposição</b> , etc.
Luminosidade	luz diurna, luz noturna, contraluz, luz artificial, etc.
<b>Posição da Luz Artificial</b>	<b>frontal, lateral, baixa, alta, etc</b>
Enquadramento	enquadramento do objeto fotografado: vista parcial, vista geral, etc. Enquadramento de seres vivos: plano geral, médio, americano, close, detalhe, etc.
Posição da Câmera	câmera alta, câmera frontal, câmera lateral, câmera baixa, vista aérea, vista submarina, vista subterrânea, microfotografia eletrônica. Distância Focal fotógrafo-objeto: <b>pequena distância focal, longa distância focal, média distância focal, etc.</b>
Composição	<b>retrato individual, retrato coletivo</b> , paisagem urbana, paisagem natural, natureza morta, etc.
Profundidade de Campo	com profundidade de campo, sem profundidade de campo, etc.

Fonte: Manini (2002, p. 91), atualizada por Alex Sandro Santos Miranda (em negrito).

Os termos em negrito no Quadro 6 foram identificados a partir de uma pesquisa de suporte com os seguintes objetivos:

- identificar novas variáveis da técnica fotográfica;
- entender as variáveis técnicas;
- atualizar o vocabulário de recursos técnicos.

A pesquisa foi realizada colhendo-se informações de *sites* especializados em técnica fotográfica: [olhares.com](http://olhares.com)<sup>1</sup>, [foto@pt](http://foto@pt)<sup>2</sup>, que contribuíram bastante para o entendimento, identificação e ilustração das variáveis técnicas. O *site* [olhares.com](http://olhares.com) publica fotografias de profissionais para serem apreciadas por fotógrafos. Os profissionais comentam as fotografias publicadas, dando sugestões e críticas; além disso, a fotografia é acompanhada de uma descrição técnica: distância focal, tempo de exposição, velocidade do obturador, etc. Estas informações técnicas somadas aos comentários de outros fotógrafos possibilitaram a identificação de novas variáveis técnicas.

Com esta pesquisa, além de encontrar fotografias para ilustrar as variáveis técnicas, conseguimos identificar novos recursos e suas variáveis ainda não contempladas na classificação de recursos técnicos de Manini (2002) e de Smit (1997).

Na próxima seção faremos uma breve descrição das técnicas fotográficas e ilustraremos algumas variáveis com fotografias.

### 4.3.1 Efeitos Especiais

- *Panning* - é uma técnica muito utilizada pelos fotógrafos de esporte para maximizar a sensação de movimento. Segundo Ramalho e Palacin (2004), o efeito especial *panning* consiste justamente em regular o obturador da câmera para longa exposição e perseguir um determinado assunto (pessoa, carro, bicicleta, etc.) em deslocamento. Perseguir o assunto com a câmera faz com que ele fique parado em relação à câmera, fazendo com que todo o resto da fotografia apresente um borramento, realçando a sensação de movimento;
- *Splash* - em inglês, *splash* significa som ou ação de espirrar ou esguichar. Em fotografia, *splash* é uma técnica que consiste em registrar, numa tomada fotográfica, o momento exato do estouro de uma bexiga, de uma gota d'água espatifando no chão, etc. Atualmente, é muito usada em publicidade, principalmente de alimentos e bebidas.

---

<sup>1</sup><http://www.olhares.com>

<sup>2</sup><http://www.fotopt.net/index.asp>

Exemplos:



Figura 8: *Panning*

Fonte: <http://www.olhares.com/mosler/foto507097.html>, 24/04/2006.



Figura 9: *Splash*

Fonte: <http://www.olhares.com/oscar/foto544643.html>, 24/04/2006.

### 4.3.2 Ótica

Sobre ótica, o Quadro 6 (p. 46) classifica quanto ao uso de objetivas e quanto ao uso de filtros. O termo “objetivas” é um jargão usado pelos fotógrafos para denominar, na verdade, um conjunto de lentes. A câmera fotográfica utiliza um conjunto de lentes para focalizar a imagem captada no fotograma. O fotograma é o local onde a imagem é formada ou focalizada e é constituído de um material especial para formar a imagem.

As objetivas são classificadas em: objetiva normal, *fish-eye*, grande angular e teleobjetiva. Para entendermos a diferença entre elas é necessário saber o conceito de ângulo de visão de uma lente, também conhecido como ângulo de abertura ou campo de visão, que é a área captada pela objetiva.

- *Objetiva Normal* - tem um ângulo de visão de  $45^\circ$  e é correspondente ao olho humano, por isso é denominada objetiva normal, servindo como um ponto de referência

para classificar as objetivas;

- *Grande Angular* - tem um ângulo de visão superior à objetiva normal e é bastante utilizada em situações onde se quer fotografar a maior área possível e se está a pouca distância do assunto; por exemplo, interiores, automóveis, etc.;
- *Fish-Eye* - possui um ângulo de visão superior à grande angular, chegando aproximadamente a 180° de abertura. Uma característica da objetiva *fish-eye*, além da captação de um campo de visão próximo a 180°, é a produção de uma fotografia arredondada e distorcida nas bordas;
- *Teleobjetivas* - têm um ângulo de visão inferior à objetiva normal. O campo de visão é reduzido, mas, em compensação, a teleobjetiva é muito eficiente para fotografar objetos que estão a longa distância; por isso as teleobjetivas são as preferidas pelos fotógrafos de esportes;
- *Objetiva Macro* - na verdade, são objetivas que têm a distância focal<sup>3</sup> variável; deste modo, conseguem obter uma distância focal também conhecida como posição de macro, que permite fotografar objetos a curta distância. A objetiva macro é muito útil para fotografar objetos pequenos, tais como insetos e pequenas plantas.

Exemplos:



Figura 10: Grande angular

Fonte: <http://www.olhares.com/anjo/foto604937.html>, 23/04/2006

---

<sup>3</sup>Distância entre o centro óptico da lente e o plano focal da câmera fotográfica focada no infinito (é a maior distância, em relação à câmera, em que os raios incidem de forma paralela sobre a objetiva).



Figura 11: *Fish eye*

Fonte: <http://alexis.vannier.free.fr/fish6.jpg>, 23/04/2006.



Figura 12: Teleobjetiva

Fonte: <http://www.olhares.com/reza/foto93612.html>, 23/04/2006.



Figura 13: Macro

Fonte: [http://www.olhares.com/tb\\_mereco\\_um\\_pouco\\_de\\_descanso/foto609671.html](http://www.olhares.com/tb_mereco_um_pouco_de_descanso/foto609671.html), 23/04/2006

A importância das objetivas na produção da fotografia torna-se evidente. O que difere uma imagem produzida com diferentes objetivas é a perspectiva. Na produção de fotografias de lances de futebol, por exemplo, torna-se indispensável o uso de teleobjetivas, pois elas têm um grande poder de aproximação do assunto; com isso é possível fazer o recorte espacial do assunto a ser fotografado, evitando que muitos elementos indesejáveis acompanham a fotografia. Nesta situação dos lances de futebol, fotografar com o uso de objetiva normal ou grande angular certamente produzirá uma imagem com muitos elementos e, além disso, os elementos aparecerão de forma pequena e distante.

Para fotografar interiores de casas, igrejas, automóveis, etc., onde o fotógrafo tem pouco espaço para se distanciar do assunto, é indispensável o uso de objetiva grande angular, uma vez que o seu campo de visão é superior a 45°. O amplo campo de visão da grande angular permite que o fotógrafo chegue próximo ao assunto a ser fotografado e tire uma fotografia em que o assunto preencha toda a imagem. Com o uso de uma objetiva normal ou uma teleobjetiva não seria possível fazer o recorte espacial completo do assunto devido à proximidade do fotógrafo em relação ao assunto e ao seu estreito campo de visão.

A finalidade do uso de objetivas é focalizar do melhor modo possível o tema a ser fotografado e, dependendo da distância entre o fotógrafo e o assunto, usar-se-á uma ou outra objetiva; conseqüentemente, a qualidade da fotografia está diretamente relacionada ao uso da objetiva adequada.

### 4.3.3 Tempo de Exposição

O tempo de exposição é o tempo em que o obturador da câmera fotográfica fica aberto para a captação da imagem, durante o clique fotográfico. É neste momento que a imagem é registrada no negativo. O tempo de exposição varia entre milésimos de segundo (exemplo: 1/1000 s) até alguns minutos. Quanto menor for o tempo de exposição, mais congelada a imagem fica. O efeito especial *splash*, que veremos adiante, é conseguido com um tempo de exposição extremamente pequeno, da ordem de 1/10000 s. Para congelarmos a imagem de automóveis em movimento, dependendo da velocidade em que se encontram, é necessário usar um baixo tempo de exposição. Neste caso, se não usarmos o tempo de exposição correto, a imagem apresentará o efeito de arraste, cuja característica principal é a produção de uma imagem borrada.

O Quadro 6 (p. 46), de recursos técnicos classifica, o tempo de exposição em instantâneo, pose e longa duração:

- *Instantâneo* - corresponde àquelas fotografias produzidas com baixo tempo de exposição. Geralmente são fotografias de pessoas, carros ou objetos quaisquer em movimento, que, com o uso da técnica tempo de exposição baixo, tornam-se congeladas ou estáticas;
- *Pose* - uma característica da pose é que a pessoa que se expõe para a fotografia fica parada esperando pelo clique fotográfico. Por isso, basta um pequeno ajuste no tempo de exposição para evitar o borramento provocado por algum movimento involuntário da pessoa fotografada;
- *Longa Exposição* - em algumas situações é interessante fotografar objetos em movimento com longo tempo de exposição para obter o efeito de arraste, que reforça a sensação de movimento;
- *Dupla Exposição* - esta técnica consiste em fazer uma tomada fotográfica e, em seguida, sem adiantar o filme, realizar outra tomada fotográfica, ou seja, realizamos duas fotografias sobre o mesmo negativo. A fotografia realizada com esta técnica apresenta duas imagens sobrepostas.

Exemplos:



Figura 14: Instantâneo

Fonte: [http://www.olhares.com/isto\\_nao\\_e\\_uma\\_bicicleta/foto475843.html](http://www.olhares.com/isto_nao_e_uma_bicicleta/foto475843.html), 22/04/2006.



Figura 15: Longa exposição

Fonte: [http://cfesas.no.sapo.pt/cf\\_fotos2.html](http://cfesas.no.sapo.pt/cf_fotos2.html), 22/04/2006.



Figura 16: Dupla exposição

Fonte: [http://www.herbario.com.br/fotoweb/dupla\\_expos.htm](http://www.herbario.com.br/fotoweb/dupla_expos.htm), 22/04/2006.

### 4.3.4 Luminosidade

A luminosidade é uma variável muito importante na produção e recepção da fotografia. O usuário final pode estar interessado em uma fotografia sobre um determinado assunto ou sobre um determinado objeto, mas gostaria que a fotografia fosse produzida com a luz do dia. O Quadro 6 (p. 46) de recursos técnicos, sistematiza a iluminação da seguinte forma:

- *Luz Noturna* - corresponde às fotografias que são produzidas à noite;
- *Luz Diurna* - corresponde àquelas fotografias que foram produzidas durante o dia. As fotos são geralmente claras e apresentam um bom contraste;
- *Contraluz* - segundo Ramalho e Palacin (2004), a contraluz ocorre quando a iluminação principal está por trás do assunto;
- *Luz Artificial* - são as fotografias produzidas com o uso de *flash* ou refletores. Outro aspecto importante a ser observado é a direção da luz. Segundo Ramalho e Palacin (2004), alterando-se a posição da fonte de luz causamos profundas mudanças na foto.

Quanto à direção da fonte de luz, ela pode ser classificada da seguinte forma:

- *Iluminação Frontal* - é aquela onde a fonte de luz e a câmera estão na mesma direção em relação ao objeto a ser fotografado. Essa iluminação valoriza a largura e a altura do objeto iluminado, porém não tem impacto na demonstração da profundidade;
- *Iluminação Lateral* - com a luz atingindo o assunto lateralmente, sua forma e volume são ressaltados, assim como sua profundidade;
- *Iluminação Baixa ou Alta* - segundo Ramalho e Palacin (2004), colocar a fonte de luz bem abaixo ou acima do assunto causa a criação de sombras e contornos que podem criar um ar de mistério ou ressaltar volumes e formas.

Exemplos:



Figura 17: Luz noturna

Fonte: <http://www.olhares.com/foto575660.html>, 23/04/2006.



Figura 18: Luz diurna

Fonte: <http://www.olhares.com/frutaria/foto488366.html>, 22/04/2006.



Figura 19: Contraluz

Fonte: [http://www.olhares.com/a\\_disputa/foto328121.html](http://www.olhares.com/a_disputa/foto328121.html), 22/04/2006.

### 4.3.5 Enquadramento

O enquadramento corresponde ao recorte espacial da realidade representada na fotografia. É realizado pelo fotógrafo com a variação do campo de visão obtido através do ajuste da distância focal da objetiva. O enquadramento pode também ser feito depois da obtenção da fotografia através da manipulação da imagem no laboratório ou por meio de *software* de edição de imagens. Este processo é conhecido como reenquadramento e é muito utilizado pelos fotojornalistas para concentrar a atenção do observador no motivo principal da fotografia. Desta forma é possível eliminar os elementos que desviam a atenção do observador do motivo principal da imagem.

Segundo Souza (2002), o enquadramento concretiza-se no plano. A fotografia é uma unidade de significação precisamente porque se consubstancia num plano. Embora as denominações e as tipologias dos planos sejam variáveis, consoante os autores, podemos considerar essencialmente a existência de quatro tipos de planos, com efeitos diferentes ao nível da expressividade fotográfica:

- *Planos Gerais* – são planos abertos, fundamentalmente informativos, e servem, principalmente, para situar o observador, mostrando uma localização concreta. São muito usados para fotografar paisagens e eventos de massas (as pessoas podem diluir-se no conjunto, mas podem também parecer personagens coletivas, com personalidade, forma e peso). Os planos gerais também podem servir, por exemplo, para fotografias em que o próprio cenário é a “personagem” (como o peso dos arranha-céus sobre as pessoas);
- *Planos de Conjunto* – planos gerais mais fechados, onde se distinguem os intervenientes da ação e a própria ação com facilidade e por inteiro;
- *Planos Médios* – servem para relacionar os objetos/sujeitos fotográficos, aproximando-se de uma visão “objetiva” da realidade; pode-se considerar um plano médio mais aberto como um plano de três quartos ou plano americano; pode-se também considerar um plano médio mais fechado como um plano próximo;
- *Grandes Planos* – enfatizam particularidades (um rosto, uma janela, etc.), sendo freqüentemente mais expressivos que informativos, embora também sejam menos polissêmicos que os planos gerais, já que estes últimos possuem mais elementos para consumo do observador. Quando o grande plano é muito fechado, denomina-se muito grande plano ou plano de pormenor.

Exemplos:



Figura 20: Plano médio

Fonte: <http://www.olhares.com/foto607928.html>, 23/04/2006.



Figura 21: Plano americano

Fonte: <http://www.olhares.com/foto602198.html>, 23/04/2006.



Figura 22: *Close* ou Grande plano

Fonte: <http://www.olhares.com/foto609031.html>, 23/04/2006.



Figura 23: Detalhe ou Plano pormenor

Fonte: <http://www.olhares.com/galerias/>, 23/04/2006.

Embora os tipos de enquadramento que acabamos de apresentar possam ser usados tanto para seres vivos quanto para objetos de um modo geral, com algumas alterações, o Quadro 6 (p. 46) faz distinção entre enquadramento de seres vivos e enquadramento de objetos fotografados.

### 4.3.6 Posição da Câmera

O ângulo de tomada ou a posição da câmera fotográfica é outro elemento importante na geração de sentido da fotografia. Através deste recurso podemos valorizar ou desvalorizar o motivo fotografado. Por exemplo, podemos retratar a idéia de um político fraco e impotente numa fotografia, simplesmente realizando uma fotografia usando, no momento da tomada fotográfica, uma câmera alta, ou seja, um ângulo de tomada de cima pra baixo. A recíproca também é verdadeira: se quisermos transmitir a idéia de um político poderoso e imponente podemos fotografá-lo usando uma câmera baixa, ou seja, ângulo de tomada de baixo pra cima. Vejamos:

- *Câmera Frontal* - a tomada da fotografia faz-se paralelamente à superfície do solo e de frente para o objeto representado, oferecendo uma leitura objetiva da realidade representada na fotografia;
- *Câmera Baixa* - a tomada da fotografia é realizada de baixo para cima; produz uma imagem que valoriza o assunto fotografado. Alguns autores denominam de plano contrapicado;
- *Câmera Alta* - a tomada da fotografia é feita de cima pra baixo. Ao contrário da câmera baixa, a câmera alta desvaloriza o motivo fotografado. Alguns autores denominam de plano picado;
- *Vista Aérea* - geralmente a tomada da fotografia é realizada com a ajuda de helicóptero ou avião. A vista aérea apresenta como característica uma visão geral do lugar fotografado e é facilmente notável que a imagem foi realizada com a ajuda de um transporte aéreo;
- *Vista Submarina* - geralmente são imagens realizadas no fundo do mar;
- *Vista Subterrânea* - geralmente são imagens realizadas dentro de grutas, cavernas, minas, etc.;

- *Distância Focal* - não é a distância focal da objetiva à que o Quadro 6 (p. 46) se refere. É a distância entre o fotógrafo e o motivo fotografado. Pode ser dividido em: pequena distância focal (o fotógrafo está bem perto do objeto fotografado), média distância focal, longa distância, etc.



Figura 24: Câmera baixa

Fonte: [http://www.olhares.com/\\_se\\_eu\\_pudesse\\_falar/foto523550.html](http://www.olhares.com/_se_eu_pudesse_falar/foto523550.html), 22/04/2006.



Figura 25: Câmera alta

Fonte: <http://www.olhares.com/foto862158.html>, 29/10/2006.



Figura 26: Vista aérea

Fonte: <http://www.ufv.br>, 23/04/2006.

### 4.3.7 Composição

Segundo Souza (2002, p. 80), entra-se no domínio da composição quando se fala da disposição dos elementos na imagem, tendo em vista a obtenção de um efeito unificado, que, em princípio, é a transmissão de uma idéia ou de uma sensação. Ou seja, entra-se no domínio da composição quando se fala da informação que é acrescentada ao enquadramento, quando se fala dos elementos da fotografia e da forma como esses elementos competem pela atenção do leitor.

A regra dos terços é uma forma clássica de compor imagens fotográficas e pictóricas. Consiste em dividir de modo imaginário a imagem em terços horizontais e verticais de modo que forme nove terços. Os pontos de interseção das linhas horizontais e verticais são pontos de atração visual da fotografia devendo ser usados para colocar o motivo principal da fotografia. A figura abaixo ilustra os pontos de atração visual segundo a regra dos terços.

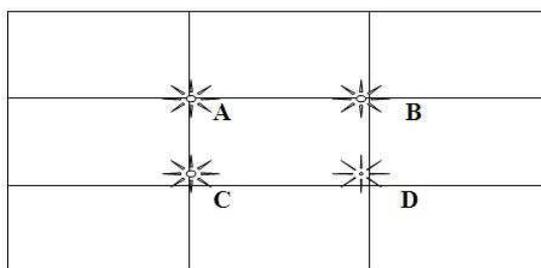


Figura 27: Regra dos terços

É possível, através da regra dos terços, estabelecer uma hierarquia dos motivos contidos numa imagem, embora não seja aconselhável produzir uma imagem que contenha muitos elementos, pois isto pode fazer com que o motivo principal fique perdido dentre os outros elementos.

O motivo principal da fotografia seguinte, o homem, está sob a linha vertical com o rosto localizado exatamente num dos pontos de atração visual, enquanto o cachorro, motivo secundário, está localizado em outro ponto de atração visual, o terço do canto direito. Desta forma o fotógrafo estabeleceu uma composição bem equilibrada entre os motivos fotografados. Este equilíbrio é conhecido como balanço. Outro detalhe que vale observar na fotografia é a linha do horizonte que coincide exatamente com a linha horizontal superior da regra dos terços.



