



UNIVERSIDADE DE BRASÍLIA
Faculdade de Ciência da Informação
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Emílio Evaristo de Sousa

USO DE ONTOLOGIA PARA RECUPERAÇÃO DA INFORMAÇÃO
DISPONIBILIZADA EM VÍDEOS POR MEIO DE INDEXAÇÃO MULTIMODAL

Brasília
2011



UNIVERSIDADE DE BRASÍLIA
Faculdade de Ciência da Informação
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Emílio Evaristo de Sousa

USO DE ONTOLOGIA PARA RECUPERAÇÃO DA INFORMAÇÃO
DISPONIBILIZADA EM VÍDEOS POR MEIO DE INDEXAÇÃO MULTIMODAL

Dissertação apresentada à banca examinadora como requisito parcial para a obtenção do título de mestre em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação da Universidade de Brasília.

Orientador: Prof. Dr. Cláudio Gottschalg Duque

Brasília
2011

UNIVERSIDADE DE BRASÍLIA
Faculdade de Ciência da Informação

Autor(a): Emílio Evaristo de Sousa

Título: Uso de ontologia para recuperação da informação disponibilizada em vídeos por meio de indexação multimodal

Área de concentração: Gestão da Informação

Linha de pesquisa: Organização da Informação

Grupo de pesquisa: Arquitetura da Informação

Dissertação submetida à Comissão Examinadora, designada pelo colegiado do Programa de Pós-Graduação em Ciência da Informação da Universidade de Brasília como requisito parcial para obtenção do título de **Mestre em Ciência da Informação**.

Aprovada por:

Professor Dr. Cláudio Gottschalg Duque
Presidente – (UnB/ PPGCInf)

Professora Dra. Ivette Kafure Muñoz
Membro Interno - (UnB/PPGCInf)

Professora Dra. Fernanda Lima
Membro Externo – (UnB/CIC)

Professor
Suplente – (UnB/PPGCInf)

Dissertação aprovada em:

Dedicatória

Dedico este trabalho primeiramente a Deus, pois sem ele não teria conseguido. À minha irmã; pelo seu apoio e ajuda constante em minha vida. Em especial, à minha mãe; pelo carinho, cuidado, proteção, compreensão, doação, perdão e amor. Você resume o que há de bom nessa vida.

Agradecimentos

À minha mãe, pelo exemplo, motivação, apoio e amor dedicado a mim.

À minha família, com quem aprendi meus valores mais importantes.

Ao Prof. Dr. Cláudio Gottschalg Duque, pela orientação, apoio, incentivo e amizade.

À Prof. Dra. Fernanda Lima e à Prof. Dra. Ivette Kafure Muñoz, pelas excelentes contribuições e comentários sobre a pesquisa.

Aos professores e colegas da FCI, pela amizade e aprendizado durante toda essa jornada.

À Lívia, pelo companheirismo, apoio, ajuda e carinho compartilhado durante esses anos. Você foi essencial nesse trabalho.

Ao “*Index Team*”, pela ajuda nos sábados, domingos e feriados na fase de indexação. Tiago, Lucas, Kalango, Rogério, Fafau, Luiz e todos os membros desse time, muito obrigado!

Ao Éder Gualberto, por toda amizade e ajuda com as ontologias, principalmente nos horários “livres” depois da meia-noite.

À Marta e à Jucilene, da secretaria da Pós-Graduação, pela ajuda, competência e gentileza nesses anos.

Aos meus amigos em Brasília (e arredores), pela simples e tão necessária amizade e pelo companheirismo.

A todos os demais que, direta ou indiretamente, contribuíram para a realização deste trabalho.

“Não consegui chegar a nada, nem mesmo tornar-me mau: nem mau, nem bom, nem canalha, nem honrado, nem herói, nem inseto. Agora, vou vivendo os meus dias em meu canto, serrazinando a mim próprio com o consolo raivoso — que para nada serve — de que um homem inteligente não pode mesmo, a sério, tornar-se algo, e de que somente os imbecis o conseguem.”