Figura 28: Exemplo da aplicação da regra dos terços

Fonte: [http://www.deec.uc.pt/exposicoes/topic\\_temas.html](http://www.deec.uc.pt/exposicoes/topic_temas.html), 19/11/2006

Além da regra dos terços, segundo Souza (2002), existem regras de composição que devem ser seguidas para se obter uma boa fotografia; dentre elas, podemos citar: simplicidade, simetria, balanço, linhas, etc., como abaixo:

- Simplicidade - as imagens simples são as melhores. Uma fotografia deve ter apenas o essencial para realçar o tema escolhido. Numa imagem não devem aparecer muitos elementos, pois ela se torna confusa e distrai o observador do tema principal;
- Simetria - ao se fotografar um objeto simétrico, é considerado uma boa prática colocar a imagem no centro do visor. Por exemplo, a torre Eiffel;
- Linhas - as linhas horizontais transmitem a idéia de serenidade. Uma linha diagonal sugere dinamismo e movimento. Objetos alinhados segundo uma linha diagonal criam a ilusão de profundidade. Algumas linhas podem guiar o olhar do observador para o tema principal da fotografia.

O Quadro 6 (p, 46) classifica a composição de fotografias em: retrato, paisagem e natureza morta.

- Retrato - segundo Souza (2002), retratar alguém consiste em procurar não apenas mostrar a faceta física exterior da pessoa ou do grupo em causa, mas também em evidenciar um traço da sua personalidade (individual ou coletiva, respectivamente). A expressão facial é sempre muito importante no retrato, já que é um dos primeiros elementos da comunicação humana. Pode-se distinguir dois tipos de retrato: o individual e o de grupo ou coletivo. Embora seja mais comum se retratar seres humanos, os animais também compõem retratos.
- Natureza Morta - o nome natureza-morta é dado a uma pintura ou fotografia de um grupo de objetos inanimados, tais como frutas, animais abatidos (comumente

usados no preparo de pratos de culinária: peixes, aves, crustáceos em geral), objetos usados dentro de casa (vasos, bules, pratos, copos, etc.).

- Paisagem - são fotografias de ambientes naturais ou urbanos.

### 4.3.8 Profundidade de Campo

A profundidade de campo é um recurso técnico usado para selecionar o foco da fotografia. Ao fotografar pessoas em um local que tenha muitos elementos na cena, é interessante usar pequena profundidade de campo para desfocar o fundo da fotografia e, deste modo, destacar a pessoa fotografada. Por outro lado, ao fotografar paisagens, é indispensável que todos os elementos da cena estejam focados; neste caso, usa-se a maior profundidade de campo possível. Segundo o quadro de recursos técnicos, podemos ter uma fotografia:

- *Com profundidade de campo* - todos os planos ficam nítidos na fotografia;
- *Sem profundidade de campo* - alguns elementos aparecem nítidos enquanto outros aparecem desfocados.

Exemplos:



Figura 29: Com profundidade de campo

Fonte: <http://www.olhares.com/frutaria/foto488366.html>, 22/04/2006.



Figura 30: Sem profundidade de campo

Fonte: [http://www.olhares.com/zig\\_002/foto502957.html](http://www.olhares.com/zig_002/foto502957.html), 22/04/2006.

## 5 Ontologias

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. Aristóteles apresenta categorias que servem de base para classificar qualquer entidade e introduz ainda o termo “*differentia*” para propriedades que distinguem diferentes espécies do mesmo gênero. A conhecida técnica de herança é o processo de mesclar *differentias* definindo categorias por gênero. O termo 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).

Segundo Gruber (1993), 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. 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.

O problema destas interpretações está relacionado com o fato dos vocabulários controlados, tesouros e glossários de termos serem construídos com base numa especificação com o uso de linguagem natural; logo, são definições ambíguas para máquinas, ou seja, não são processáveis por computador. De acordo com a definição de Borst (1997), que veremos em seguida, os vocabulários controlados, tesouros e glossários de termos não podem ser

considerados como ontologias, pelo fato de não serem processáveis por computador.

Para a finalidade deste trabalho adotaremos a definição de ontologia proposta por Borst (1997, p. 12, *apud* Almeida 2003) “é uma especificação formal e explícita de uma conceitualização compartilhada”. Em outras palavras, é uma estrutura conceitual de um domínio, especificada em linguagem formal, isto é, processável por computador. “Especificação explícita” diz respeito a conceitos, propriedades, relações, funções, restrições, axiomas, explicitamente definidos; “compartilhado” quer dizer conhecimento consensual; e “conceitualização” diz respeito a um modelo abstrato de algum fenômeno do mundo real. Segundo Gruber (1993), o corpo do conhecimento representado formalmente é baseado numa conceitualização: os objetos, conceitos e outras entidades que são assumidas existir em alguma área de interesse e os relacionamentos entre eles. Uma conceitualização é um visão abstrata e simplificada do mundo que queremos representar para alguma finalidade. Cada ontologia define um conjunto de classes, propriedades, relações e objetos constantes para algum domínio do discurso, e inclui axiomas para restringir a interpretação.

A definição de ontologia proposta por Borst (1997) restringe o universo do que pode ser considerado uma ontologia, a partir dos termos “formal” e “compartilhada”. Para Borst (1997) é necessário que uma ontologia seja processável por computador e que, além disso, a especificação seja consensual. A Figura 31 foi empregada por McGuinness (2001) para ilustrar as interpretações potenciais do termo ontologia com base na definição de Gruber (1993). Lendo o gráfico da esquerda para direita, notamos a presença de glossários de termos, vocabulários controlados, tesouros e hierarquias informais, que podem ser interpretados como ontologia com base na definição de Gruber (1993), mas que, de acordo com a definição de Borst (1997), não podem ser considerados ontologias. Do lado direito da linha vermelha encontramos algumas propriedades que segundo McGuinness (2001) são necessárias para considerar alguma coisa como uma ontologia – hierarquia formal com base na relação *is-a*, restrições de valores, herança de propriedades, instância formal, etc.

## 5.1 Ontologias e Usos

A literatura é rica na descrição de ontologias e suas finalidades. A aplicação geralmente está relacionada de alguma forma com a comunicação entre sistemas ou pessoas, além de ser muito referenciada a questão da reusabilidade do conhecimento que a ontologia incorpora. Uschold e Gruninger (1996) subdividem os tipos de usos em: comunicação entre pessoas e organizações, interoperabilidade entre sistemas, compartilhamento do

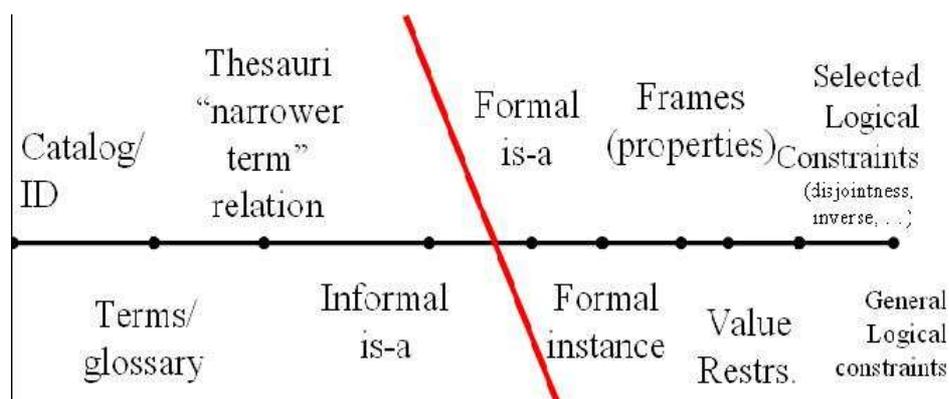


Figura 31: Espectro de ontologias.

Fonte: McGuinness (2001)

conhecimento e reusabilidade.

- Comunicação entre pessoas - ontologias reduzem a confusão conceitual e terminológica por meio de uma estrutura conceitual unificadora; logo, possibilita o entendimento e a comunicação entre pessoas com diferentes vocabulários e necessidades. Em pesquisa científica, por exemplo, podem existir vários paradigmas que estudam o mesmo problema, porém com jargões diferentes e especializados. Uma maneira de estabelecer uma melhor forma de troca de informações entre os pesquisadores destes paradigmas é por meio do desenvolvimento de uma ontologia. Assim, os pesquisadores irão dispor de um vocabulário comum para auxiliar a troca informações e, com isso, facilitar o desenvolvimento científico;
- Interoperabilidade entre sistemas - interoperabilidade é, segundo Heffin e Hendler (2000), a dificuldade de integrar recursos que foram desenvolvidos usando diferentes vocabulários e diferentes perspectivas dos dados. Muitas aplicações de ontologias são destinadas ao problema de interoperabilidade. Segundo Uschold e Gruninger (1996), ontologias podem ser usadas para integrar repositórios de dados existentes através da padronização da terminologia entre diferentes usuários e ferramentas de *software* ou pelo fornecimento de um fundamento semântico para proceder às traduções entre diferentes usuários. Existe a necessidade de criar um ambiente para melhorar a busca de informação em sistemas distribuídos. Além disso, procura-se melhorar a precisão dos mecanismos de busca por meio da indexação semântica dos dados;

- Compartilhamento do conhecimento e reusabilidade - refere-se ao fato de que, quando uma ontologia é formalmente especificada, é útil torná-la pública para que outras pessoas ou *softwares* possam reusá-la em suas aplicações, evitando, deste modo, a “reinvenção da roda”.

A Figura 32 ilustra um ambiente de comunicação entre agentes com linguagens distintas; neste caso, existem quatro linguagens rotuladas como L1, L2, L3 e L4. Para o estabelecimento de uma comunicação mútua entre os agentes são necessários  $O(n^2)$  tradutores, ou seja, 16 tradutores para este caso específico com 4 linguagens.

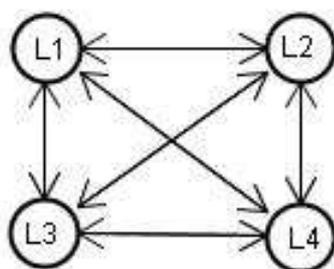


Figura 32: Exemplo de comunicação entre 4 linguagens.

Fonte: Uschold e Gruninger (1996)

Quando o número de linguagens aumenta, o número de tradutores cresce de modo quadrático, o que torna a rede de comunicação muito complexa. Isto pode ser resolvido com a eliminação da confusão conceitual e terminológica, por meio da definição de uma estrutura conceitual compartilhada, ou seja, por meio de uma ontologia que pode funcionar com uma estrutura conceitual unificadora para diferentes contextos. Basicamente, ontologias representam o conhecimento de um domínio que pode servir como uma conceitualização compartilhada capaz de viabilizar a comunicação entre pessoas ou *software* que possuem diferentes perspectivas da informação. Desta forma, ontologias agem como uma interlíngua. O exemplo a seguir, descrito em Uschold e Gruninger (1996), mostra uma ontologia fazendo um papel de uma interlíngua. A Figura 33 representa um ambiente com quatro linguagens distintas L1, L2, L3 e L4 e uma ontologia integrando todas elas. As linguagens podem ser vistas não somente como linguagem natural, mas como modelos de representação de domínios, modelos de banco de dados, protocolos de comunicação, linguagens de programação, ou seja, o termo linguagem aqui está sendo usado num sentido amplo.

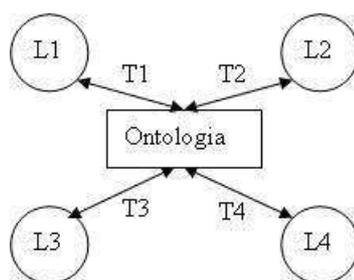


Figura 33: Exemplo de ontologia como interlíngua.

Fonte: Uschold e Gruninger (1996)

De acordo com a Figura 33 para estabelecer a comunicação entre duas línguas diferentes com o uso de uma ontologia, fazendo o papel de uma interlíngua, são necessários apenas dois tradutores entre cada linguagem e a ontologia; isto significa que para o estabelecimento de uma comunicação entre  $n$  linguagens são necessários apenas  $O(n)$  tradutores, ao invés de  $O(n^2)$  tradutores, situação que acontece quando não há um vocabulário comum. Neste exemplo, ontologias auxiliam na tradução de linguagens reduzindo o número de tradutores para  $O(n)$ .

Noy e McGuinness (2003) apresentam algumas razões para construir ontologias:

- Compartilhar entendimento comum da estrutura de informação entre pessoas ou agentes de *software*;
- Habilitar o reuso do conhecimento do domínio;
- Separar o conhecimento do domínio do conhecimento operacional;
- Analisar o conhecimento do domínio.

## 5.2 Tipos de Ontologias

Algumas propostas de classificação de ontologias são baseadas no grau de formalismo, função, estrutura e conteúdo. Quanto ao grau de formalismo, Uschold e Gruninger (1996) argumentam que se uma ontologia é uma estrutura para comunicação entre pessoas, então a representação da ontologia pode ser informal, contanto que não seja ambígua. Entretanto, se a ontologia é para ser usada por ferramentas de *software* ou agentes inteligentes, então a semântica da ontologia deve ser muito mais precisa. O Quadro 7 apresenta os tipos

de ontologias com base no grau de formalismo de especificação. Os tipos variam desde ontologias informais, expressas em linguagem natural, até ontologias rigorosamente formais, que são especificadas em linguagem formal, como, por exemplo, lógica de primeira ordem, que utiliza axiomas e prova de teoremas. As ontologias semi-informais, geralmente, são ontologias especificadas numa linguagem formal, com menos rigor de especificação que ontologias formais. Além da classificação quanto ao grau de formalidade mostrado no Quadro 7, existem outras classificações – quanto à função (Quadro 8), estrutura (Quadro 9) e conteúdo (Quadro 10).

Quadro 7: Tipos de ontologias quanto ao grau de formalidade.

<b>Classificação</b>	<b>Descrição</b>
Altamente informais	Expressa livremente em linguagem natural.
Semi-informais	Expressa em linguagem natural de forma restrita e estruturada.
Semi-formais	Expressa em uma linguagem artificial definida formalmente.
Rigorosamente formais	Os termos são definidos com semântica formal, teoremas e provas.

Fonte: Uschold e Gruninger (1996).

Quadro 8: Tipos de ontologias quanto à função.

<b>Classificação</b>	<b>Descrição</b>
Ontologias de Domínio	Reutilizáveis no domínio, fornecem vocabulário sobre conceitos, seus relacionamentos, sobre atividades e regras que os governa.
Ontologias de Tarefa	Inclui um vocabulário sistematizado de termos, especificando tarefas que podem ou não estar no mesmo domínio.
Ontologias Gerais	Incluem um vocabulário relacionado a coisas, eventos, espaço, causalidade, comportamento, funções, etc.

Fonte: Almeida e Bax (2003).

Quadro 9: Tipos de ontologias quanto à estrutura.

<b>Classificação</b>	<b>Descrição</b>
Ontologias de Alto Nível	Descrevem conceitos gerais relacionados a todos os elementos da ontologia (espaço, tempo, matéria, objeto, evento, ação, etc.) os quais são independentes do problema ou domínio.
Ontologias de Domínio	Descrevem o vocabulário relacionado a um domínio, como, por exemplo, medicina, automóveis, etc.
Ontologias de Tarefa	Descrevem uma tarefa ou atividade, como, por exemplo, diagnósticos ou compra, mediante a inserção de termos especializados na ontologia.

Fonte: Haav e Lubi (2001).

Quadro 10: Tipos de ontologias quanto ao conteúdo.

<b>Classificação</b>	<b>Descrição</b>
Terminológicas	Especificam termos que serão usados para representar o conhecimento em um domínio (por exemplo, os léxicos) .
Informação	Especificam a estrutura de registros de banco de dados (por exemplo, os esquemas de banco de dados.)
Ontologias de modelagem do conhecimento	Especificam conceitualizações do conhecimento, têm uma estrutura interna semanticamente rica e são refinadas para o uso no domínio do conhecimento que descrevem.
Ontologias de aplicação	Contêm as definições necessárias para modelar o conhecimento em uma aplicação.
Ontologias de domínio	Expressam conceitualizações que são específicas para um determinado domínio do conhecimento.
Ontologias genéricas	Similares às ontologias de domínio, mas os conceitos que as definem são considerados genéricos e comuns a vários campos.
Ontologias de representação	Explicam as conceitualizações que estão por trás dos formalismos de representação do conhecimento.

Fonte: Van Heijst, Schreiber e Wielinga (2002) *apud* Almeida e Bax (2003).

## 5.3 Metodologias de Construção de Ontologia

Uschold e Gruninger (1996) argumentam que não existe uma metodologia-padrão de desenvolvimento de ontologias; não há muito material publicado nesta área. Noy e McGuinness (2003) esclarecem que não há um modo correto para modelar uma ontologia, existem alternativas viáveis. A melhor solução depende quase sempre da aplicação que se tem em mente e das extensões que se antecipa. O desenvolvimento de ontologias é, necessariamente, um processo iterativo; conceitos, na ontologia, devem ser orientados à objetos (físicos ou lógicos) e relacionamentos, no seu domínio de interesse; os objetos estão mais para substantivos enquanto os relacionamentos para verbos. Gruber (1993) propõe alguns critérios para o desenvolvimento de ontologias, com ênfase no compartilhamento e na reusabilidade:

- Clareza – a ontologia deve, de modo eficaz, comunicar o significado pretendido dos termos definidos. As definições devem ser objetivas e independentes do contexto social ou computacional e, sempre que possível, expressar as definições através de axiomas lógicos. Todas as definições devem ser documentadas em linguagem natural;
- Coerência – a ontologia deve ser coerente. Os axiomas definidos devem ser logicamente consistentes. A coerência deve aplicar-se também aos conceitos que são definidos informalmente, como aqueles descritos na documentação e nos exemplos em linguagem natural. Se uma sentença que possa ser inferida dos axiomas contradisser uma definição ou um exemplo dado informal, então a ontologia é incoerente;
- Extensibilidade – uma ontologia deve ser projetada prevendo os usos do vocabulário compartilhado. Deve oferecer um fundamento conceitual para uma gama de tarefas previstas, e a representação deve ser realizada de modo que se possa estender e especializar a ontologia monotonicamente, ou seja, deve ser possível definir novos termos para usos específicos baseados no vocabulário existente, de modo que não seja necessário uma revisão das definições existentes;
- Polarização de codificação mínima – a conceitualização deve ser especificada num nível de conhecimento independente de uma linguagem de codificação. A polarização de codificação acontece quando escolhas de representação são realizadas por conveniência de notação ou implementação. A polarização de codificação deve ser minimizada, pois agentes de compartilhamento de conhecimento são implementados em sistemas e estilos de representação diferentes;

- Compromisso ontológico mínimo – uma ontologia deve ter um compromisso ontológico mínimo suficiente para auxiliar as atividades de compartilhamento de conhecimento pretendido. A ontologia deve definir somente aqueles termos que são essenciais para comunicação da consistência do conhecimento.

Uschold e Gruninger (1996) propõem uma metodologia de desenvolvimento de ontologias que estabelece os seguintes passos: identificar a finalidade e o escopo, construção, validação e documentação.

### 5.3.1 Identificação da finalidade e do escopo da ontologia

É importante, esclarecer, para que a ontologia está sendo construída e quais serão os seus usos. Noy e McGuinness (2003) sugerem como ponto inicial no desenvolvimento de uma ontologia a definição do domínio e escopo. Isto equivale a responder às seguintes questões: que domínio a ontologia cobrirá? para que a ontologia está sendo construída? Para que tipo de questões as informações na ontologia devem fornecer respostas? Quem usará e manterá a ontologia? As respostas para estas questões podem mudar durante o processo de projeto da ontologia, mas em qualquer momento elas ajudarão a limitar o escopo da ontologia. Uschold e Gruninger (1996) argumentam que uma forma para determinar o escopo de uma ontologia é esboçar uma lista de questões às quais a ontologia deve ser capaz de fornecer respostas. Esta lista foi denominada de questões de competência.

### 5.3.2 Construção da ontologia

Segundo Noy e McGuinness (2003), em termos práticos, o desenvolvimento de uma ontologia consiste em: definir classes; arranjar as classes numa hierarquia taxonômica (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. Para Uschold e Gruninger (1996), a construção de uma ontologia envolve os seguintes passos: captura, codificação e integração de ontologias existentes.

#### 5.3.2.1 Definição das classes e hierarquias de classes

Inicialmente é importante obter uma lista de termos (termo-significado), sem se preocupar se eles são classes ou propriedades. Estes termos corresponderão às classes ou

propriedades de classes da ontologia. As classes representam conceitos do domínio. As propriedades das classes correspondem às características dos conceitos. Uma vez definida a classe, deve-se descrever a sua estrutura interna que compreende a criação de suas propriedades. A definição da classe compreende, dentre outras coisas, a criação das propriedades que irão descrevê-la. Por exemplo, a classe Carro representa o conceito carro no mundo real; carro possui características como cor, modelo, nome, etc., que, na ontologia, representam propriedades da classe Carro. Tipicamente, a definição de uma classe requer o seguinte: a definição de um nome para a classe (ex. Aluno); a definição de suas propriedades e de seus tipos de dados como, por exemplo, *string*, numérico, booleano, etc.; e restrição de valores.

#### Classe Aluno

Nome:                   Tipo: String  
Matrícula:            Tipo: Numérico  
Telefone:             Tipo: String

Uma classe representa um conceito do mundo real que representa uma coleção de objetos. Uma instância é a identificação de um único elemento desta coleção. Por exemplo, o conceito Aluno representa todos os Alunos. Uma instância da classe **Aluno** é a identificação de um aluno unicamente. A criação de uma instância requer: a escolha de uma classe, a criação de uma instância individual da classe, e o preenchimento dos valores das propriedades da classe. Exemplo de instância da classe **Aluno**:

#### Aluno

Nome: Alex Sandro  
Matrícula: 0001  
Telefone: 783-838

Segundo Uschold e Gruninger (1996), a definição de hierarquias de classes pode ser realizada por meio das seguintes alternativas:

- *Top-down* - o processo de desenvolvimento inicia com a definição dos conceitos mais gerais para os mais específicos;
- *Bottom-up* - o processo de desenvolvimento inicia com a definição de classes mais específicas, as folhas da hierarquia, com subsequente agrupamento destas classes dentro de conceitos mais gerais;

- *Middle-out* - o processo de desenvolvimento é uma combinação das técnicas *top-down* e *bottom-up*. Elabora uma lista dos conceitos mais importantes do domínio; em seguida generaliza ou especializa os conceitos.

A alternativa depende fortemente da visão da pessoa do domínio. Se o desenvolvedor tem uma visão sistemática *top-down* do domínio, então pode facilmente usar a alternativa *top-down*. A alternativa *middle-out* é mais fácil para muitos desenvolvedores de ontologias, desde que os conceitos do meio “*in the middle*” tendem a ser os conceitos mais descritivos do domínio (Rosch, 1978, *apud* McGuinness, 2001). Para Uschold e Gruninger (1996), a opção por uma das técnicas tem vários efeitos: a alternativa *bottom-up* resulta num alto nível de detalhe que aumenta o esforço, os riscos de inconsistências e retrabalho; a alternativa *top-down* resulta num maior controle do nível de detalhe, entretanto começar pelo topo pode resultar numa escolha arbitrária de categorias do nível superior que pode resultar na instabilidade do modelo, implicando em retrabalho e um maior esforço; a alternativa *middle-out* define os conceitos mais fundamentais antes de movê-los para dentro de conceitos mais abstratos ou mais específicos, facilita o relacionamento dos termos de diferentes áreas e, além disso, reduz o potencial de retrabalho.

A hierarquia de classes é projetada empregando a relação “is-a” (é-um) para a definição de classe subclasse. Uma classe A é subclasse da classe B se toda instância de A é também instância de B. Quando dizemos que a classe A é subclasse da classe B, isto equivale a dizer que A *é-um* B, por exemplo, a classe Jeep é subclasse da classe Carro, ou seja, Jeep *é-um* Carro. As subclasses de uma classe arbitrária X descrevem uma visão mais condensada da realidade do que X descreve. Uma forma de garantir que a hierarquia de classe está correta é por meio da relação “is-a” ou “*kind of*” (do tipo); verifica-se foi aplicada a relação *é-um* entre a subclasse e a classe superior. O relacionamento de subclasse é transitivo, isto significa que se B é subclasse de A e C é subclasse de B, então C é subclasse de A.

### 5.3.2.2 Codificação da ontologia

Por codificação entende-se a representação explícita da conceitualização capturada no estágio anterior em alguma linguagem formal. Isto envolve: validar os termos básicos que serão usados para especificar a ontologia (ex.: classes, entidades, relações), escolher uma linguagem de representação e escrever o código.

### 5.3.2.3 Integração de ontologias existentes

É necessário avaliar a possibilidade de usar ontologias já existentes para poupar esforços e minimizar custos. Muitas ontologias já estão disponíveis na Internet e podem ser importadas dentro de um ambiente de desenvolvimento. O formalismo o qual uma ontologia é expressa não tem sido problema desde que muitos sistemas de representação de conhecimento podem importar e exportar ontologias. Mesmo se um sistema de representação do conhecimento não pode trabalhar diretamente com um formalismo particular, a tarefa de tradução de uma ontologia num formalismo para outro não é difícil.

### 5.3.3 Avaliação de ontologia

A avaliação das definições da ontologia é técnica e deve ser executada durante todo o ciclo de vida da ontologia. O objetivo é detectar a ausência de algumas propriedades bem definidas nas definições. Segundo Gómez-Pérez (1994), avaliação inclui os seguintes passos:

- Verificar a estrutura ou arquitetura da ontologia – o objetivo é verificar se as definições seguiram os critérios do projeto do ambiente na qual elas estão incluídas. Ontologias construídas com o uso do ambiente Ontolíngua devem satisfazer aos cinco critérios de projeto estabelecidos por Gruber (1993);
- Verificar a sintaxe das definições – detectar o quanto antes estruturas sintaticamente incorretas. O ambiente de desenvolvimento deve fornecer analisador sintático que verifica automaticamente a presença ou ausência de documentação em linguagem natural, palavras-chave erradas nas definições formais e ausência de laços infinitos entre as definições;
- Verificar o conteúdo das definições – detectar o que a ontologia define, não define, ou define incorretamente, e o que pode ser inferido, não pode ser inferido ou pode ser inferido incorretamente.

Segundo Gómez-Pérez (1994), verificar o conteúdo das definições lida com o problema dos três Cs: consistência, completude e concisão:

*Consistência* - refere-se à incapacidade de obter conclusões contraditórias a partir de dados de entrada válidos. Uma ontologia é semanticamente consistente somente se suas definições são semanticamente consistentes. Uma definição é semanticamente consistente

somente se: 1) a definição individual é consistente, o que significa que os significados das definições formais e informais são consistentes com o mundo real, e consistentes entre si, e 2) elas não são sentenças contraditórias que podem ser inferidas usando outras definições e axiomas que podem ou não ser da mesma ontologia;

*Completeness* - refere-se à extensão, ao grau ou a cobertura da informação da ontologia independente do usuário. A definição de completude depende do nível de acordo de granularidade da ontologia como um todo. Podemos dizer que uma definição é completa se nada foi esquecido. Completude de definição lida com completude de suas definições formais e informais;

*Conciseness* - refere-se à utilidade e precisão de toda informação obtida da ontologia. Concisão não implica ausência de redundância. Algumas vezes algum grau de redundância pode ser útil nas definições. Podemos garantir que uma dada definição numa ontologia é concisa se: primeiro não existem redundâncias na sua definição formal e informal; segundo redundâncias explícitas não existem entre definições; terceiro, redundância não podem ser derivada através de axiomas associados as definições. Além disso, o conjunto de propriedades de uma classe são precisamente definidas. A explicação textual das definições e exemplos não são considerados conhecimentos redundantes das definições formais.

### 5.3.4 Documentação

É desejável que ontologias sejam documentadas para facilitar o compartilhamento de conhecimento. Segundo (Skuce,1995 *apud* Uschold, 1996), uma das principais barreiras relacionadas ao compartilhamento de conhecimento é a documentação inadequada de ontologias e bases de conhecimentos. É necessário que todos os conceitos da ontologia sejam documentados, ou seja, as classe, as propriedades, etc.

## 5.4 Exemplos de Ontologias

Nesta seção descreveremos sucintamente alguns exemplos de ontologias conhecidas como, por exemplo, WordNet, CYC e ODP.

### 5.4.1 CYC

O projeto CYC, cujo nome deriva de *enCYClopædia*, fornece um fundamento para inferência do senso comum através do desenvolvimento de ontologias para um ampla

variedade de aplicações de domínios específicos. As ontologias são organizadas dentro de conjunto de módulos conhecidos como microteorias. É reconhecido como maior esforço direcionado a desenvolver uma ontologia de alto nível com maior amplitude possível. A ontologia Cyc pode ser vista como exemplo extremo de alto nível de formalização. Possui centenas de milhares de conceitos formalizada através de axiomas lógicos, regras e outras declarações que especificam restrições de instâncias e classes. *CycLanguage* é usada para expressar a ontologia Cyc. A sintaxe da linguagem CycL é derivada do cálculo de predicado de primeira ordem através de uso de conceitos de segunda ordem.

### 5.4.2 WordNet

A *WordNet* foi desenvolvida pelo Laboratório de Ciência Cognitiva da Universidade de Princeton. É um sistema de referência léxico *on-line* onde substantivos, verbos e adjetivos são organizados dentro de conjunto de sinônimos. A *WordNet* divide o léxico dentro de cinco categorias: substantivos, verbos, adjetivos, advérbios e termo função. Ela organiza a informação lexical com base no significado das palavras ao invés da forma. Logo, os relacionamentos semânticos são usados para organização. Os relacionamentos usados na organização da ontologia são sinônimos, antônimos, homônimos, metonímia e relacionamentos morfológicos para reduzir a forma da palavra.

### 5.4.3 Open Directory Project

O *Open Directory Project* – ODP é o maior e mais compreensivo diretório da *Web* editado por humanos que organiza *sites* dentro de hierarquias de categorias. O ODP é um mecanismo de descrição de *sites* baseado em categorias. As categorias do ODP são construídas com a ajuda de dezenas de milhares de editores voluntários espalhados pelo mundo. Um usuário pode publicar um *site* simplesmente iniciando do top da hierarquia do ODP e navegando até encontrar uma categoria que melhor descreva seu *site*. Em seguida o usuário pode descrever a URI através do preenchimento de um título e de uma descrição textual.

## 5.5 Linguagens Formais

Ontologia é uma especificação formal sobre um domínio; logo, requer linguagem lógica formal para expressá-la. A maioria das linguagens para formalização de ontologias derivou

de duas alternativas: lógica de predicado de primeira ordem (FOL) e XML-RDF. Nesta seção descreveremos as idéias por trás destas alternativas e abordaremos as linguagens mais representativas. Em relação à primeira alternativa, abordaremos KIF e CycL. Em relação à segunda alternativa, abordaremos XML, RDF, RDF *Schema* e OWL.

### 5.5.1 Linguagens de predicado de primeira ordem

Nesta seção discutiremos as idéias por trás de KIF e CycL, que são conhecidas como linguagens de ontologia. Ambas estendem linguagem de predicado de primeira ordem por meio de conceito de segunda ordem. Linguagens de representação de conhecimento baseada em FOL – *First Order Logic* emergiram das idéias da comunidade de matemáticos e cientistas da computação, que desejavam definir uma linguagem expressiva para sistemas de computadores. O objetivo era definir uma linguagem que fosse tão expressiva quanto a linguagem natural, mas que não tivesse a imprecisão e ambigüidade desta. KIF e CycL são linguagens baseadas neste tipo de alternativa. KIF inicialmente não foi projetada como uma linguagem de ontologia, mas como uma linguagem de troca de conhecimento. Entretanto, é qualificada como um linguagem de ontologia por causa do seu alto nível de generalidade. CycL, por outro lado foi desenvolvida como uma linguagem de ontologia para representar o conhecimento embutido na ontologia de senso comum CYC.

KIF (*Knowledge Interchange Format*) é uma linguagem orientada para computadores que foi projetada como um formato de troca de conhecimento entre diferentes sistemas de computação para facilitar o compartilhamento de conhecimento. Ela pode ser aplicada para especificação de ontologias, para comunicação de agentes de *softwares*, para automatizar deduções e satisfazer restrições. Tipicamente, KIF trabalha do seguinte modo: um programa lê uma base de conhecimento em KIF e a converte para sua própria linguagem de implementação. O programa realiza algum processamento computacional na base de conhecimento que foi convertida para sua linguagem interna. Depois, quando o programa necessita comunicar-se com outro, mapeia os dados da sua linguagem de volta para KIF. O outro program recebe a informação via-KIF e, em seguida, converte para sua própria linguagem. Desta forma ,KIF trabalha como uma interlíngua entre diferentes sistemas de computação.

A especificação KIF identifica três importantes características para a linguagem: possui semântica declarativa, isto é, o significado das expressões na linguagem de representação pode ser entendido como está; é logicamente compreensível, isto é, qualquer sentença em cálculo de predicado de primeira ordem pode ser expresso; e fornece representação

do conhecimento sobre o conhecimento, conseqüentemente, usuários da linguagem podem introduzir novas construções de representação do conhecimento sem alterar a linguagem.

CycL é uma linguagem de representação de ontologia que foi desenvolvida como parte do projeto Cyc, cujo objetivo é construir uma grande base de conhecimento para prover os computadores de senso comum. CycL é o meio de representação da ontologia Cyc. Como KIF, CycL é também baseada em FOL e também estende FOL com características de segunda ordem.

### 5.5.2 XML, RDF, RDF Schema e OWL

XML (*Extensible Markup Language*) foi desenvolvida para ser uma linguagem legível por máquina que permite a estruturação sintática de documentos. Portanto, XML não pode tratar questões relacionadas à semântica dos documentos. Para isso foram desenvolvidas linguagens que se beneficiaram das vantagens da estrutura sintática do XML, como RDF Schema e OWL (*Web Ontology Language*). Ambas são baseadas em RDF (*Resource Description Framework*), que é um modelo de dados desenvolvido para descrever recursos *Web* com metadados. RDF não é uma linguagem, mas um modelo de dados que é independente de qualquer domínio ou implementação.

O modelo de dados RDF é um grafo e consiste de nós e arcos. Os nós correspondem aos objetos ou recursos e os arcos correspondem às propriedades. Os identificadores dos nós e arcos são URIs (*Uniform Resource Identifiers*). Os recursos são todas as coisas que são descritas por expressões RDF. Um recurso pode ser uma imagem, um documento HTML, parte de uma página *Web*, uma coleção de páginas HTML, etc. Uma propriedade são atributos específicos que descrevem os recursos.

Uma propriedade com o seu valor associado a um recurso faz uma declaração sobre o recurso. *Statement* (declaração) RDF consiste de um recurso específico junto com uma propriedade mais o valor da propriedade para aquele recurso. Logo, uma declaração RDF é composta de três partes: *subject*, *predicate* e *object*. O *subject* (recurso) representa um recurso que será descrito; o *predicate* (predicado) representa uma característica do recurso; e o *object* (objeto) representa o valor do *predicate*. O objeto de uma declaração, isto é, o valor da predicação pode ser outro recurso, uma simples *string* ou algum outro tipo de dado definido pelo XML.

RDF Schema foi desenvolvida com a finalidade de definir o vocabulário usado nos modelos RDF. RDF Schema oferece um conjunto de primitivas de modelagem, como

*rdfs:Class*, *rdf:Property* ou relacionamento *rdf:subClassOf* para definir o vocabulário RDF para aplicações específicas. Em RDF *Schema* é possível definir classes de classes, classes de propriedades, classes de literais que são *string*, numérico, booleanos, etc. Usando propriedades RDF *Schema*, como *rdf:type*, *rdfs:subClassOf* e *rdfs:subPropertyOf*, é possível definir relacionamentos *instanceOf* entre recursos e classes. Usando as propriedades *rdfs:domain* e *rdfs:range* é possível restringir os recursos que podem ser *subject* ou *objects* da propriedade.

RDF *Schema* foi considerada como a primeira iniciativa para definir uma linguagem de definição de ontologias para *Web*; no entanto existem muitas coisas que não podem ser expressas com RDF *Schema* como, por exemplo, união, interseção, complemento e disjunção de classes. Não há como declarar restrição de cardinalidade e as propriedades não podem ser declaradas como transitivas, simétricas ou inversas. Entretanto, os pesquisadores defendem que estas características são essenciais para uma linguagem de ontologia, pois sem elas não há como fornecer suporte para inferência. Devido a estas limitações foi proposto pelo grupo de trabalho da *World Wide Web Consortium* (W3C) a linguagem OWL, que a descreve como “linguagem projetada para uso em aplicações que necessita processar o conteúdo da informação ao invés de simplesmente apresentá-la para humanos”.

Existem três tipos da linguagem OWL:

- OWL *Full* é a mais expressiva. É compatível com RDF, logo todo documento OWL *Full* é um documento RDF válido; entretanto, ela não é dedutível, logo, não é possível executar inferência automática com OWL *Full*;
- OWL DL está relacionada com Lógica de Descrição. É computacionalmente completa e dedutível, conseqüentemente é possível automaticamente verificar inconsistências numa ontologia OWL DL;
- OWL *Lite* é a menos expressiva, destinada para usos em situações onde uma hierarquia de classes simples e restrições simples são necessária.

### 5.5.3 Editor de ontologias Protégé

Existem vários ambientes de desenvolvimento de ontologias, mas optamos por adotar neste trabalho o editor de ontologias *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*. *Protégé* é um ambiente de desenvolvimento de ontologias com funcionalidade para edição de classes, instâncias e propriedades. O *Protégé* suporta a edição de ontologias OWL através do *plugin* OWL, desde quando a linguagem OWL se tornou a linguagem-padrão para especificação de ontologias para *Web*. O *plugin* OWL pode ser usado para carregar e salvar arquivos OWL em diferentes formatos, para editar ontologias OWL através de janelas gráficas customizadas e para executar inferência baseado em Lógica de Descrição.

A interface do usuário consiste de vários painéis, chamados de *tabs*, cada um deles mostra um aspecto diferente da ontologia numa visão especializada. A maioria das *tabs* existentes fornece uma visão estilo *explorer* do modelo, com uma árvore do lado esquerdo e detalhes do recurso selecionado do lado direito.

## 6 A Ontologia de Descrição de Imagens

Este capítulo compreende a definição de uma ontologia de descrição de imagens embasada nas técnicas de representação de imagens por conceito, que representa a informação iconográfica através de conceitos. A ontologia de descrição é uma proposta de representação do conteúdo imagético, somado com um vocabulário de técnicas fotográficas que são empregadas na produção da fotografia e que será representado na ontologia para permitir a descrição da forma como conteúdo da imagem é mostrado. O objetivo da ontologia é empregá-la numa ferramenta (*software*) de indexação e recuperação de fotografias para permitir a descrição e recuperação de imagens. Neste capítulo descreveremos a construção desta ontologia. O processo de desenvolvimento foi baseado na metodologia proposta por Uschold e Gruninger (1996), e nos critérios estabelecido por Gruber (1993). Basicamente, os procedimentos adotados para construção da ontologia foram: identificação da finalidade e escopo, elaboração de uma lista de termos, construção da ontologia, implementação e teste. Além da ontologia, foi desenvolvida uma ferramenta de descrição e recuperação de fotografias que utiliza a ontologia de descrição internamente para permitir ao indexador descrever e recuperar imagens. Esta ferramenta foi desenvolvida com a finalidade de empregar a ontologia para descrever e recuperar fotografias, além de validar a ontologia.

### 6.1 Finalidade e Escopo da Ontologia

A nossa proposta – a ontologia de descrição de imagens – tem como objetivo melhorar a descrição de imagens e, conseqüentemente a precisão de sistemas de recuperação de fotografias. A idéia é enriquecer a descrição de imagens por meio da associação de conceitos e instâncias da ontologia à imagem, desta forma, melhorar a precisão da recuperação de sistemas de informação fotográficas. Outra finalidade está relacionada com o emprego da ontologia de descrição como vocabulário comum entre o indexador e o usuário, por meio de um sistema de indexação e recuperação que combina as técnicas de recuperação com base em facetas e ontologias, de modo que a descrição e a busca sejam realizadas por meio

da seleção de conceitos extraídos da árvore de categorias da ontologia. Adicionalmente, a ontologia pode ser usada para guiar o processo de descrição, através da associação de hierarquia de categorias ao campo de descrição numa interface de usuário do sistema de indexação.