Fiódor Dostoiévski

Memórias do subsolo

RESUMO

Recuperar informação disponibilizada em vídeos longos pode ser uma experiência demorada. A recuperação da informação em vídeos apresenta alguns problemas, pois é realizada, principalmente, através das descrições textuais feitas pelos usuários. O objetivo desta pesquisa é promover, com maior índice de precisão (Cleverdon, 1962), a recuperação da informação disponibilizada em vídeos por meio da indexação multimodal e uso de ontologia. Foi desenvolvida uma ontologia de domínio, associada com uma indexação multimodal para avaliação da precisão em um sistema de recuperação da informação audiovisual. Observou-se nos experimentos um aumento significativo da precisão nas configurações que utilizaram a ontologia para recuperar informação audiovisual. A utilização de uma ontologia de domínio da Doutrina Espírita alinhada a indexação multimodal, em um sistema de recuperação de trechos de vídeos de palestras sobre esse tópico, permitiu melhorar a eficiência do mesmo em termos de precisão.

Palavras-chave: Recuperação da informação. Indexação multimodal. Ontologia.

ABSTRACT

Retrieving information available on long videos can be a time consuming experience. Information retrieval on video presents some problems because users mainly do it through textual descriptions. The objective of this research is to promote, with more precision (Cleverdon,1962), information retrieval available on video through the use of multimodal indexing and use of ontology. A domain ontology was developed, associated with an multimodal index for evaluation of accuracy in an audiovisual information retrieval system. A significant increase in accuracy in the experiments was observed in settings that used the ontology to retrieve audiovisual information. The use of a domain ontology of the Spiritist Doctrine aligned with multimodal indexing in a retrieval system for sections of videos of lectures on this topic, allowed for increased efficiency in terms of the same precision.

Keywords: Audiovisual Information Retrieval. Multimodal Indexing. Ontology.

LISTA DE FIGURAS

Figura 1 -	Forma de Indexação utilizada pela Virage	9
Figura 2 –	Sistema de busca Yovisto	10
Figura 3 –	ImageScape	14
Figura 4 –	Tela do analisador sintático PALAVRAS	37
Figura 5 –	Tela do protótipo	40
Figura 6 –	Trecho indexado relacionado ao tempo do segmento	41
Figura 7 –	Palavras-chave após indexação multimodal e processamento pelo <i>parser</i> PALAVRAS	43
Figura 8 –	Parte da hierarquia da classe <i>Medium</i> na ontologia do domínio	45
Figura 9 –	Representação da área administrativa para configuração das consultas	46
Figura 10 –	Representação generalizada do método utilizado.	47

LISTA DE GRÁFICOS

Gráfico 1 -	Resultado da Precisão utilizando a configuração Completa	52
Gráfico 2 –	Resultado da Precisão utilizando a configuração Palavras	53
Gráfico 3 –	Resultado da Precisão utilizando a configuração Completa_Onto	54
Gráfico 4 –	Resultado da Precisão utilizando a configuração Palavras_Onto	54
Gráfico 5 –	Resultado de Precisão do experimento comparando as configurações Completa e Completa_Onto	55
Gráfico 6 –	Resultado de Precisão do experimento comparando as configurações Palavras e Palavras_Onto	56

LISTA DE TABELAS

Tabela 1 -	Configurações de execução do experimento	47
Tabela 2 –	Resultado da Precisão nas diferentes configurações realizadas no experimento	50

SUMÁRIO

1 INTRODUÇÃO	3
1.1 DEFINIÇÃO DO PROBLEMA.....	4
1.2 OBJETIVOS DA PESQUISA	4
1.2.1 Objetivo geral	4
1.2.2 Objetivos específicos	5
1.3 JUSTIFICATIVA	5
2 REVISÃO DA LITERATURA	7
2.1 TRABALHOS CORRELATOS	7
2.2 RECUPERAÇÃO DA INFORMAÇÃO	11
2.2.1 Recuperação do conteúdo audiovisual	11
2.2.2 Recuperação da informação em palestras	15
2.3 INDEXAÇÃO	16
2.3.1 Indexação da informação audiovisual	17
2.3.2 Indexação multimodal	19
2.3.2.1 <i>Indexação do conteúdo imagético em vídeos</i>	20
2.3.2.2 <i>Transcrição automática do discurso</i>	22
2.4 ONTOLOGIAS	24
2.4.1 Tipos de ontologia	27
2.4.2 Métodos de desenvolvimento de ontologia	28
2.4.2.1 <i>Cyc</i>	28
2.4.2.2 <i>Uschold e King</i>	29
2.4.2.3 <i>Gruninger e Fox</i>	29
2.4.2.4 <i>Methontology</i>	30
2.4.2.5 <i>Método 101</i>	30
2.4.3 Ferramentas de manipulação de ontologia	30
2.5 PROCESSAMENTO DE LINGUAGEM NATURAL	32
2.5.1 O analisador sintático PALAVRAS	36
3 METODOLOGIA	39
3.1 DELIMITAÇÃO DO ESTUDO	39
3.2 CARACTERIZAÇÃO DO UNIVERSO ESTUDADO	39

3.3 CARACTERIZAÇÃO DA AMOSTRA	39
3.4 MÉTODOS, TÉCNICAS E INSTRUMENTOS	40
3.5 COLETA DE DADOS	46
3.6 TRATAMENTO DOS DADOS.....	48
4 RESULTADOS.....	50
5 DISCUSSÃO, CONCLUSÃO E CONSIDERAÇÕES FINAIS	57
5.1 DISCUSSÃO	57
5.2 CONCLUSÃO	58
5.3 CONSIDERAÇÕES FINAIS	59
5.4 TRABALHOS FUTUROS	60
REFERÊNCIAS	61
APÊNDICES	76

1 Introdução

A *World Wide Web* já se tornou uma grande biblioteca virtual, onde a informação sobre qualquer assunto está disponível a qualquer hora e em qualquer lugar, com ou sem custo, e em diferentes formatos como textos, áudios, vídeos, etc.

No atual estágio da internet, a quantidade de informação e conteúdo disponibilizado por meio de vídeos vem crescendo a cada dia. Segundo dados da *ComScore*¹, somente em setembro de 2010, foram realizadas 10,6 bilhões de buscas no *Google*. O *site* do Yahoo chegou a receber 2,7 bilhões de buscas no mesmo período.

No âmbito dos conteúdos multimídias, o esforço volta-se para conseguir a recuperação da informação com eficiência através desses motores de busca e outras ferramentas automáticas. Repositórios de vídeos e de imagens estão entre os sites mais acessados da Internet. O *YouTube*, que oferece busca e compartilhamento de vídeos, já é o terceiro site mais acessado do mundo, segundo *Alexa, The Web Information Company*².

Diante de tal realidade, encontrar a informação precisa no *YouTube* é uma tarefa desafiante devido ao tamanho do repositório e a sua grande base de dados textuais (Sureka, 2010).

Assim, Brascher (2002) afirma que:

A necessidade de recuperação de informações armazenadas em grandes repositórios de informação disponíveis na Internet e de responder com maior precisão às buscas realizadas diretamente pelos usuários finais, têm levado a um esforço no sentido de adicionar informação semântica às páginas Web. (BRASCHER, 2002, p.1).

Porém, nesses ambientes onde o usuário pode buscar informação, não ocorre um processo ontológico³ ou de indexação do conteúdo dos vídeos.

As ontologias são estruturadas de tal maneira que permitem um considerável ganho de qualidade quando empregadas num sistema de classificação. Elas oferecem maiores possibilidades estruturais (classes;

¹Pesquisa mensal realizada nos Estados Unidos em setembro de 2010.

²Os dados da Alexa são atualizados diariamente. Acesso em 23/05/2011.

³Entenda por processo ontológico a representação formal de um vocabulário especializado em um domínio.

instâncias; parte-todo; pai-filho; etc.) das que são oferecidas por outros sistemas, como por exemplo, thesauri. A ideia é que o índice, criado a partir de estruturas conceituais geradas por meio do resultado de extensa análise de linguagem natural, apresente um melhor desempenho para as respostas às consultas de usuários (DUQUE, 2006).

Os sistemas de recuperação da informação atuais dependem do uso de termos de indexação na busca para obter os documentos relevantes. Sendo as coleções de documentos indexadas com base no conteúdo textual, os documentos não textuais (imagens, vídeos) dependem de uma indexação manual (associação de termos de indexação) após a interpretação de seu conteúdo.

Para melhorar a recuperação é preciso indexar o conteúdo visual e o áudio, e usar formas mais elaboradas de extração do significado. O sistema de recuperação deve ser capaz de processar documentos e criar descrições que relacionem os objetos, usando termos, ontologias subjacentes e a análise de audiovisual.

Avançar para além da recuperação de texto requer a análise de segmentos de áudio e vídeo. A criação de ontologias explorando tanto descritores audiovisuais com a combinação de ontologias de domínio poderá solucionar o problema?

Essa preocupação pode-se apresentar nos termos seguintes: como armazenar e manipular (processar, gerir) o conhecimento nos sistemas de informação, de maneira formalizada de tal modo feita que possa ser usada por mecanismos para realizar uma tarefa determinada? (ROBREDO, 2005).