A ontologia de descrição de imagens é uma especificação formal – ou seja, interpretável por computadores – de alguns modelos conceituais de representação da imagem. Dentre os modelos conceituais de representação que vimos anteriormente, optamos por representar na ontologia: o modelo conceitual de Shatford (1986), que propõe algumas categorias para representação do conteúdo da imagem (Quem, Que, Quando, Onde, Sobre); um vocabulário de técnicas fotográficas que será usado para descrever a forma com que o conteúdo da imagem é mostrado de acordo com Manini (2002) e Lacerda (1993); atributos extrínsecos à imagem, ou seja, que estão fora do conteúdo da imagem serão representado com base na classificação de Layne (1994) (atributos biográfico, de relacionamento, de conteúdo e de exemplo); e na classificação de metadados de Gilliland-Swetland (2000) (administrativo, descritivo, de preservação, técnico e de uso). O modelo de Jaimes e Chang (2000) não será representado na ontologia, apesar de suas contribuições para a representação da imagem, pois os primeiros quatro níveis do sistema de classificação estabelecem categorias de descritores que envolvem técnicas de Processamento de Imagens para extraí-los e que estão fora do escopo deste trabalho, já que nosso foco é a descrição de imagens com base em conceitos. Embora os níveis posteriores (5-10) sejam categorias de descrição com base em conceitos, estes estão melhor representados no modelo conceitual de Shatford (1986), pois o sistema de classificação não contempla as categorias Onde e Quando.

A ontologia de descrição de imagens pode ser classificada da seguinte forma: quanto ao grau de formalidade, ela é formal, pois é destinada ao uso em sistemas de informação de imagens; quanto ao uso, ela está relacionada com o compartilhamento de informação entre o indexador e o usuário de imagens e com a reusabilidade dos conceitos e instâncias para descrever e recuperar imagens; quanto à estrutura, ela é de alto nível, pois descrevem conceitos gerais relacionados ao conteúdo da imagem (espaço, tempo, objeto, evento, ação, etc.) de acordo com as categorias de Shatford (1986), que representa o conteúdo da imagem por meio das categorias de alto nível: Quem (objetos), Que (ação, evento), Quando (tempo) e Onde (local). A classificação da ontologia de descrição em alto nível, praticamente define o escopo da ontologia, ou seja, não desenvolveremos um esquema de categorias exaustivo para representar o conteúdo da imagem, pois isto é inviável, já que o conteúdo da imagem pode ser sobre qualquer coisa. No entanto, definiremos apenas as categorias principais embasadas nas definições de Shatford (1986), de modo que possa

ser estendida para representar domínios específicos. A complexidade da representação do conteúdo da imagem está melhor descrita nas palavras de Leung, Hibler e Mwara (1992): “não é a escala de assuntos que determina a dificuldade, mas o fato de que a natureza das aplicações e os usuários potenciais serem desconhecidos”.

## 6.2 Lista de Termos

A literatura existente sobre metodologia de construção de ontologias recomenda como ponto inicial para o desenvolvimento de uma ontologia enumerar os termos principais do domínio sem realizar julgamentos se eles são classes ou propriedades. Uma boa técnica neste estágio é produzir uma lista de termos (termo e significado) mais relevantes. Existem técnicas para obter esta lista de termos. Segundo Uschold e Gruninger (1996), a lista de termos pode ser obtida através de uma sessão de *brainstorming*<sup>1</sup> (ou “tempestade de idéias”). Neste trabalho, obtemos os conceitos mais relevantes a partir da revisão de literatura sobre a representação da imagem, ou seja, a partir do estudo das técnicas de representação de imagens.

A classificação de metadados de Gilliland-Swetland (2000) em administrativos, descritivos, de preservação, técnicos e de uso somada com a classificação de atributos de imagens de Layne (1994) em: biográficos, de conteúdo, de relacionamento e de exemplos foram empregadas para guiar o processo de obtenção dos termos mais relevantes da representação da imagem, pois estas classificações englobam tanto atributos intrínsecos à imagem, ou seja, que fazem parte do conteúdo da imagem, quanto os atributos extrínsecos à imagem, ou seja, que estão fora do conteúdo da imagem, como data, título, etc. Com base nestas classificações identificamos e classificamos os atributos como mostrado na lista abaixo:

- Atributos biográficos: autor, título, data, local, cultura, histórico;
- Atributos de conteúdo ou descritivos: a representação conceitual da imagem de Shatford (1986), somada com a categoria da dimensão expressiva definida por Manini (2002) que estão apresentadas no Quadro 4 (p. 33) são exemplos de atributos descritivos ou de conteúdo;
- Atributos de relacionamento: publicação (jornal, revista, exposição, etc.);

---

<sup>1</sup>Brainstorming (ou “tempestade de idéias”) mais que uma técnica de dinâmica de grupo é uma atividade desenvolvida para explorar a potencialidade criativa do indivíduo, colocando-a a serviço de seus objetivos.

- Atributos administrativos: doador, coleção, indexador, data de indexação, copiraites, publicador e identificador;
- Atributos técnicos: formato, tamanho da imagem;
- Atributo de preservação: qualidade.

A partir da classificação anterior especificamos a lista de termos, mostrado no Quadro 11, que representa tanto o conteúdo visual quanto o conteúdo não visual da imagem embasado nos seguintes autores: Shatford (1986), Manini (2002), Layne (1994) e Gilliland-Swetland (2000). A lista de termos serve como insumo para a próxima etapa do processo de construção da ontologia que consiste em definir os conceitos a partir da definição de propriedades e relacionamentos. A próxima seção consiste em modelar a ontologia de descrição.

Quadro 11: Lista de Termos.

Autor	O nome do fotógrafo.
Título	Título dado para a imagem.
Data	Data da produção da fotografia.
Local	Local da produção da fotografia.
Cultura	Cultura do local da produção da fotografia.
Histórico	Descrição textual relacionada com a “viagem” da imagem.
Palavras-chave	Lista de palavras-chave que representam o conteúdo da imagem.
Descrição	Descrição textual da informação presente na imagem.
Quem	Objetos animados e inanimados. Seres vivos, artefatos, construções, acidentes naturais, etc.
Que	Atividades, eventos e ações das pessoas na imagem.
Quando	Tempo linear ou cíclico, datas e períodos específicos, tempos recorrentes.
Onde	Tipos de lugares geográficos, arquitetônicos ou cosmográficos.
Sobre	Conceitos abstratos.
Uso	Uso da fotografia em exposições, jornais, revistas, etc.
Técnica Fotográfica	Variáveis da técnica fotográfica como, close, plano americano, câmera baixa, luz diurna, etc.
Doador	Doador da imagem: pessoa ou instituição.
Indexador	Nome da pessoa que descreveu a imagem.
Data Indexação	Data da descrição da imagem.
Copirraites	Direitos autorais da imagem.
Coleção	O nome da coleção de que a fotografia faz parte.
Publicador	Entidade responsável por tornar a imagem disponível.
Identificador	Referência não ambígua para a imagem num determinado contexto.
Formato	Nome do formato da imagem, como jpg, gif, eps, etc.
Tamanho	Tamanho do arquivo da imagem em <i>bytes</i> .
Qualidade	Estado da qualidade da imagem, como, por exemplo, bom, regular, excelente.

## 6.3 Desenvolvimento da Ontologia

Nesta seção modelaremos e implementaremos a ontologia de descrição de imagens. A modelagem será realizada com o uso da linguagem UML (*Unified Modeling Language*) que é muito utilizada para a especificação de sistemas orientado a objetos. A linguagem UML possui um conjunto de diagramas e artefatos que são usados para representar conceitos, propriedades, relações e fluxo de processos de um domínio; logo, podemos empregá-la para especificação de uma ontologia. A modelagem de uma ontologia consiste numa definição mais formal da representação do domínio; isto significa que os conceitos são definidos a partir de suas propriedades e relações com outros conceitos de modo que seja possível a implementação da ontologia por meio de uma linguagem formal como *RDF Schema*. Depois da realização da modelagem, contruímos a ontologia com o editor de ontologias *Protégé*. A escolha do *Protégé* foi devido à sua portabilidade e suporte para construção de ontologias usando a linguagem OWL, que é a linguagem pela qual optamos para implementar a ontologia, em função de sua expressividade e por ser a linguagem-padrão de especificação de ontologias para *Web*.

### 6.3.1 Ontologia do conteúdo visual

Iniciaremos a modelagem da ontologia de descrição de imagens com a representação ontológica do conteúdo visual da imagem, através de diagrama de classe, artefato da linguagem UML. No capítulo sobre a representação da imagem, vimos vários modelos de descrição do conteúdo, que geralmente são um conjunto de atributos empregados para extrair o conteúdo informacional da imagem. Dentre os modelos de descrição de imagens que são destinado para representação do conteúdo visual, e que foram vistos anteriormente na seção sobre a representação do conteúdo visual, optamos por representar: o modelo de descrição do conteúdo de imagens de Shatford (1986), por ser uma referência na literatura de descrição de imagens por conceitos, mesmo sendo datado de 1986; a questão da forma com que o conteúdo da imagem é mostrado a partir de variáveis da técnica fotográfica, como proposto por Smit (1996) e Manini (2002); alguns atributos do esquema de descrição de Jörgensen (1996). Entretanto, a classificação de imagens de Jaimes e Chang (2000) não será representada na ontologia de descrição, pois os níveis sintáticos deste sistema se referem à descrição de imagens com base em características de baixo nível, como cor, textura e forma de figuras geométricas, que exigem técnicas de Processamento de Imagens para extraí-los; logo, este tipo de descrição está fora do escopo deste trabalho.

A base da modelagem da ontologia de descrição é o modelo de descrição de imagens de Shatford (1986), que define quatro categorias (QUEM, QUE, QUANDO, ONDE e SOBRE) para extrair o conteúdo informacional da imagem. Estas categorias são conceitos de alto nível que determinaram a técnica de desenvolvimento da hierarquia de classes em *top-down*. Como já dissemos, a ontologia de descrição de imagens é de alto nível, isto significa que definirá apenas conceitos gerais do conteúdo da imagem de modo que possam ser estendidos para domínios específicos.

Nas próximas seções, realizaremos a modelagem e a implementação da ontologia da conteúdo visual que é a representação ontológica das categorias Quem, Que, Onde, Quando, Sobre e Técnica Fotográfica. Tratamos cada categoria do conteúdo visual como uma ontologia independente, de forma que denominamos de Ontologia Quem, Ontologia Onde, etc. Para cada categoria, primeiramente realizamos a modelagem através do diagrama de classe UML; em seguida implementamos a ontologia com o uso do editor de ontologias *Protégé*.

### 6.3.1.1 Construção da ontologia Quem

Segundo Shatford (1986), a categoria QUEM descreve objetos animados e inanimados e seres concretos. A autora esclarece que a categoria QUEM representa respostas das seguintes questões: de QUEM é esta imagem? De que Objetos? De que Seres? Smit (1997) complementa a definição anterior argumentando que a categoria QUEM se refere à identificação do “objeto focado”: seres vivos, artefatos, construções, acidentes naturais, etc. Com base nas definições anteriores, obtívemos os conceitos principais do nível superior da hierarquia Quem: **pessoa, construção, objetos, paisagem, animal, inseto, vegetal, acidentes naturais**. A partir desta lista de conceitos, definimos o primeiro nível da hierarquia da classe QUEM, como mostrado no diagrama de classe, Figura 34 (p. 93). A classe QUEM é a classe raiz da hierarquia; é uma classe abstrata, isto é, não possui propriedades que a caracterizam e não pode ser instanciada, mas pode ser estendida. É descrita pelas suas relações de subclasses. A hierarquia foi construída baseada na relação *is-a* (é-um). Para cada classe (pessoa, construção, animal, vegetal, paisagem, objeto) foram definidas suas propriedades com a especificação do seu tipo, além da definição da relação (*is-a*). A título de exemplo, especificamos no diagrama de classe, Figura 34 (p. 93), uma instância da classe **Atriz** através da relação *rdf:type* para ilustrar a herança de propriedades. A classe **Atriz** é subclasse da classe **Pessoa**; isto significa que **Atriz** é uma (*is-a*) **Pessoa**, ou seja, **Atriz** herda as propriedades da classe **Pessoa**, que são as

características nome, cor, cultura e estágio de vida e que estão ilustradas na instância da classe **Atriz** identificada por **instancia2:Atriz**.

Para a identificação das propriedades dos conceitos, baseamos-nos em dois tipos de características dos objetos: intrínsecas e extrínsecas. As propriedades intrínsecas são aquelas inerentes ao objeto, como, por exemplo, raça e altura de uma pessoa. Por outro lado, as propriedades extrínsecas não fazem parte do objeto, mas estão associadas a eles de alguma forma, como, por exemplo, nome, ocupação de uma pessoa. No diagrama de classe, Figura 34 (p. 93), as propriedades são colocadas dentro de sua respectiva classe, isto é a forma como a linguagem UML representa as propriedades de um conceito, devido à sua relação estreita com a modelagem de sistemas orientados a objetos em que as propriedades têm o escopo limitado pela sua classe. Entretanto, a linguagem RDF *Schema* não possui esta restrição, ou seja, as propriedades são definidas independentes e associadas aos conceitos por meio da relação *rdfs:domain* que define a quais classes a propriedade pertence.

No diagrama de classe, Figura 34 (p. 93), as propriedades são acompanhadas do seu tipo, que é uma restrição do seu valor. Por exemplo, o conceito **Pessoa** é descrito pelas propriedades nome, raça e cultura. A propriedade nome é do tipo *string*, isto significa que o valor que ela pode assumir é um *string* de texto. Em RDF *Schema* o tipo é definido pela relação *rdfs:range* que pode ser tipo de dados primitivos como *string*, numérico, booleano ou tipo classe que pode assumir como valor qualquer instância de sua classe.

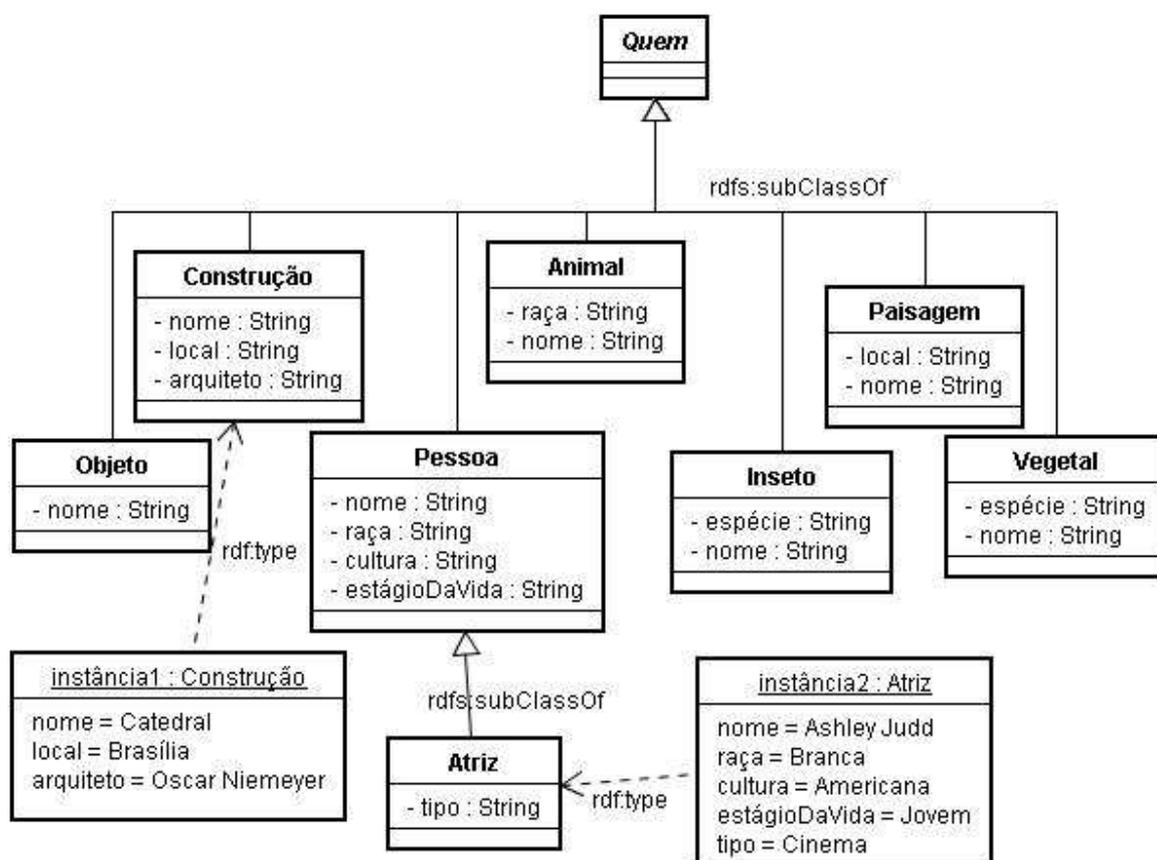


Figura 34: Diagrama de classe Quem.

O diagrama de classe mostrado na Figura 34 (p. 93) é uma representação de alto nível do objeto focado numa fotografia, através das classes **Pessoa**, **Construção**, **Objeto**, **Paisagem**, **Animal**, **Vegetal** e **Inseto**, que foram identificadas com base nas definições de Shatford (1986) e Smit (1997). O objetivo da hierarquia de classe QUEM é ser útil na descrição do objeto focado da fotografia. Embora a modelagem da hierarquia defina apenas o nível superior, ela facilita a extensão da hierarquia, pois as classes superiores são conceitos mais genéricos das suas subclasses; isto significa que ela pode ser estendida observando a relação *is-a* para atender a domínios específicos. No entanto, isto está além do escopo do nosso trabalho, visto que a idéia aqui é desenvolver uma ontologia de descrição de imagens de alto nível embasada nas técnicas de descrição de imagens existentes na literatura de representação de imagens por conceitos.

A implementação da ontologia Quem no editor de ontologia *Protégé* foi realizada por meio do *Plugin* OWL que habilita o editor *Protégé* para o desenvolvimento de ontologias usando a linguagem OWL. A construção da ontologia Quem foi baseada no diagrama de classe Quem, Figura 34 (p. 93), consistiu em: criar as classes, criar as propriedades das classes, definir o tipo das propriedades e a sua cardinalidade e organizar as classes numa hierarquia taxonômica através da relação *rdfs:subClassOf*. Foi verificada a possibilidade de reusar alguma ontologia já existente para estender as categorias, mas não tivemos sucesso nesta busca. A Figura 35 (p. 95), é um recorte da tela de edição de classes do editor *Protégé*, que mostra do lado esquerdo a hierarquia de classe da ontologia Quem, com a classe **Pessoa** selecionada. No lado direito da figura estão as propriedades da classe **Pessoa** acompanhadas da especificação do seu tipo. Desdobramos a classe **Pessoa** através de conceitos que representam a ocupação da pessoa e que estão de acordo com a relação *is-a*.

Como resultado da implementação da ontologia no ambiente *Protégé*, obtivemos uma especificação da ontologia através da linguagem OWL que será integrada posteriormente na ontologia de descrição por meio de mecanismo de importação da linguagem OWL.

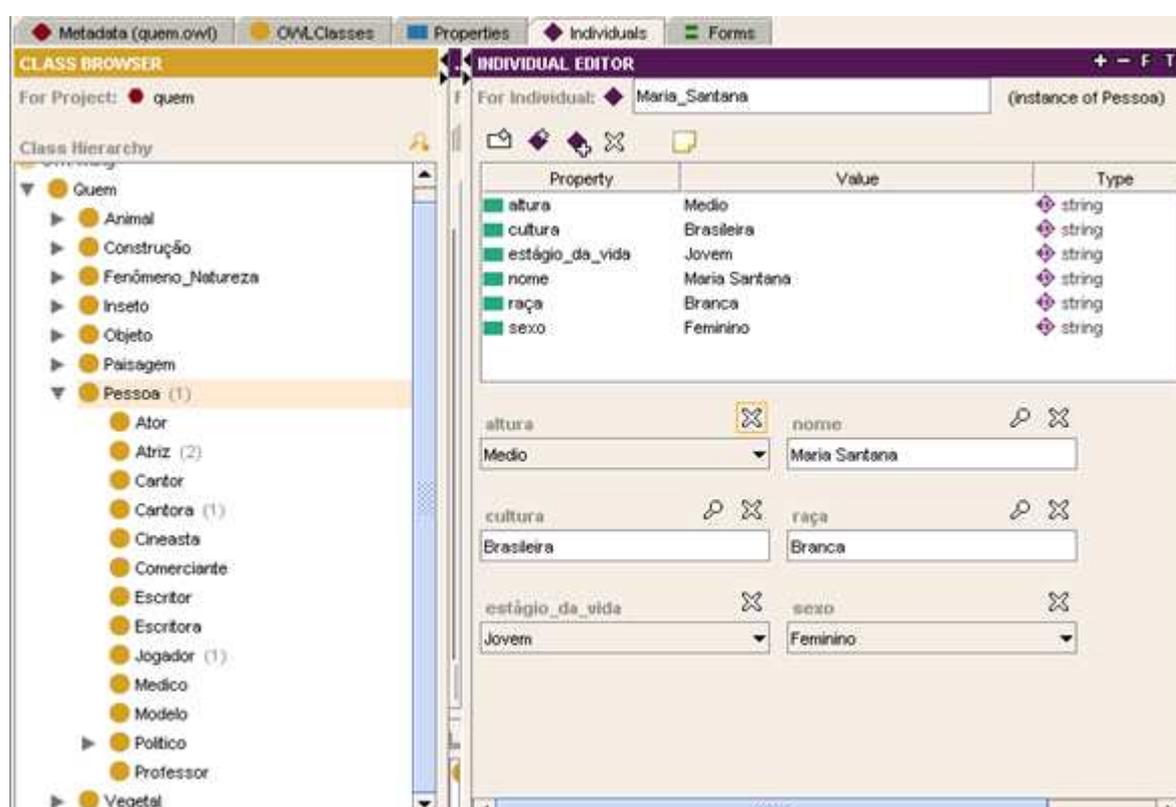


Figura 35: Ontologia Quem.

### 6.3.1.2 Construção da ontologia Onde

A categoria Onde representa, segundo Shatford (1986, p. 53), a localização da imagem no espaço; tipos de lugares geográficos, arquitetônicos e cosmográficos. A autora esclarece que, no nível específico (DE Específico), a faceta Onde consiste de termos nomeando a localização geográfica ou cosmográfica específica como Nova York, Los Angeles, Marte, Plutão. Enquanto no nível genérico (DE Genérico), a faceta Onde refere-se a tipos de lugares geográficos, arquitetônicos ou cosmográficos como por exemplo, caverna, planeta, *shopping*, interiores, floresta, etc.

No diagrama de classe Onde, Figura 36 (p. 97) os lugares específicos estão representados pelas classes abstratas **Local Geográfico** e **Local Cosmográfico**. A hierarquia de classe **Local Geográfico** pode ir até um nível de profundidade, de modo que possa representar cidades individuais. Enquanto as classes **Local Arquitetônico** e **Local Geográfico Genérico** representam tipos de lugares de modo genérico sem especificação de sua localização geográfica como, por exemplo, rio, floresta, cidade, etc.

Durante a implementação do diagrama de classe Onde no *Protégé* foram eliminadas algumas classes como **Local Geográfico** e **Local Cosmográfico** por não representarem conceitos descritivos de imagens; mas mantivemos a sua estrutura (hierarquia), que representa os conceitos descritivos. A Figura 37 (p. 98), mostra a ontologia Onde implementada no *Protégé*. Consiste de uma hierarquia simples sem muitos relacionamentos e facilmente implementável no editor de ontologia *Protégé*.

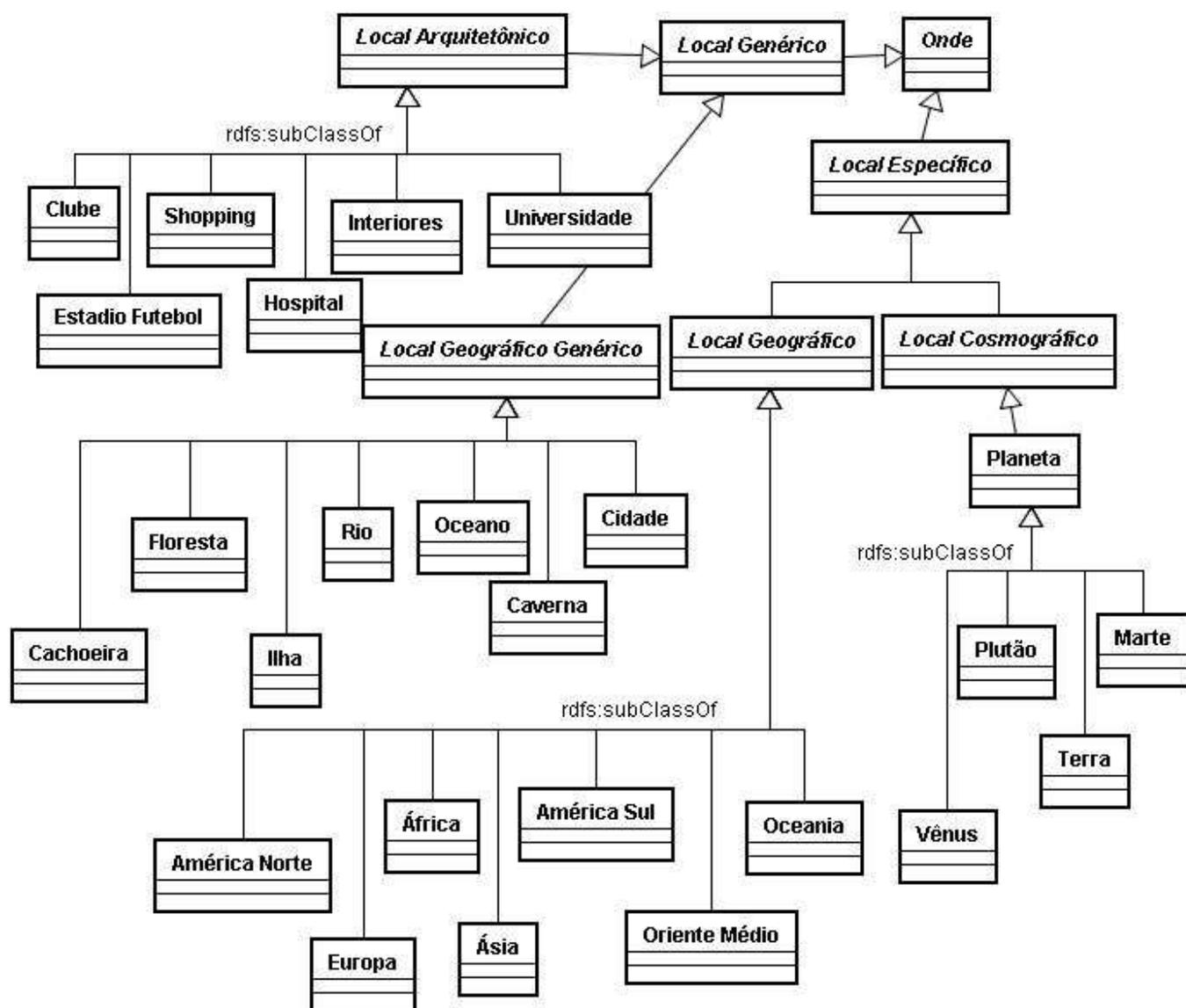


Figura 36: Diagrama de classe Onde.

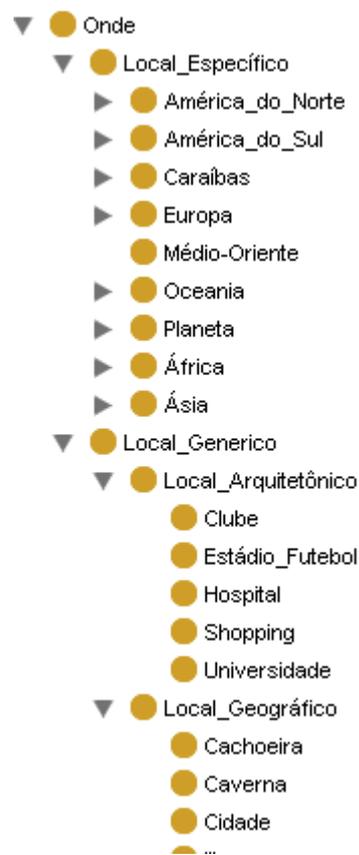


Figura 37: Ontologia Onde.

### 6.3.1.3 Construção da ontologia Quando

Segundo Shatford (1986), a faceta Quando corresponde à localização da imagem no tempo, que pode ser linear ou cíclico. A autora esclarece que o tempo linear corresponde a uma data específica ou período, enquanto tempo cíclico significa tempo recorrente, como estação do ano ou momento do dia, por exemplo, verão, inverno, noite, tarde, etc. Com base nestas definições elaboramos o diagrama de classe apresentado na Figura 38, que representa o tempo cíclico. Quanto ao tempo linear, será representado posteriormente como uma propriedade, pois não configura um conceito.

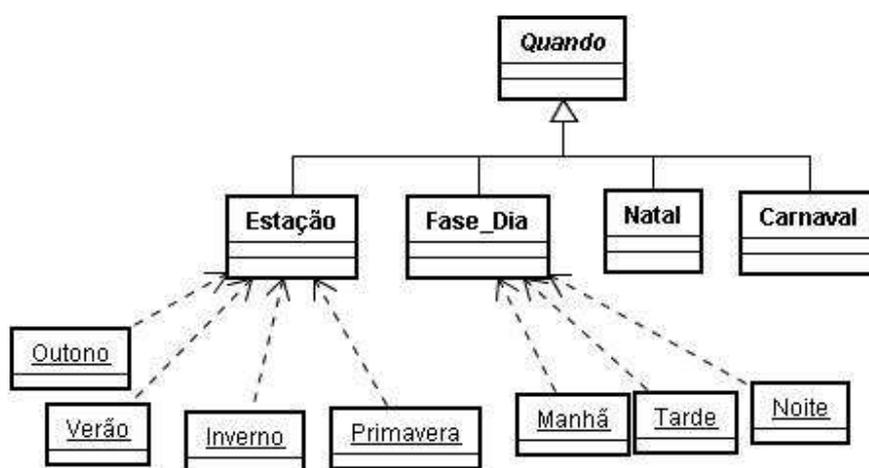


Figura 38: Diagrama de classe Quando.

A implementação da ontologia apresentada na Figura 39 é basicamente uma reprodução do diagrama de classe exibido na Figura 38, com o uso do editor *Protégé*.



Figura 39: Ontologia Quando.

### 6.3.1.4 Construção da ontologia Que

Segundo Shatford (1986, p. 52), a categoria QUE é composta de termos que designam eventos, ações, condições, emoções: qualquer termo que responda às seguintes questões “O que os objetos ou seres estão fazendo?” e “Quais são as suas condições ou estado?”. Smit (1997) esclarece que a categoria QUE se refere à descrição de atitudes ou detalhes relacionados ao “objeto focado” quando este é um ser vivo (por exemplo, cavalo correndo, criança trajando roupa do século XVIII). Com base nestas definições, elaboramos o diagrama de classe abaixo, cujo objetivo é representar os eventos das quais as pessoas estão participando na fotografia, as atividades, o estado emocional e as condições físicas dos objetos presentes na imagem.

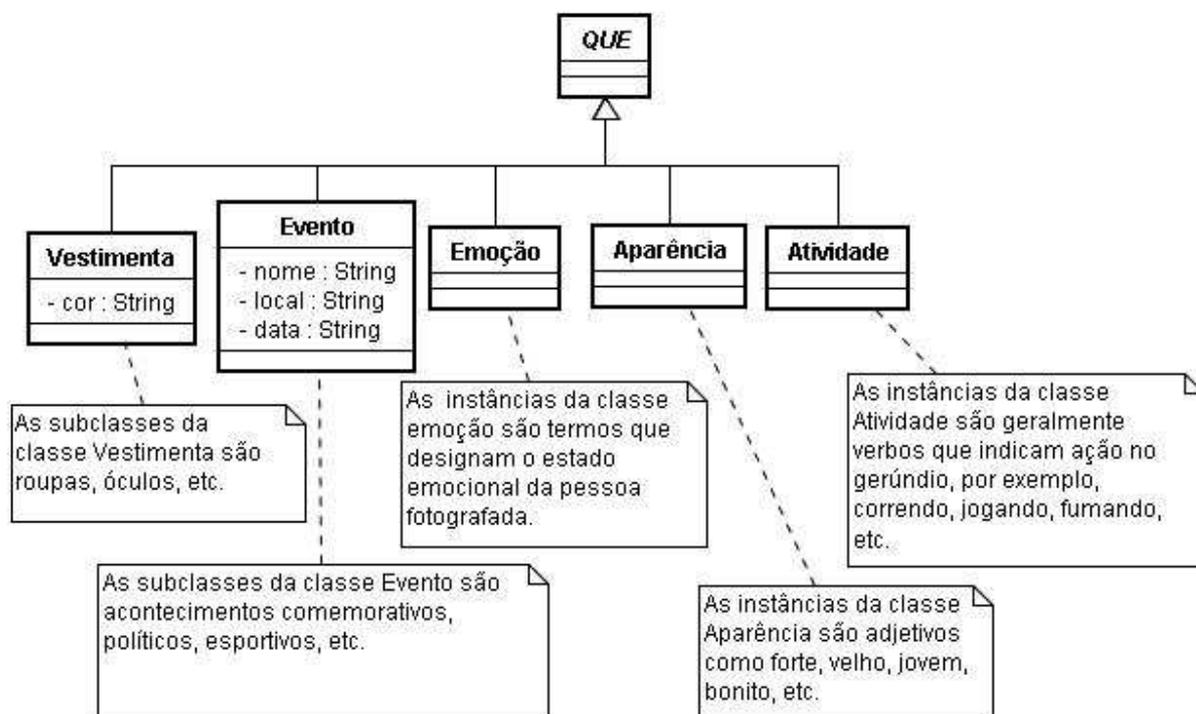


Figura 40: Diagrama de classe Que.

A implementação do diagrama de classe Que resultou na ontologia exibida na Figura 41 (p. 101) e nas Figuras 42, 43 (p. 102). Apenas as classes Evento e Vestimenta possuem propriedades, enquanto as outras são desprovidas de propriedades por serem verbos ou adjetivos.



Figura 41: Ontologia Que.



Figura 42: Ontologia Que.



Figura 43: Ontologia Que.

### 6.3.1.5 Construção da ontologia Sobre

A faceta Sobre é, segundo Shatford (1986), o resultado da análise das facetas Quem, Que, Onde e Quando da seguinte forma: para a faceta Quem verifica-se: os seres ou objetos funcionam como símbolos de outros seres ou objetos? Representam a manifestação de uma abstração? Por exemplo, ponte (DE Genérico) e ponte das bandeiras (DE Específico) podem suscitar Urbanização. Quanto à faceta Onde verifica-se: o lugar simboliza um lugar diferente ou mítico? O lugar representa a manifestação de um pensamento abstrato? Por exemplo: selva (DE Genérico) e Amazonas (DE Específico) podem suscitar paraíso (supõem um contexto que permita esta interpretação). Quanto a faceta QUE verifica que idéias abstratas as ações podem simbolizar? Por exemplo, jogo de futebol (DE Genérico) e copa do mundo (DE Específico) podem simbolizar esporte. Portanto, o vocabulário da faceta Sobre consiste de termos que designa idéias abstratas como esporte, urbanização, política, saúde, educação, segurança, trabalho, etc., ou emoções como, por exemplo, amor, felicidade, paz, sucesso, que representam o significado expressivo da imagem. Deste modo, a faceta está mais associada ao assunto da imagem. Com base nestas definições elaboramos o diagrama de classe a seguir para representar a categoria Sobre.

No diagrama de classe da Figura 44 (p. 103) representamos o nível superior da

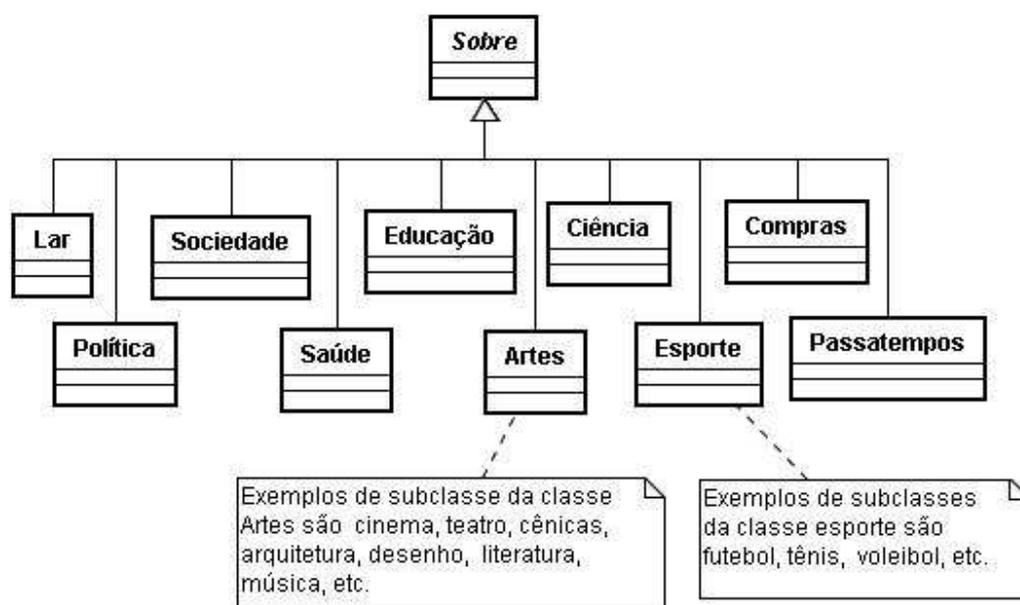


Figura 44: Diagrama de classe Sobre.

hierarquia de classe Sobre. As classes da ontologia Sobre são desprovidas de propriedades por representarem conceitos abstratos ou emoções que não possuem características como as classes da ontologia Quem, que representam objetos.

A categoria Sobre foi desenvolvida usando categorias de metadados de várias fontes; dentre elas, podemos citar o ODP – *Open Directory Project* e o IPTC – *International Press Communications Council*. O ODP é um mecanismo de busca baseado em categorias que são desenvolvidas pela contribuição de muitos editores espalhados pelo mundo. O projeto ODP assim, como é conhecido na comunidade da Internet, disponibiliza suas categorias no modelo RDF para serem reusadas por outras aplicações. Neste trabalho aproveitamos somente algumas categorias do ODP, pois observamos que muitas delas não se aplicam ao domínio de indexação de imagens. Uma das facilidades que encontramos ao usar as categorias do ODP foi devido ao fato de estarem especificadas em RDF, permitindo transformá-las facilmente num esquema de categorias usando a linguagem OWL. Também foram usadas as categorias do IPTC, que é um consórcio de agência, publicadores e vendedores de notícias. Um dos objetivos do IPTC é estabelecer padrões de troca de informação jornalística; ele possui um esquema de categorias de assuntos jornalísticos especificado em XML. Várias subcategorias da categoria Sobre foram criadas com base no IPTC. A Figura 45 (p. 104) mostra o nível superior de classes da ontologia Sobre, enquanto a Figura 46 (p. 104) mostra as subclasses da classe **Artes**.



Figura 45: Ontologia Sobre.



Figura 46: Ontologia Sobre: categoria artes.

### 6.3.1.6 Construção da ontologia Técnica Fotográfica

A ontologia de descrição de imagens acrescenta a representação da forma com que o conteúdo da imagem é mostrado a partir de variáveis da técnica fotográfica como ângulo de tomada, enquadramento, composição, profundidade de campo, *close*, etc., que serão usados como elementos de descrição da imagem. O objetivo do vocabulário da técnica fotográfica é descrever a técnica fotográfica usada na produção da fotografia que, segundo Manini (2002), está associada à expressão fotográfica. A partir deste vocabulário pretendemos complementar a descrição do conteúdo da imagem através de variáveis que estão associadas à noção de forma.