1.1 Definição do problema

De que forma a utilização de descritores audiovisuais, conjuntamente com ontologias de domínio, poderá auxiliar na recuperação da informação disponibilizada em vídeos?

1.2 Objetivos da pesquisa

1.2.1 Objetivo Geral

Promover, com maior precisão, a recuperação de uma informação específica existente em um repositório de vídeos por meio da indexação multimodal e uso de

ontologia.

1.2.2 Objetivos Específicos

- Realizar um estudo bibliográfico, revisar a literatura pertinente para selecionar as técnicas de recuperação de informação (RI) mais adequadas para o problema;
- Projetar a utilização da ontologia como parte do processo de RI;
- Indexar vídeos para obter anotações textuais de forma automática a partir de documentos audiovisuais;
- Desenvolver uma ontologia de domínio que contemple os conceitos e termos utilizados nas palestras do II Congresso Espírita Brasileiro;
- Mensurar a precisão do método proposto, utilizando medida padrão, analisando principalmente a diferença entre a recuperação: (1) sem o uso da ontologia,(2) com o uso da mesma.

1.3 Justificativa

O número de usuários interessados na procura de conteúdos em vídeo cresceu significativamente nos últimos anos. Pode-se considerar como exemplo a página do *YouTube*. Criado em 2005, o *YouTube* é um *site* que proporciona hospedagem e visualização de vídeos e já alcançou a terceira posição no ranking dos *sites* mais acessados do mundo em menos de dois anos⁴. De acordo com o *YouTube*, hoje são enviadas 24 horas de vídeo a cada minuto, quantidade essa quatro vezes maior que há 3 anos atrás. Com a crescente quantidade de informação sendo disponibilizada nesses repositórios de vídeos, como recuperar a informação relevante dentre tantos vídeos disponibilizados?

Devido a esse crescimento no interesse pela busca de informação em suportes especiais, identifica-se a necessidade de desenvolver ferramentas voltadas

4 Dados retirados da *Alexa Company*.

para a web que facilitem a busca de informação dentre a grande quantidade de conteúdo ofertado atualmente. O potencial de informação em conteúdos multimídia vem tendo o seu uso aproveitado recentemente por meio de vídeo-aulas, vídeo-conferências, notícias, etc (Armstrong, 2005; Babin,1989). A importância do acesso rápido à informação por meio de novas mídias pode estabelecer uma forma mais dinâmica de comunicação e aprendizagem (Negroponte, 1995).

A recuperação da informação audiovisual geralmente baseia-se na análise individual de um dos descritores disponíveis — áudio, imagem ou texto, e o propósito desta pesquisa é analisar o desempenho da recuperação da informação audiovisual por meio de uma indexação multimodal (áudio e imagem), utilizando conjuntamente uma ontologia de domínio com a intenção de possibilitar uma organização e recuperação mais eficiente da informação em grandes repositórios. Pretende-se ainda avaliar a precisão da busca por trechos específicos nos vídeos.

2 REVISÃO DA LITERATURA

2.1 Trabalhos Correlatos

Modalidade refere-se ao tipo de canal de comunicação usado para transportar ou adquirir a informação. Igualmente cobre a maneira que uma idéia é expressada ou percebida, ou a maneira em que uma ação é executada. Estudos voltados para a indexação de vídeos podem ser encontrados com abordagens unimodal⁵ usando indexação visual, de áudio, ou texto para adicionar informação semântica aos vídeos. Ao invés de usar somente uma modalidade, a indexação de vídeos multimodal procura classificar um documento de vídeo baseado em uma análise de duas ou mais modalidades.

Somente recentemente, foram reportadas abordagens usando combinações multimodais. Entretanto, trabalhos que relacionassem a indexação multimodal, uso de ontologias e vídeos, não foram encontrados. Esses temas são frequentemente encontrados na literatura, mas de maneira isolada.

Em estudo desenvolvido em 2001, Alatan, Akansu e Wolf aplicaram um método utilizando Cadeia Escondida de Markov (HMM⁶) para dividir cenas que contêm diálogos de uma classe de dados de audiovisual (ficção de entretenimento: filmes, séries de TV). Cada *shot*⁷ é classificado com as duas faixas de áudio (via classificação de fala, silêncio e música) e conteúdo visual (rosto e informações de localização).