O desenvolvimento do diagrama de classe da técnica fotográfica foi baseado no vocabulário de técnicas fotográfica apresentado no Quadro 6 (p. 46), que classifica as variáveis da técnica fotográfica e que pode ser usado para representar a forma como o conteúdo da imagem é mostrado. O diagrama de classe da técnica fotográfica, Figura 47 (p. 105), representa as classes e suas instâncias através da relação *rdf:type*. Algumas instâncias não foram representadas no diagrama devido à limitação de espaço na página.

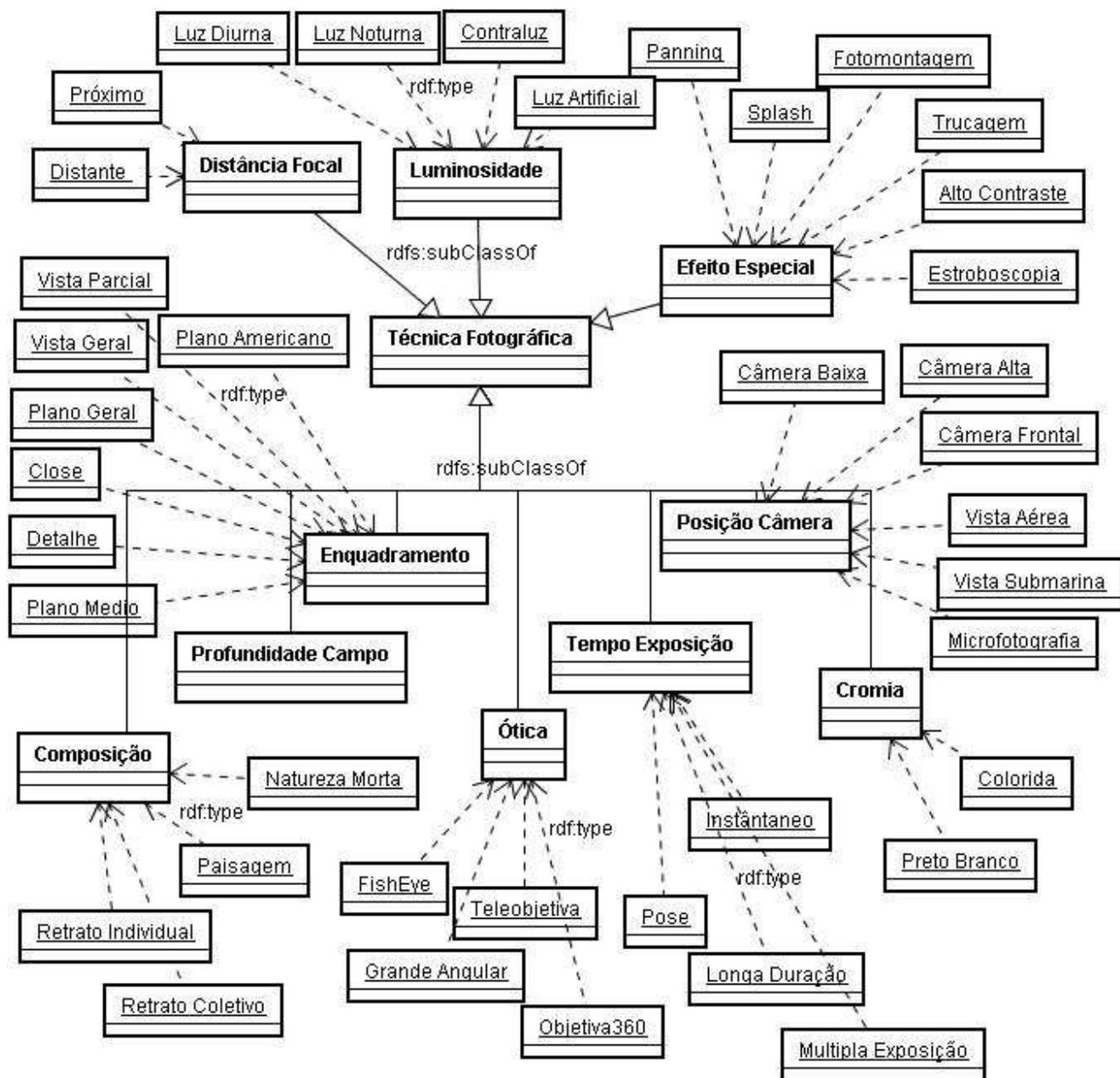


Figura 47: Diagrama de classe Técnica Fotográfica.

A modelagem do diagrama de classe da técnica fotográfica, Figura 47 (p. 105) consiste de uma hierarquia de classes e instâncias com ausência de propriedades de conceitos. Basicamente, a construção da ontologia no *Protégé* compreendeu a criação de classes e instâncias, resultando numa hierarquia de classe conectada pela relação *rdfs:subClassOf* (*is-a*). A Figura 48 (p. 106) mostra a ontologia técnica fotográfica construída no editor *Protégé*. O objetivo da ontologia da técnica fotográfica é ser útil na descrição e na busca de fotografias com base na forma com que o conteúdo da imagem é mostrado.

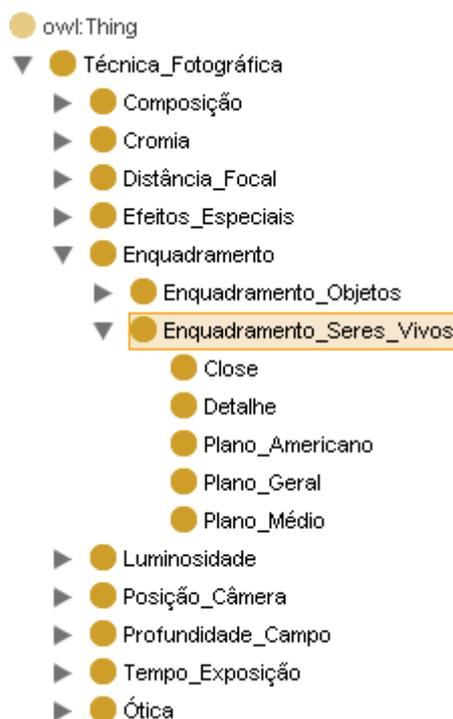


Figura 48: Ontologia Técnica Fotográfica.

### 6.3.2 Ontologia do conteúdo não visual

A lista de termos apresentado no Quadro 11 (p. 89) reúne os atributos da imagem de acordo com a classificação de atributos de Layne (1994) e com a classificação de metadado de Gilliland-Swetland (2000). Estes atributos são representados na ontologia de descrição como propriedades da classe **Fotografia** que é o componente principal do esquema de descrição. O diagrama de classe mídia, Figura 49 (p. 108), tem como classe superior a classe **Mídia**, que representa tipos de mídia como imagem, áudio, vídeo, etc., embora nosso foco seja somente fotografias. A classe **Mídia** possui alguns atributos que são comuns a todos os tipos de mídia, como título, doador e copirraites. A classe **Fotografia** estende a classe **Mídia** e agrega todas as propriedades da lista de termos do Quadro 11

(p.89) que representa tanto o conteúdo visual quanto o conteúdo não visual.

As propriedades da classe **Fotografia** são acompanhadas da especificação do seu tipo, ou seja, da escala de valores que elas podem assumir. As propriedades do conteúdo não visual foram especificadas como sendo do tipo *string*, ou seja, podem armazenar valores como palavras ou texto, enquanto as propriedades do conteúdo visual são tipo instância; isto significa que os valores são instâncias de classe, por exemplo, a propriedade **quem** é definida na classe **Fotografia** da seguinte forma **quem:Quem** tem a seguinte semântica: a palavra **quem** do lado esquerdo aos dois pontos é apenas o nome da propriedade, enquanto a palavra **Quem** do lado direito é o tipo da propriedade; isto significa que os valores da propriedade **quem** são instâncias de qualquer classe da ontologia Quem; por exemplo, um possível valor para a propriedade **quem** é uma instância da classe **Pessoa**.

As propriedades **quem, como (que), quando, onde, sobre e técnica fotográfica** da classe **Fotografia** estabelecem a ligação entre a classe **Fotografia** e as ontologias Quem, Quando, Onde, Que, Sobre e Técnica Fotográfica por meio dos seus tipos. Basicamente, a classe **Fotografia** compreende o esquema de descrição de imagens; ela é responsável por gerar a descrição da fotografia através da criação de instâncias da sua classe que consistem em preencher os valores para as propriedades do tipo *string* e associar conceitos ou instâncias de classe da ontologia para as propriedades do tipo instância: quem, como (que), quando, onde, sobre e técnica fotográfica.

A Figura 50 (p. 109) mostra a ontologia de descrição de imagens que consiste basicamente da classe Fotografia e das ontologias Quem, Onde, Que, Quando, Sobre e Técnica Fotográfica que fornecem o vocabulário para descrever as propriedades da classe **Fotografia**. A Figura 50 (p. 109) mostra a classe **Fotografia** selecionada. No lado direito estão algumas propriedades da classe **Fotografia**. A ontologia de descrição integra as outras ontologias (Quem, Que, Onde, Quando, Sobre e Técnica Fotográfica) por meio do mecanismo de importação da linguagem OWL, como pode ser visto na presença das classes superiores destas ontologias na ontologia de descrição. A ontologia de descrição é responsável pela geração da descrição da fotografia, que é produzida por meio da criação de instâncias da classe **Fotografia**, que consiste em preencher os valores para as propriedades do conteúdo não visual e associar conceitos e instâncias para as propriedades do conteúdo visual. A especificação da ontologia de descrição de imagens através da linguagem OWL, pode ser encontrada no Apêndice (p. 123). A implementação utilizou o metadado *Dublin Core* para representar alguns atributos do conteúdo não visual como título, palavras-chave, identificador, etc. A vantagem de usar os atributos do *Dublin Core*

está relacionada com o fato de que a semântica é bem conhecida.

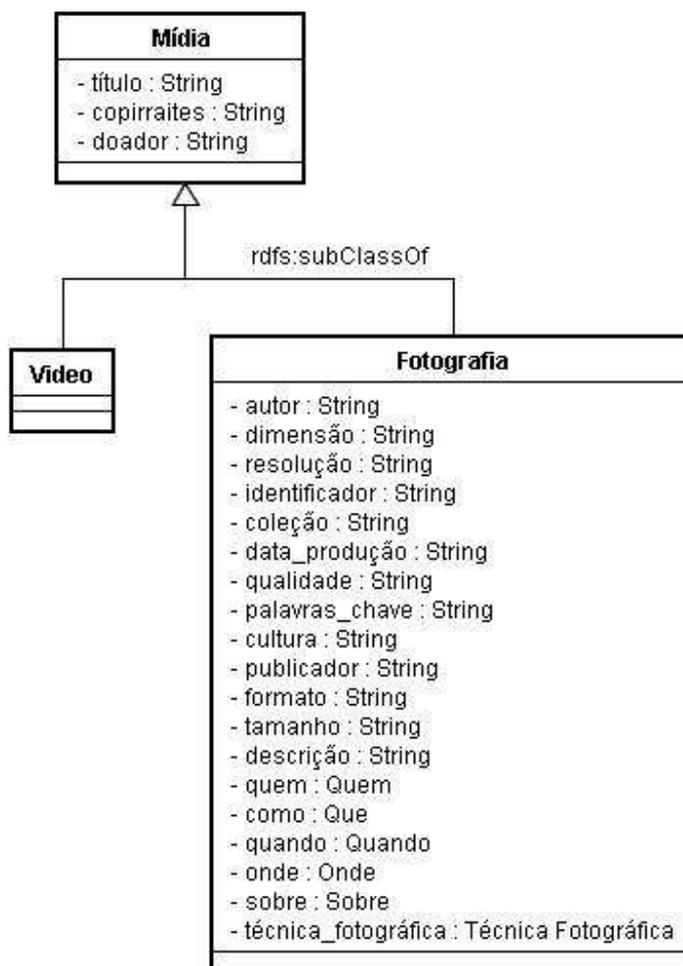


Figura 49: Diagrama de classe Mídia.

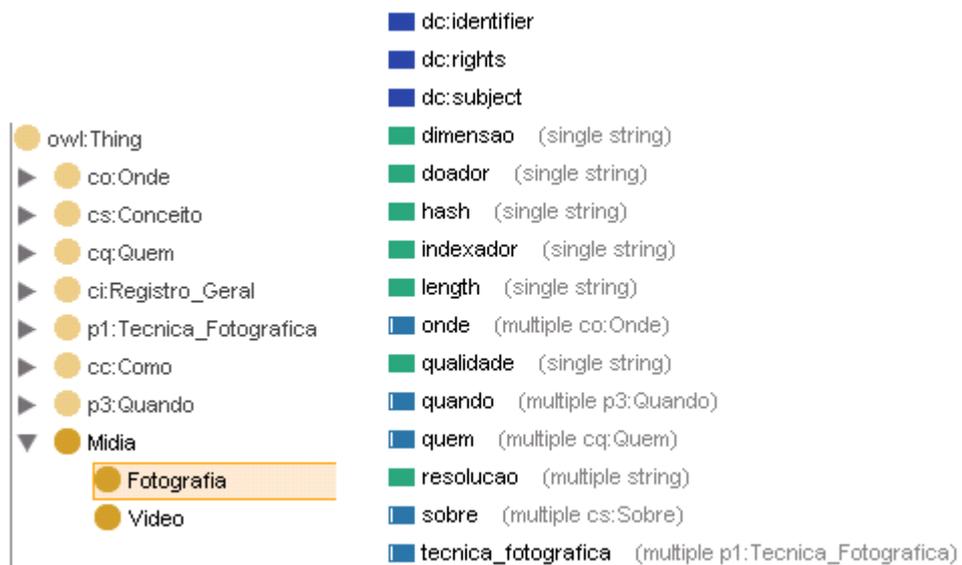


Figura 50: Ontologia de Descrição de Imagens.

A título de exemplo, vamos realizar uma descrição da fotografia, Figura 51, com o uso da ontologia de descrição no editor *Protégé*. Inicialmente, seleciona-se a classe **Fotografia** na ontologia de descrição de imagens e, em seguida, seleciona-se a aba *Individuals* que permite a criação de instância de conceitos. A aba *Individuals*, Figura 52 (p. 111) mostra as propriedades da classe **Fotografia** para entrada de valores. Para o preenchimento das propriedades do tipo *string* é necessário apenas que o usuário entre com o valor da propriedade. Enquanto para as propriedades do tipo instância como quem, onde, como, quando, sobre e técnica fotográfica, o *Protégé*, permite que o usuário selecione uma instância da ontologia para associar à propriedade; se não houver uma instância que represente o conceito desejável, o usuário pode criá-la; neste caso o *Protégé* lê o metadado do conceito e abre uma janela extra, Figura 53 (p. 111) com os campos das propriedades do conceito de modo que o usuário possa entrar com os valores. Depois de gerar a instância do conceito, ela é associada à propriedade da fotografia acompanhada de sua descrição. Após o preenchimentos dos campos da instância da classe **Fotografia**, o *Protégé* grava a descrição no formato RDF, como mostrado na Figura 54 (p. 112).



Figura 51: Ivete Sangalo no carnaval de Salvador.

Fonte: [http://pt.wikipedia.org/wiki/Ivete\\_Sangalo](http://pt.wikipedia.org/wiki/Ivete_Sangalo)



Figura 52: Geração de instância da classe Fotografia no *Protégé*.



Figura 53: Geração de instância da classe Cantora no *Protégé*.

O metadado RDF, mostrado na Figura 54 é resultado da descrição da fotografia Figura 51 (p. 110) no editor *Protégé*, de acordo com os passos descritos anteriormente. O diferencial da nossa proposta de descrição de imagens com base em ontologia torna-se evidente no metadado, mais precisamente no conteúdo das propriedades de instância; por exemplo, a propriedade **quem** recebeu como valor uma instância do classe **Cantora** que descreve uma cantora com base em suas propriedades, que definimos para este conceito e na herança de propriedades de sua classe superior **Pessoa**. Logo, a propriedade quem foi descrita não apenas com nome da pessoa, mas a partir de um conjunto de propriedades e de relações da classe **Cantora**.

```
<Fotografia rdf:ID="Fotografia_1">
  <dc:title rdf:datatype="xsd:string">Ivete Sangalo no carnaval de Salvador</dc:title>
  <dc:description rdf:datatype="xsd:string">
    Ivete Sangalo cantando no carnaval de salvador em cima de um trio elétrico.
  </dc:description>
  <colecacao rdf:datatype="xsd:string">Artistas Bahianos</colecacao>
  <tecnica_fotografica rdf:resource="ct#Teleobjetiva"/>
  <dc:creator rdf:datatype="xsd:string">Macos da Silva</dc:creator>
  <quem>
    <cq:Cantora rdf:ID="Ivete_Sangalo">
      <cq:nome rdf:datatype="xsd:string">Ivete Sangalo</cq:nome>
      <dc:title rdf:datatype="xsd:string">Ivete Sangalo</dc:title>
      <cq:estagio_da_vida rdf:datatype="xsd:string">Jovem</cq:estagio_da_vida>
      <cq:sexo rdf:datatype="xsd:string">Feminino</cq:sexo>
      <cq:cultura rdf:datatype="xsd:string">Brasileira</cq:cultura>
      <cq:cor rdf:datatype="xsd:string">Branca</cq:cor>
      <cq:altura rdf:datatype="xsd:string">1,80</cq:altura>
    </cq:Cantora>
  </quem>
  <como>
    <cc:Carnaval rdf:ID="CarnavalSalvador">
      <cc:local rdf:datatype="xsd:string">Salvador</cc:local>
      <cc:data rdf:datatype="xsd:date">2007-01-15</cc:data>
      <dc:title rdf:datatype="xsd:string">Carnaval de 2007</dc:title>
    </cc:Carnaval>
  </como>
  <quando><j.0:Verão rdf:ID="Verão"/></quando>
  <onde><co:Brasil rdf:ID="Salvador"/></onde>
  <sobre><cs:Música rdf:ID="Música"/></sobre>
  <indexador rdf:datatype="xsd:string">Alex Miranda</indexador>
  <dc:format rdf:datatype="xsd:string">jpg</dc:format>
  <dc:rights rdf:datatype="xsd:string">Wikipédia</dc:rights>
  <qualidade rdf:datatype="xsd:string">Bom</qualidade>
  <doador rdf:datatype="xsd:string">João da Silva</doador>
  <dimensao rdf:datatype="xsd:string">300x200</dimensao>
</Fotografia>
```

Figura 54: Descrição da fotografia em RDF.

Similarmente, à propriedade **quem**, a propriedade **como (que)** foi descrita por meio de uma instância da classe **Carnaval** que estende a classe **Evento**; logo, herda as propriedades título, local e data do evento. Esta forma de descrever imagens por meio da associação de instâncias da ontologia à propriedade da imagem permite uma descrição mais detalhada do conteúdo da imagem e, conseqüentemente, possibilita melhorar a precisão da recuperação, pois todos os campos do metadado são pesquisáveis; além disso, a estrutura do metadado é um grafo conectado semanticamente, o que permite aos agentes de *software* realizarem inferência e extrair informações.

## 6.4 Teste da Ontologia de Descrição de Imagens

Para testar a nossa ontologia de descrição de imagens, optamos por construir um *software* de descrição e busca que utiliza a ontologia de descrição de imagens internamente para auxiliar o indexador na tarefa de descrição, e o usuário na tarefa de busca. A idéia é verificar a usabilidade da ontologia de descrição numa ferramenta construída especificamente para descrever e recuperar fotografias através da ontologia. Este *software* foi denominado de *Splash*, que é o nome de uma técnica fotográfica que consiste em realizar uma fotografia de um evento que acontece em frações de segundos. O *Splash*, Sistema de Descrição e Busca de Imagens, foi desenvolvido usando a linguagem de programação Java, a linguagem de consulta SPARQL, a ontologia de descrição de imagens em OWL e o *framework* Jena, que permite manipular uma ontologia através da linguagem Java (editar classes, instâncias, propriedade e realizar consultas). O *Splash* consiste de dois módulos: descrição de imagens e busca de imagens. Inicialmente descreveremos como foi construído o módulo de descrição, em seguida apresentamos um exemplo de descrição de uma fotografia com o uso do *Splash*. E, finalmente, apresentamos o módulo de busca, que combina as técnicas de busca com base em facetas e em ontologias.