O resultado desta classificação baseada em *shots* é um símbolo audiovisual para ser usado pelo diagrama de estados HMM para alcançar a análise da cena. Após as simulações com topologias HMM circular, e da esquerda para a direita, observa-se que ambos estão com desempenho muito bom com entradas multimodal. Além disso, para a topologia circular, as comparações entre a formação e conjuntos diferentes de observação mostram que as informações áudio e rosto em conjunto dão resultados mais consistentes entre conjuntos diferentes de observação.

5 Unimodal é a utilização de somente uma modalidade.

6 Utilizaremos a sigla inglesa para denominar a Cadeia Escondida de Markov – *Hidden Markov Model*.

7 *Shot* é originalmente referido como uma sequência contínua de quadros em movimento. Nesta pesquisa, *shot* será uma sequência de quadros que capture o mesmo *slide* da apresentação.

Babaguchi, Kawai e Kitahashi (2002) propuseram a indexação de vídeo baseada em evento, que é uma espécie de indexação pelo seu conteúdo semântico. Como os dados de vídeo são compostos de informações multimodal, como visual, auditiva e textual (legendas), foi apresentada uma estratégia de colaboração intermodal, ou seja, foram utilizados os diferentes canais de informação por meio de um processo colaborativo para a extração de um único significado. O objetivo da pesquisa foi melhorar a confiabilidade e a eficiência na análise de conteúdo de vídeo.

Os mesmos autores citados acima focalizaram ainda na correspondência temporal entre imagem e legendas, com método proposto para buscar intervalos de tempo em que os eventos possam ter ocorrido através da extração de palavras-chave da atual legenda e, em seguida aos *shots* no fluxo visual. Os resultados para o vídeo transmitido de esportes dos jogos de futebol americano indicam que a colaboração intermodal⁸ é eficaz para a indexação de vídeo para eventos, como *touchdown* e *field goal*.

Huang et al. (1999) apresentaram quatro métodos diferentes para a integração de informação de áudio e imagem para a classificação de vídeo baseado em um modelo de Cadeia de Markov (HMM): concatenação direta, o produto HMM, HMM em duas fases, e a integração, através de redes neurais. Seus resultados mostraram melhorias significativas em comparação com o uso de uma única modalidade.

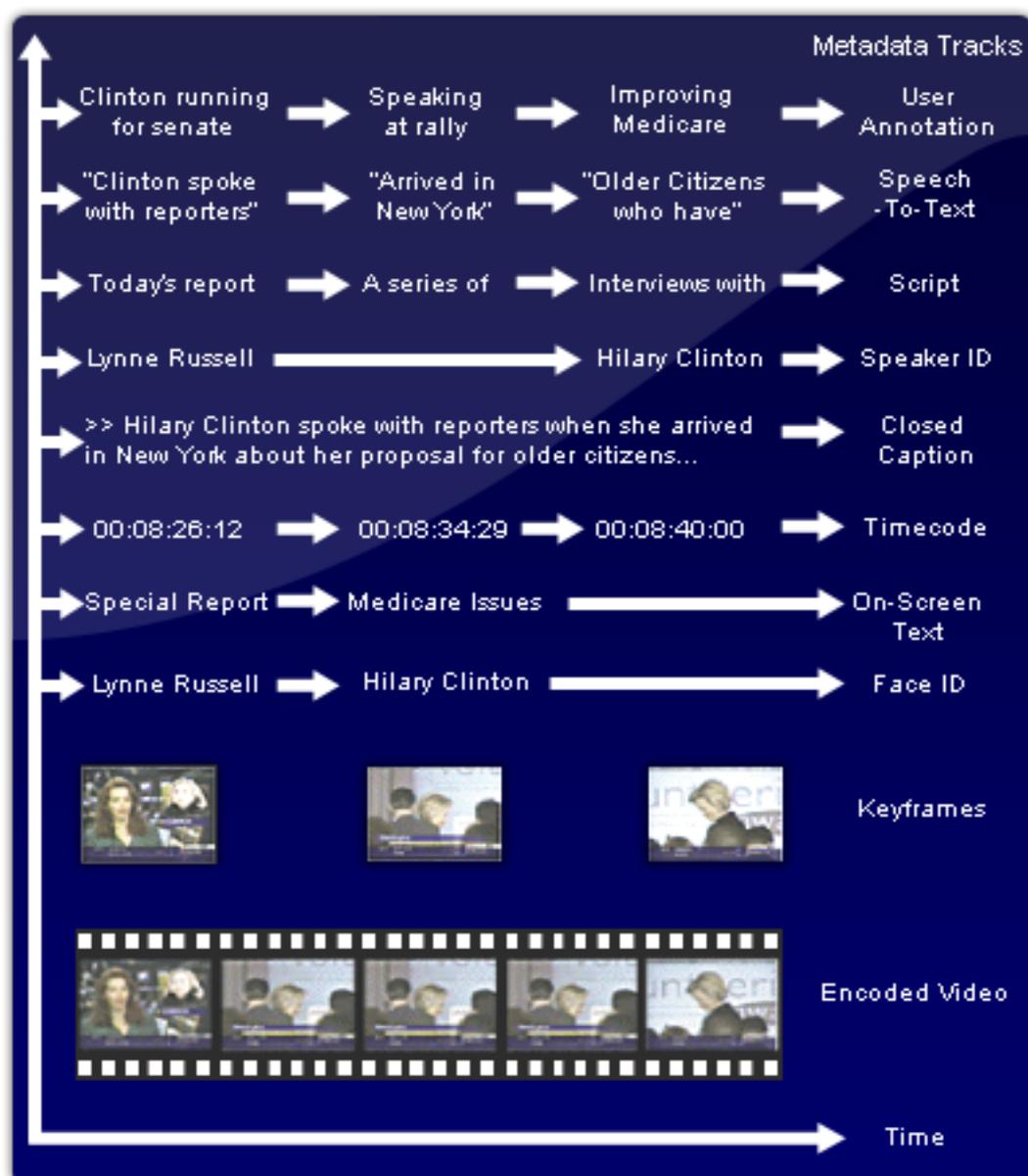
Uma pesquisa desenvolvida por Pfeiffer, Lienhart e Effelsberg (2001), resultou em um sistema para identificar cenas por meio de *shots* de acordo com os diálogos detectados, as configurações de áudio e similares. Resultados da pesquisa mostraram que a identificação automática para determinados tipos de cenas são confiáveis.