### 6.4.1 *Splash*: módulo de descrição de imagens

O *Splash* carrega ontologia de descrição de imagens internamente usando o *framework* Jena que permite executar operações de edição de conceitos, instâncias, propriedades, além de consultas sobre a ontologia. O módulo de descrição fornece uma interface de usuário simples para auxiliar o indexador descrever a fotografia.

O módulo de descrição de imagens do *Splash* permite que o indexador selecione uma fotografia no sistema de arquivos de um computador e a descreva através de formulários

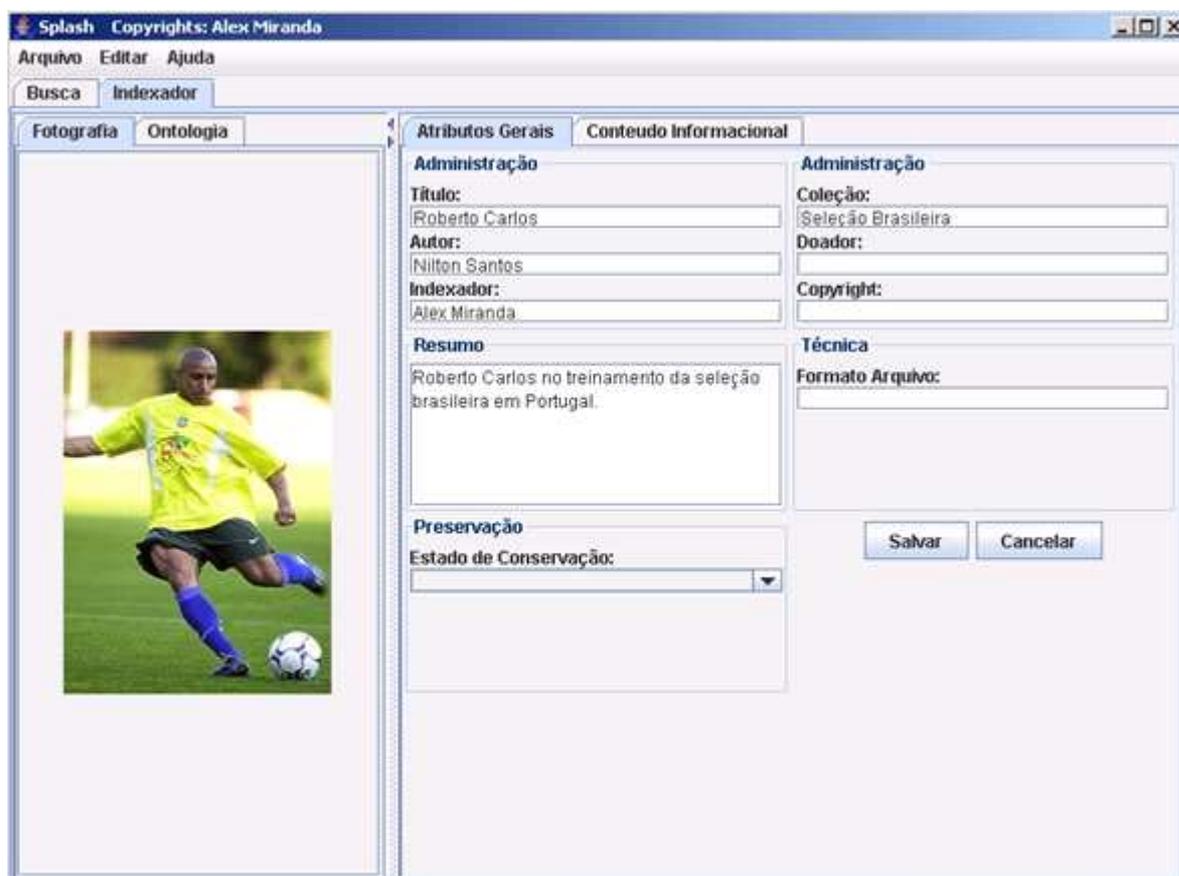
de entrada de dados que permitem o preenchimento dos valores das propriedades da classe **Fotografia** da ontologia de descrição de imagens. A interface de descrição de fotografias é composta por dois formulários de entrada de dados. O primeiro permite que o indexador descreva as propriedades do conteúdo não visual, enquanto o segundo permite que o indexador descreva o conteúdo visual da imagem através da seleção de conceitos ou instâncias da ontologia de descrição de imagens.

A título de exemplo, descrevemos a fotografia abaixo, Figura 55, com o uso do *Splash*, que consiste inicialmente em fornecer os valores para as propriedades do conteúdo não visual (título, autor, doador, resumo, etc.) como mostrado na interface do usuário do sistema na Figura 56 (p. 115). Após o preenchimento das propriedades do conteúdo não visual, o indexador seleciona o formulário de descrição do conteúdo visual, Figura 57 (p. 116), que apresenta as propriedades do conteúdo visual da imagem, mais a propriedade técnica fotográfica. O *Splash* associa as hierarquias do conteúdo visual (Quem, Quando, Onde, Como e Sobre) aos respectivos campos da propriedade, no formulário de entrada de dados, de modo que o usuário possa navegar dentro da hierarquia da ontologia para selecionar conceitos e instâncias. Para descrever o conteúdo visual, o indexador seleciona a instância que expressa o conceito desejável a partir da árvore de categorias da ontologia; assim, o sistema associa a instância à respectiva propriedade do metadado. Após a realização da descrição, o indexador salva a descrição e, com isso, o *Splash* gera a descrição em RDF similar à descrição obtida por meio do ambiente *Protégé* Figura 54 (p. 112).



Figura 55: Roberto Carlos.

Foto: Nilton Santos/Divulgação Ambev Roberto Carlos durante o treino da seleção brasileira em Portugal.



The screenshot shows a software window titled "Splash Copyrights: Alex Miranda" with a menu bar containing "Arquivo", "Editar", and "Ajuda". Below the menu bar are two buttons: "Busca" and "Indexador". The main interface is divided into two tabs: "Fotografia" and "Ontologia". The "Fotografia" tab is active and displays a photograph of a soccer player (Roberto Carlos) in a yellow jersey, kicking a ball on a green field. To the right of the image is a metadata form with two main sections: "Atributos Gerais" and "Conteúdo Informacional".

**Atributos Gerais**

**Administração**

Título: Roberto Carlos  
Autor: Nilton Santos  
Indexador: Alex Miranda

**Resumo**

Roberto Carlos no treinamento da seleção brasileira em Portugal.

**Preservação**

Estado de Conservação: [dropdown menu]

**Conteúdo Informacional**

**Administração**

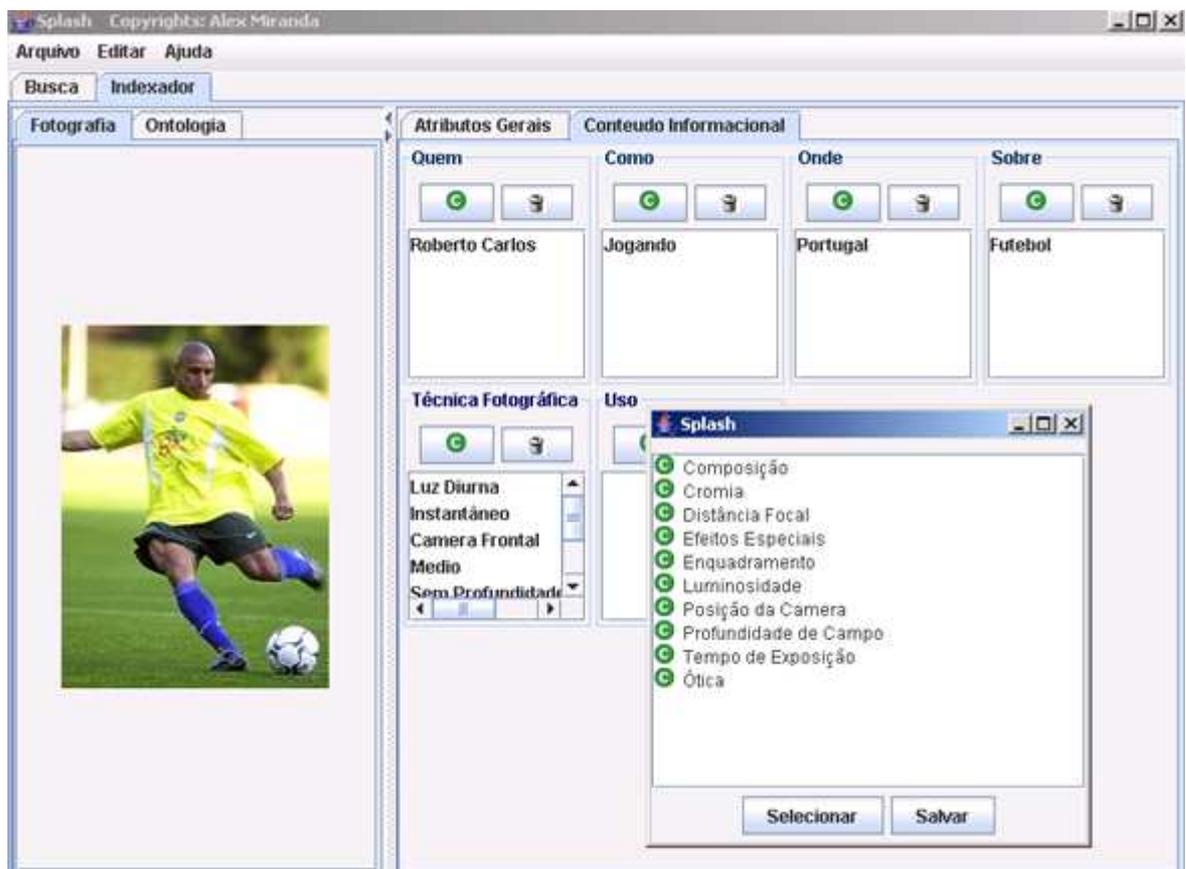
Coleção: Seleção Brasileira  
Doador: [text field]  
Copyright: [text field]

**Técnica**

Formato Arquivo: [text field]

Buttons: Salvar, Cancelar

Figura 56: *Splash*: formulário de descrição do conteúdo não visual.

Figura 57: *Splash*: formulário de descrição do conteúdo visual.

A interface de descrição do *Splash* é muito simples de forma que qualquer pessoa pode usá-la para descrever imagens, diferentemente da geração da descrição com editor de ontologias *Protégé* que possui uma interface muito complexa para usuários finais, em função de sua finalidade que é prover uma ambiente para desenvolvimento de ontologias de modo flexível; logo, não é uma ferramenta destinada para o uso comum por usuário leigos.

### 6.4.2 **Splash: módulo de recuperação de imagens**

O módulo de recuperação explora a ontologia de descrição de imagens para auxiliar o usuário final a formular sua expressão de busca. A construção do módulo de recuperação foi baseada nas idéias de Layne (1994), que argumenta que a indexação de imagens deve fornecer acesso às imagens com base em seus atributos; além disso a indexação de imagens deve fornecer acesso não somente a imagens individuais mas a grupo de imagens; e esclarece ainda que este agrupamento deve ser baseado nos atributos da imagem ou nos atributos do que é representado na imagem. De acordo com Layne (1994), a ontologia de descrição de imagens permite a indexação com base nos seus atributos, intrínsecos e extrínsecos; além disso, a ontologia de descrição permite agrupamento de imagens através da associação de conceitos. Ao associar um conceito da ontologia a uma imagem, estamos agrupando imagens relacionadas a um mesmo conceito, pois um conceito representa um conjunto de instâncias.

Com o objetivo de fornecer acesso não somente a imagens individuais, mas a grupo de imagens, construímos o mecanismo de busca de imagens com base em facetar<sup>2</sup>. Esta técnica foi primeiramente empregada num sistema de recuperação de imagens médicas denominado *HiBrowse* que consiste em apresentar o conteúdo de um repositório de dados através de facetar, de modo que o usuário possa usá-las extensivamente para recuperar informação. Segundo Pollitt (1998), um sistema de busca baseado em facetar é uma ferramenta destinada a satisfazer as necessidades identificadas pelo *Classification Research Group*:

[...] um índice, uma classificação, um mecanismo automático, ou qualquer outro sistema de ‘recuperação de informação’, é uma ferramenta projetada para ajudar o usuário a encontrar seu caminho sobre a massa de informação publicada relacionada a um certo campo do conhecimento. O usuário pode ter um entendimento detalhado do padrão de conhecimento do assunto que ele explora, ou ele pode ter somente um

---

<sup>2</sup>Faceta é uma visão da árvore de categorias através de uma interface gráfica num sistema de recuperação de informação.

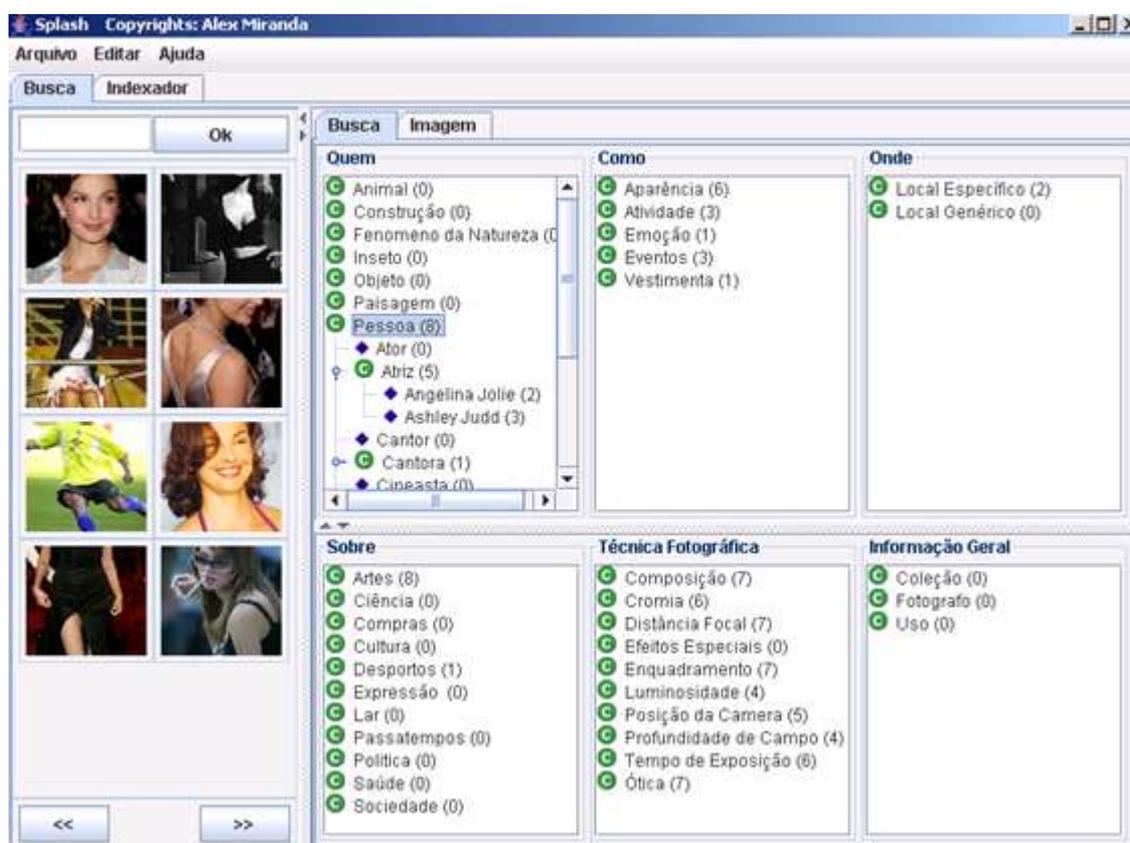
entendimento vago e confuso. Um sistema de recuperação de informação deve ser projetado, primeiramente, para ajudar, mesmo um usuário ignorante, passar de uma vaga formulação do assunto em mente, para uma formulação precisa no sistema; e então, atingindo esta formulação precisa, encaminhá-lo para literatura referente. (*Classification Research Group* 1957 apud POLLIT, 1998)

O módulo de recuperação do *Splash*, Figura 58 (p. 119), apresenta as árvores de categorias do conteúdo visual da imagem (Quem, Que, Quando, Onde, Sobre e Técnica Fotográfica) através de facetadas que permitem ao usuário realizar sua expressão de busca por meio da seleção de conceitos ou instâncias da ontologia ou por meio de uma combinação de seleção de conceitos ou instâncias de várias facetadas. Ao lado do nome de cada conceito ou instância há um valor numérico que corresponde à soma do número de imagens indexadas por aquela hierarquia de conceitos ou pela instância, e serve para guiar o usuário no momento da pesquisa fornecendo a noção de imagens existentes naquele ramo da árvore de categorias.

A pesquisa no *Splash* consiste numa pesquisa multi-facetada, isto é, o usuário faz a seleção de categorias de interesse de diferentes facetadas; com base nesta seleção, o sistema constrói a expressão de busca. Suponhamos que as categorias selecionadas sejam  $C_1, \dots, C_n$  e as subcategorias de  $C_i, i = 1 \dots n$ , sejam  $S_{i,1}, S_{i,2}, \dots, S_{i,k}$ , respectivamente, então, a consulta resulta na seguinte expressão booleana AND ( $\wedge$ ) - OR ( $\vee$ ):

$$(S_{1,1} \vee \dots \vee S_{1,k}) \wedge (S_{2,1} \vee \dots \vee S_{2,l}) \wedge \dots \wedge (S_{n,1} \vee \dots \vee S_{n,m})$$

A expressão de consulta do *Splash* é construída, desta forma através da linguagem de consulta SPARQL, que permite realizar consultas em repositórios de metadados RDF. Desta modo, o *Splash* permite tanto recuperação de imagens individuais quanto recuperação através de grupos de imagens. As imagens individuais são recuperadas através da seleção de instâncias de classe que geralmente correspondem às folhas da árvore de categorias; enquanto a recuperação com base em grupo de imagens é natural na ontologia, visto que os conceitos representam grupos de imagens, ou seja, ao selecionar um conceito o usuário está realizando uma busca com base em grupos de imagens. A pesquisa multi-facetada, ou seja, quando o usuário combina a seleção de conceitos de várias facetadas, reduz drasticamente o número de imagens recuperadas de modo que seja possível atender às necessidades específicas dos usuários.

Figura 58: *Splash*: módulo de busca de imagens.

## 7 Conclusão

Verificamos durante a revisão de literatura a inexistência de padrões de metadados que permitem descrever a informação contida na imagem de modo representativo. Metadados como *Dublin Core* são destinados à descrição geral de recursos digitais; logo, não atendem as especificidades da representação da imagem. Neste trabalho empregamos algumas técnicas de representação de imagens por conceito no desenvolvimento de uma ontologia de descrição de imagens. Esta ontologia consistiu em representar o conteúdo da imagem por meio de uma representação ontológica de alto nível, além de definir internamente um esquema de descrição de imagens que permite que uma ferramenta de *software* possa interpretá-la para descrever e recuperar fotografias.

O grau de complexidade de especificação da ontologia de descrição de imagens está relacionado com a complexidade de representação do conteúdo da imagem, enquanto a complexidade de implementação do projeto - ontologia de descrição mais o sistema protótipo - está relacionado com o número de tecnologias envolvidas para a implementação. O desenvolvimento da ontologia exigiu habilidades com a linguagem OWL e com editor *Protégé*. Já o desenvolvimento do sistema protótipo exigiu habilidades com a linguagem de programação Java e com a linguagem de consulta SPARQL, além de conhecimento em programação de computadores.

A ontologia foi desenvolvida levando em consideração alguns critérios propostos por Gruber (1993). O critério *compromisso ontológico mínimo* foi empregado durante a definição das classes através da representação das propriedades essenciais dos conceitos, ou seja, apenas as propriedades mais relevantes de um determinado conceito foram representadas. A ontologia de descrição de imagens é uma ontologia de alto nível; isto significa que foi projetada por meio da representação de conceitos gerais do conteúdo da imagem, sem levar em consideração um domínio específico; deste modo, fornece uma estrutura geral para ser adaptada ou especializada para atender a um domínio específico. Pode guiar o desenvolvimento da representação ontológica de um domínio específico, já que os conceitos gerais da representação da imagem estão definidos e embasados nas técnicas de

representação de imagem por conceitos; assim, a ontologia está de acordo com o critério de *extensibilidade*. A especificação da ontologia usando a linguagem UML permite que a modelagem da ontologia seja independente da linguagem de codificação, como RDF *Schema*, desta forma atendemos ao critério de *polarização de codificação mínima*. Na definição das classes empregamos termos legíveis de modo que as hierarquias de conceitos sejam facilmente compreensíveis por indivíduos.

Mostramos a geração da descrição de imagens através da ontologia de descrição no ambiente *Protégé* e constatamos que a associação de instância de classes da ontologia para a descrição de uma propriedade da imagem enriquece a descrição da imagem de modo significativo, o que representa o diferencial da nossa proposta de descrição em relação à descrição com base em atributos ou texto. Diferente das alternativas de descrição com base em atributos ou texto, em que a descrição é realizada a partir da atribuição de valores em linguagem natural, a descrição com base em ontologia resulta num metadado descritivo semanticamente estruturado de modo que seja possível para agentes de *software* interpretá-lo e extrair informação.

Demonstramos como a ontologia de descrição de imagens pode ser usada em um sistema de descrição e busca de imagens através do protótipo *Splash*. Combinamos as técnicas de busca com base em facetas e ontologias e verificamos que esta solução pode auxiliar o usuário na formulação de sua expressão de busca e guiá-lo dentro do repositório de imagens através da contextualização da pesquisa fornecida pela árvore de categorias. A pesquisa multi-facetada permitiu a recuperação com base no conteúdo visual da imagem e na técnica fotográfica através da ontologia de descrição. Ela permite que o usuário formule sua expressão de busca por meio da seleção de conceitos de várias categorias, o que aumenta a restrição da pesquisa reduzindo drasticamente o número de registros.

Basicamente, nosso trabalho contribui com uma ontologia de descrição de imagens de alto nível, que, além de ser uma representação da imagem com base em modelos conceituais de descrição de imagens por conceitos, é também um esquema de descrição que pode ser usado por ferramentas de *software* para descrever e recuperar fotografias, como foi demonstrado através do protótipo *Splash*. A ontologia de descrição pode ser usada para descrever imagens de um modo geral. Entretanto, se existe a necessidade de uma representação de um domínio específico, então ela pode guiar o desenvolvimento da ontologia de um domínio específico a partir do desdobramento dos conceitos gerais da ontologia de descrição de imagens (Quem, Quando, Onde, Sobre e Como) para representar os conceitos do domínio de interesse.

## Referências

- ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. v. 32, n. 3, p. 7–20, 2003. Disponível em: <<http://www.scielo.br/pdf/ci/v32n3/19019.pdf>>. Acesso em: 5 out. 2006.
- ALVES, M. D. R.; SOUZA, M. I. F. Dublin core e marc 21: um estudo de correspondência de elementos de metadados. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 4, n. 2, p. 20–38, 2007.
- ARMITAGE, L. H.; ENSER, P. G. B. Analysis of user need in image archives. *Journal of Information Science*, v. 23, n. 4, p. 287–299, 1996.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. *Modern information retrieval*. Addison Wesley, 1999.
- BENITEZ, A. B. et al. Object-based multimedia content description schemes and applications for mpeg-7. 2000. Disponível em: <[http://www.ctr.columbia.edu/papersneadvent/00/ICJ00\\_ana.pdf](http://www.ctr.columbia.edu/papersneadvent/00/ICJ00_ana.pdf)>. Acesso em: 2 fev. 2007.
- BERNERS-LEE; HENDLER, J.; LASSILA, O. The semantic web. *Scientific American*, 2001.
- BONIFACIO, A. S.; HEUSER, C. A. Metadados semânticos para buscas em bibliotecas digitais. 2001. Disponível em: <<http://www.uel.br/pessoal/ailton/Trabalhos/SemanaAcad-Ailton.html>>. Acesso em: 5 jun. 2006.
- BRIN, S.; PAGE, L. The anatomy of a large-scale hypertextual web search engine. Disponível em: <<http://infolab.stanford.edu/~backrub/google.html>>. Acesso em: 10 jan. 2006.
- BURFORD, B.; BRIGGS, P.; EAKINS, J. P. A taxonomy of the image: on the classification of content for image retrieval. 2003. Disponível em: <<http://vcj.sagepub.com/cgi/content/abstract/2/2/123>>. Acesso em: 2 jul. 2007.
- CASTAÑÓN, C. A. B. *Recuperação de imagens por conteúdo através de análise multiresolução por Wavelets*. Tese (doutorado) — Universidade de São Paulo, São Paulo, 2003.
- CHAMPIN, P.-A. *RDF Tutorial*. [S.l.], 2001. Disponível em: <<http://bat710.univ-lyon1.fr/~champin/rdf-tutorial/>>. Acesso em: 2 jun. 2006.
- ESTORNILO, J. *A representação da imagem: indexação por conceito e por conteúdo*. Dissertação (Monografia) — Universidade de São Paulo, São Paulo, 2004.

- FERNEDA, E. *Recuperação de informação: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação*. Tese (doutorado) — Universidade de São Paulo, São Paulo, 2003.
- GILLILAND-SWETLAND, A. J. Introduction to metadata: setting the stage. 2000. Disponível em: <<http://www.slis.kent.edu/~mzeng/metadata/Gilland.pdf>>. Acesso em: 12 nov. 2006.
- GÓMEZ-PÉREZ, A. Some ideas and examples to evaluate ontologies. 1994. Disponível em: <<http://citeseer.ist.psu.edu/gomez-perez94some.html>>. Acesso em: 13 jan. 2007.
- GÓMEZ-PÉREZ, A.; CORCHO, O. Ontology languages for the semantic web. *IEEE INTELLIGENT SYSTEMS*, 2002. Disponível em: <<http://www.cs.umbc.edu/771-/papers/ieeeIntelligentSystems/webservices/ontologyLanguages.pdf>>. Acesso em: 13 jan. 2007.
- GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing. *International Journal Human-Computer Studies*, Italy, n. 43, p. 907–928, 1993.
- HAAV, H.-M.; LUBI, T.-L. A survey of concept-based information retrieval tools on the web. 2001. Disponível em: <<http://www.science.mii.lt/ADBIS/local2/haav.pdf>>. Acesso em: 22 ago. 2007.
- HANDSCHUN, S.; STAAB, S. Annotation of the shallow and the deep web. 1999. Disponível em: <<http://www.aifb.uni-karlsruhe.de/WBS/sha/papers/annotbook-cream.pdf>>. Acesso em: 5 jul. 2004.
- HEFLIN, J.; HENDLER, J. Semantic interoperability on the web. 2000. Disponível em: <<http://www.cs.umd.edu/projects/plus/SHOE/pubs/extreme2000.pdf>>. Acesso em: 12 jun. 2004.
- HOLLINK, L. et al. Classification of user image descriptions. 2003. Disponível em: <<http://www.cs.vu.nl/~guus/papers/Hollink04a.pdf>>. Acesso em: 12 out. 2006.
- HOLLINK, L. et al. Semantic annotation of image collections. 2003. Disponível em: <<http://www.cs.vu.nl/~guus/papers/Hollink03b.pdf>>. Acesso em: 15 out. 2006.
- HYVÖNEN, E.; HOLI, M.; VILJANEN, K. Designing and creating a web site based on rdf content. 2004. Disponível em: <<http://www.seco.tkk.fi/publications/2004/hyvonen-holi-et-al-designing-and-creating-2004.pdf>>. Acesso em: 10 jul. 2004.
- HYVÖNEN, E.; SAARELA, S.; VILJANEN, K. Ontogator: combining view and ontology based search with semantic browsing. 2003. Disponível em: <<http://www.cs.helsinki.fi/u/eahyvone/publications/xmlfinland2003/yomXMLFinland2003.pdf>>. Acesso em: 5 jul. 2004.
- HYVÖNEN, E.; SAARELA, S.; VILJANEN, K. Ontology based image retrieval. 2003. Disponível em: <<http://www2003.org/cdrom/papers/poster/p199/p199-hyvonen.html>>. Acesso em: 5 jul. 2004.
- JAIMES, A. et al. Experiments in indexing multimedia data at multiple levels. 2000. Disponível em: <<http://www.ee.columbia.edu/~ajaimes/Pubs/ideamart00.pdf>>. Acesso em: 15 nov. 2006.

- JAIMES, A.; CHANG, S. Conceptual framework for indexing visual information at multiple levels. 2000. Disponível em: <[http://www.ctr.columbia.edu/~ajaimes/Pubs/spie00\\_internet.pdf](http://www.ctr.columbia.edu/~ajaimes/Pubs/spie00_internet.pdf)>. Acesso em: 13 nov. 2006.
- JÖRGENSEN, C. Indexing images: Testing an image description template. 1996. Disponível em: <<http://www.asis.org/annual-96/ElectronicProceedings/jorgensen.html>>. Acesso em: 12 abr. 2006.
- JÖRGENSEN, C. et al. A conceptual framework and empirical research for classifying visual descriptors. 2001. Disponível em: <[http://www.ee.columbia.edu/~ajaimes/Pubs/jasis01\\_pyramid.pdf](http://www.ee.columbia.edu/~ajaimes/Pubs/jasis01_pyramid.pdf)>. Acesso em: 12 nov. 2006.
- KNUBLAUCH, H.; MUSEN, M. A. Editing description logic ontologies with the protégé owl plugin. 2004. Disponível em: <<http://protege.stanford.edu/plugins/owl/publications/DL2004-protege-owl.pdf>>. Acesso em: 12 abr. 2007.
- LACERDA, A. L. *Os sentidos da imagem: fotografias em arquivos pessoais*. Rio de Janeiro: Acervo, 1993. 41-54 p.
- LAFON, Y.; BOS, B. *Describing and retrieving photos using RDF and HTTP*. [S.l.], 2002. Disponível em: <<http://www.w3.org/TR/photo-rdf/>>. Acesso em: 22 out. 2006.
- LASSI, M. Automatic thesaurus construction. 2001. Disponível em: <<http://www.adm-hb.se/~mol/gslt/thesauri.pdf>>. Acesso em: 10 out. 2006.
- LAYNE, S. S. Some issue in the indexing of images. *Journal of the American Society for Information Science*, 1994.
- LEUNG, C. H. C.; HIBLER, D.; MWARA, N. Picture retrieval by content description. *Journal of Information Science*, n. 18, p. 111–119, 1992.
- MANINI, M. P. *Análise documentária de fotografias: um referencial de leitura de imagens fotográficas para fins documentários*. Tese (doutorado) — Universidade de São Paulo, São Paulo, 2002.
- MCGUINNESS, D. L. Ontologies come of age. Stanford, 2001. Disponível em: <[http://www.hec.unil.ch/myoo/soc/support/lectures/Ontologies\\_come\\_of\\_age.pdf](http://www.hec.unil.ch/myoo/soc/support/lectures/Ontologies_come_of_age.pdf)>. Acesso em: 18 ago. 2007.
- NOLETO, A. D.; BRITO, F. P. Rdf e rdf schema na representação de páginas do portal. 2003. Disponível em: <<http://www.ulbra-to.br/ensino/43020/artigos/anais2003/anais/rdfscheme-encoinfo2003.pdf>>. Acesso em: 11 jun. 2004.
- NOY, N. F.; MCGUINNESS, D. L. Ontology development 101: a guide to creating your first ontology. Stanford, 2003. Disponível em: <[http://protege.stanford.edu/publications/ontology\\_development/ontology101.pdf](http://protege.stanford.edu/publications/ontology_development/ontology101.pdf)>. Acesso em: 10 fev. 2005.
- PANOFSKY, E. *Significado nas artes visuais*. São Paulo: Debates, 1991.
- POLLITT, A. S. *The key role of classification and indexing in view based searching*. UK: [s.n.], 1998. Disponível em: <<http://www.view-based-systems.com/papers/pollifla.htm>>. Acesso em: 22 jun. 2004.

- RAMALHO, J. A.; PALACIN, V. *Escola de fotografia*. São Paulo: Futura, 2004.
- SCHAFER, J. B.; KONSTAN, J.; RIEDLL, J. *E-commerce recommendation applications*. [s.n.]. Disponível em: <<http://www.grouplens.org/papers/pdf/ECRA.pdf>>. Acesso em: 15 out. 2006.
- SCHREIBER, A. T. *Ontology-based photo annotation*. [s.n.], 2001. Disponível em: <<http://www.cs.vu.nl/~guus/papers/Schreiber01a.pdf>>. Acesso em: maio de 2004.
- SCHREIBER, A. T.; BLOK, I. I.; CARLIER, D. *A mini-experiment in semantic annotation*. [s.n.]. Disponível em: <<http://www.cs.vu.nl/~guus/papers/Schreiber02b.pdf>>. Acesso em: 15 out. 2006.
- SHATFORD, S. *Describing a Picture: A Thousand Words are Seldom Cost Effective*. New York: Cataloging and Classification Quartely, 1984. 13-30 p.
- SHATFORD, S. *Analysing the subject of a picture: a theoretical approach*. New York: Cataloging and Classification Quartely, 1986. 39-62 p.
- SMIT, J. W. *Análise documentária: a análise da síntese*. Brasília: IBICT, 1989. 101-113 p.
- SMIT, J. W. *A representação da imagem*. Rio de Janeiro: Informare, 1996. 28-36 p.
- SMIT, J. W. *Propostas para a indexação de informação iconográfica*. Brasília: Mimeo, 1997.
- SOUZA, J. P. *Fotojornalismo: uma introdução à história, às técnicas e à linguagem da fotografia na imprensa*. Porto: [s.n.], 2002. Disponível em: <<http://www.bocc.ubi.pt/pag/sousa-jorge-pedro-fotojornalismo.pdf>>. Acesso em: 12 fev. 2007.
- SOUZA, J. R. *Reconhecimento de padrões usando indexação recursiva*. Tese (doutorado) — Universidade Federal de Santa Catarina, Florianópolis, 1999. Disponível em: <<http://teses.eps.ufsc.br/defesa/pdf/1753.pdf>>. Acesso em: 13 out. 2006.
- SOUZA, R. R.; ALVARENGA, L. A web semântica e suas contribuições para a ciência da informação. 2004. Disponível em: <<http://www.ibict.br/cionline/viewarticle.php?id=71>>. Acesso em: 15 out. 2006.
- STYRMAN, A. *Ontology based image annotation and retrieval*. Dissertação (mestrado) — University Helsinki, Helsinki, 2005. Disponível em: <<http://www.cs.helsinki.fi/u/astyrman/gradu.pdf>>. Acesso em: 13 out. 2006.
- TRAINA, A. J.; TRAINA, C. J. Técnicas e aplicações de recuperação de imagens por conteúdo. Disponível em: <[http://iris.sel.eesc.usp.br/wvc2005/Minicurso\\_CBIR\\_slides.pdf](http://iris.sel.eesc.usp.br/wvc2005/Minicurso_CBIR_slides.pdf)>. Acesso em: 10 out. 2006.
- USCHOLD, M.; GRUNINGER, M. *Ontologies: principles, methods and applications*. 1996. Disponível em: <<http://www.upv.es/sma/teoria/sma/onto%20-%2096-ker-intro-ontologies.pdf>>. Acesso em: 13 fev. 2007.
- USCHOLD, M.; KING, M. *Towards a methodology for building ontologies*. 1995. Disponível em: <<http://citeseer.ist.psu.edu/uschold95toward.html>>. Acesso em: 13 fev. 2007.

---

WIELINGA, B. J. et al. From thesaurus to ontology. 2001. Disponível em: <<http://www.cs.vu.nl/~guus/papers/Wielinga01a.pdf>>. Acesso em: 11 out. 2006.

## APÊNDICE A – Esquema de indexação de imagens em RDF Schema

```

<?xml version="1.0"?>
<rdf:RDF
  xmlns:cc="http://www.assm.com/como.owl#"
  xmlns:cs="http://www.assm.com/foto.owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:p2="http://www.owl-ontologies.com/assert.owl#"
  xmlns:cq="http://www.assm.com/quem.owl#"
  xmlns:j.0="http://www.assm.com/quando.owl#"
  xmlns:ct="http://www.assm.com/tecnica_fotografica.owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:co="http://www.assm.com/onde.owl#"
  xmlns="http://www.assm.com/index.owl#"
  xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ci="http://www.assm.com/info.owl#"
  xml:base="http://www.assm.com/index.owl">
  <owl:Ontology rdf:about="">
    <owl:imports rdf:resource="file:/C:/DEV-TOOLS/workspace/splash/xml/onde.owl"/>
    <owl:imports rdf:resource="file:/C:/DEV-TOOLS/workspace/splash/xml/foto.owl"/>
    <owl:imports rdf:resource="file:/C:/DEV-TOOLS/workspace/splash/xml/quem.owl"/>
    <owl:imports rdf:resource="file:/C:/DEV-TOOLS/workspace/splash/xml/info.owl"/>
    <owl:imports rdf:resource="file:/C:/DEV-TOOLS/workspace/splash/xml/tecnica_foto

```

```
<owl:imports rdf:resource="file:/C:/DEV-TOOLS/workspace/splash/xml/como.owl"/>
<owl:imports rdf:resource="file:/C:/DEV-TOOLS/workspace/splash/xml/quando.owl"/>
</owl:Ontology>
<owl:Class rdf:ID="Fotografia">
  <dc:date rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  ></dc:date>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Midia"/>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Video">
  <rdfs:subClassOf rdf:resource="#Midia"/>
</owl:Class>
<owl:ObjectProperty rdf:ID="sobre">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdfs:range rdf:resource="http://www.assm.com/foto.owl#Sobre"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="tecnica_fotografica">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdfs:range rdf:resource="http://www.assm.com/tecnica.owl#Tecnica_Fotografica"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="quando">
  <rdfs:range rdf:resource="http://www.assm.com/quando.owl#Quando"/>
  <rdfs:domain rdf:resource="#Fotografia"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="como">
  <rdfs:range rdf:resource="http://www.assm.com/como.owl#Como"/>
  <rdfs:domain rdf:resource="#Fotografia"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="quem">
  <rdfs:range rdf:resource="http://www.assm.com/quem.owl#Quem"/>
  <rdfs:domain rdf:resource="#Fotografia"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="uso">
  <rdfs:domain rdf:resource="#Fotografia"/>
```

```
<rdfs:range rdf:resource="http://www.assm.com/info.owl#Uso"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="onde">
  <rdfs:domain rdf:resource="#Fotografia"/>
</owl:ObjectProperty>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/date"/>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/rights">
  <rdfs:domain rdf:resource="#Fotografia"/>
</rdf:Property>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/creator">
  <rdfs:domain rdf:resource="#Fotografia"/>
</rdf:Property>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/format">
  <rdfs:domain rdf:resource="#Fotografia"/>
</rdf:Property>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/subject">
  <rdfs:domain rdf:resource="#Fotografia"/>
</rdf:Property>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/identifier">
  <rdfs:domain rdf:resource="#Fotografia"/>
</rdf:Property>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/type"/>
<rdf:Property rdf:about="http://purl.org/dc/elements/1.1/contributor"/>
<owl:DatatypeProperty rdf:ID="resolucao">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="data_indexacao">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#date"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="colecacao">
  <rdfs:domain rdf:resource="#Fotografia"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="endereco_imagem">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
```

```
</owl:DatatypeProperty>
<owl:FunctionalProperty rdf:ID="indexador">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="length">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="dimensao">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="hash">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="qualidade">
  <rdfs:domain rdf:resource="#Fotografia"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="doador">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:domain rdf:resource="#Fotografia"/>
</owl:FunctionalProperty>
</rdf:RDF>

<!-- Created with Protege (with OWL Plugin 3.3.1, Build 430)
http://protege.stanford.edu -->
```