Tradicionalmente, conteúdo audiovisual tem sido considerado um recurso que demanda alto custo computacional, exigindo consideráveis horas de trabalho para extrair retornos tangíveis. Como consequência, a maioria das organizações não conseguiram utilizar as informações valiosas contidas nesses recursos, tais como

⁸ Entenda intermodal como a interligação/cruzamento dos diferentes canais de informação indexados.

reuniões gravadas, vídeos de treinamento e conteúdos de radiodifusão. Um exemplo disso é a Autonomia Virage, apresentada na figura 1. Usando motores de análise de áudio e imagem avançados que assistem, ouvem e lêem um sinal de vídeo em tempo real, a Autonomia Virage indexa o arquivo de vídeo para extrair o significado contido na informação.

Figura 1 – Forma de indexação utilizada pela Virage



Fonte: <http://www.virage.com>

Um outro exemplo é o *Yovisto*, uma plataforma Web 2.0 para *upload*, compartilhamento, pesquisa, etiquetagem, e discussão de vídeos acadêmicos. Além da função de um motor de busca de vídeo, o *Yovisto* permite a busca em vídeos de palestras e fornece um mecanismo automatizado para gerar um índice de texto completo dos vídeos. Com *Yovisto* é possível descrever o conteúdo de vídeo de forma colaborativa, além disso, os usuários podem colocar palavras-chave ou etiquetas (*tags*) em qualquer posição dentro de um vídeo (*tagging* colaborativo) e podem iniciar ou continuar discussões pertencentes a qualquer posição dentro do vídeo. Além disso, os usuários podem escrever ou editar páginas wiki para enriquecer o conteúdo de vídeo com mais informações, como imagens, hiperlinks e texto.

A funcionalidade de pesquisa com base em conteúdo do *Yovisto*, como podemos observar na figura 2, é o mais importante benefício adicional para o usuário, e permite uma busca eficiente através de seu repositório.

Figura 2 – Sistema de busca Yovisto

The image shows the Yovisto website interface. At the top, the logo 'yovisto' is displayed in white on a blue background, with the tagline 'Academic video search' below it. A navigation bar contains links for 'Home', 'Search', 'MyYovisto', 'Universities', 'Lectures', 'Videos', and 'Speaker'. Below the navigation bar, there are links for 'standard search' and 'advanced search'. The main content area displays a search result for a video titled '1 06 Labor Relations and International HRM' by Carol Himelhoch. The video is described as 'MGT 360, Human Resource Management' and 'Structure and Functions Local Unions'. A small thumbnail image of the video is shown, along with the text '[video:11774] 4 Views'. At the bottom of the video player area, the score '(Score:1.9969164)' and duration 'Duration: 00:23:46' are displayed.

Fonte: <http://www.yovisto.com>

2.2 Recuperação da informação

De acordo com Araújo Júnior (2007), a recuperação da informação é reconhecida como a recuperação de referências de documentos em resposta às solicitações (demandas expressas por informações). Esse processo pode ser realizado em três etapas: coleta, indexação e ordenação (Duque, 2005).

Segundo Baeza-Yates (1996), as ferramentas de recuperação de informação, geralmente, trabalham com técnicas de indexação capazes de indicar e acessar mais rapidamente documentos de uma coleção de documentos. Ribeiro-Neto e Baeza-Yates (1999) mostram que a motivação para construção de estruturas léxicas (thesaurus⁹) está baseada na utilização de um vocabulário preciso que serve, no contexto da recuperação de informação, para coordenar a indexação e recuperação de documentos.

2.2.1 Recuperação do conteúdo audiovisual

Os conteúdos audiovisuais reúnem naturalmente um conjunto de informações complexas com diversos aspectos semânticos de diferentes tipos e formatos. Tal multiplicidade de informação em um mesmo vídeo tende a dificultar um registro multimídia. Um autor usa de canais visuais, auditivos e textuais para expressar suas idéias, o que torna o índice de um vídeo intrinsecamente multimodal (Snoek e Worrying, 2001).

Deve-se saber que o som presente no vídeo possui diversas formas, como música, fala e sons de ambientes. Além disso, quando o assunto é imagem, pode estar digitalizada em cores, texturas, formas e apresentar diferentes modificações por meio de efeitos especiais. Esta complexidade gerou uma série de estudos na área de recuperação do conteúdo audiovisual.

A primeira geração de sistemas de recuperação de informação audiovisual tinha os atributos de dados de imagem extraídos manualmente. Essa representação

9 Thesaurus é uma estrutura léxica que organiza e relaciona semanticamente as palavras de uma linguagem.

com base em um atributo, implica um alto nível de abstração da imagem e do conteúdo visual inspirado no nível conceitual. Sistemas de representação, como os modelos relacionais e modelos orientados a objetos têm sido utilizados.

Os motores de busca que estiveram em prática, utilizaram-se de duas técnicas textuais e tradicionais, como as linguagens de consulta SQL ou a recuperação de texto completo. O custo da indexação manual sempre foi muito alto, e ainda há o problema da subjetividade das descrições.

Os primeiros anos da Recuperação da informação audiovisual eram freqüentemente baseadas em visão computacional – Ballard e Brown (1982); Levine (1985), e Haralick e Shapiro (1993) – e algoritmos que se concentraram na busca de similaridade baseada em imagens, vídeo e áudio. Influentes e populares exemplos desses sistemas são QBIC (Flickner et al., 1995) e Virage (Bach et al. 1996), apresentados em meados dos anos 90.

Passados alguns anos, o conceito básico da busca por similaridade foi transferido para vários motores de busca da Internet, incluindo imagem Webseek (Smith e Chang, 1997) e Webseer (Frankel et al. 1996). Esforço significativo também foi colocado sobre a integração direta dos recursos de busca por similaridade com as bases de bancos de dados, principalmente de grandes empresas como Informix datablades, DB2 Extenders IBM, Oracle cartridges (Bliujute et al. 1999; Egas et al. 1999) para fazer a recuperação da informação audiovisual mais acessível para a indústria privada.

Na área de recuperação audiovisual, o foco principal em meados da década de 90 foi baseado no limite de detecção robusta, onde as abordagens mais comuns envolviam limiarização pela distância entre os histogramas de cor correspondente a dois quadros consecutivos em um vídeo (Flickner et al. 1995).

Hanjalic et al. (1997) propuseram um método que superou o problema dos limites subjetivos do usuário. Tal abordagem não era dependente de qualquer parâmetro manual, resultava em um conjunto de quadros-chave com base em um modelo objetivo para a informação de fluxo do vídeo.

Haas et al. (1997) descreveram um método de usar o movimento dentro do vídeo para determinar os locais em que ocorriam quadros de fronteira. Seu método superou as abordagens em histogramas do período e também realizou classificação

semântica dos *shots* de vídeo em categorias tais como *zoom-in*, *zoom-out*, *pan*, e assim por diante.

O mais recente guia prático de detecção de transição de vídeo é dado por Lienhart (2001). Perto da virada do século 21, os investigadores observaram que a similaridade baseada em recursos de busca por algoritmos não era tão intuitiva ou amigável ao usuário como eles esperavam. Pode-se dizer que os sistemas construídos por cientistas da pesquisa foram, essencialmente, sistemas que só poderiam ser utilizados de forma eficaz por cientistas.

Uma nova direção foi voltada para projetar sistemas que sejam fáceis de utilizar e que possam trazer o conhecimento multimídia para uma gama de bibliotecas, bases de dados e coleções para o mundo. Para que isso se tornasse possível, observou-se que a próxima evolução dos sistemas precisavam compreender a semântica de um consulta, não apenas o baixo nível de recursos computacionais subjacentes. Esse problema geral foi chamado de "lacuna semântica". De uma perspectiva de reconhecimento de padrões, isto significa traduzir o facilmente calculável baixo nível de conteúdo baseado em recursos de mídia para o alto nível de conceitos ou termos que seria intuitivo para o usuário.

Exemplos de como diminuir a lacuna semântica para o conceito único dos rostos humanos foram demonstrados por Rowley et al. (1996) e Lew e Huijsmans (1996). Talvez o mais antigo sistema baseado em recuperação de conteúdo pictórico, que abordou o problema da lacuna semântica na interface de busca, indexação e resultados foi o motor de busca ImageScape (Lew, 2000). Neste sistema, o usuário pode fazer consultas diretas para vários objetos visuais, tais como o céu, as árvores, a água, e assim por diante, usando ícones posicionados espacialmente em um índice contendo mais de dez milhões de imagens e vídeos com quadros-chave. O sistema utilizou a teoria da informação para determinar as melhores características para minimizar a incerteza na classificação, é o que observa-se na figura 3.

Figura 3 - ImageScape

Fonte: <http://skynet.liacs.nl/imagescape/index.html>

Os motores de busca baseados em similaridade foram úteis em uma variedade de contextos (Smeulders, 2000). Alguns exemplos de busca por similaridade foram utilizados para pesquisar bancos de dados de marcas (Eakins et al, 2003), encontrar *shots* de um vídeo com conteúdo visual e de movimento semelhante, ou para DJs procurando músicas com semelhantes ritmos (Foote, 1999), e detecção automática de conteúdo pornográfico (Forsyth e Fleck, 1999; Bosson et al, 2002). Intuitivamente, as aplicações com maior sucesso são aquelas em que as características básicas, tais como a cor e textura, são encontradas nas imagens e vídeos; ou o ritmo dominante, melodia, ou espectro de frequência de áudio (Foote, 1999) são altamente correlacionados com os objetivos de busca de uma aplicação específica.

Os sistemas de recuperação de informação de terceira geração estão à procura de informações de alto nível em busca de imagens, áudio e vídeo. Quem são os personagens, suas funções, ações e suas relações lógicas, além do impacto emocional do observador, é a informação que deve ser extraída automaticamente, sem ou com mínima intervenção manual, a fim de apoiar a recuperação semântica baseada em metas. Muito mais do que simples imagens, recuperação de vídeo geralmente é significativa somente se executada em altos níveis de representação.

A princípio, oferecer informações importantes transmitidas por diversos canais distintos auxilia o usuário a encontrar a informação multimodal que ele procura. A representação de informações através de vários canais de transmissão que ocorre nos vídeos, como imagem, áudio e contexto aparentemente deve ser mais fácil de recuperar. Discutiremos isso na seção 2.3.2.

2.2.2 Recuperação da informação em palestras

O ensino e a aprendizagem em sala de aula tradicional pode ser visto como uma atividade de autoria multimídia. Palestras, seminários e eventos educativos compartilham do mesmo fim. Os principais fluxos de atividade incluem o que é ouvido (áudio), o que é visto (vídeo) e que é discutido (*slides* da apresentação), em uma sala de aula.

Um dos objetivos essenciais da educação à distância é proporcionar qualidade na aprendizagem, que seja ao mesmo tempo compatível e comparável ao ambiente de sala de aula tradicional (Moore & Kearsley, 1996; Moran, 2000). Para atingir este objetivo, um problema fundamental é como indexar efetivamente as atividades presentes em sala de aula, palestras e seminários para o ambiente multimídia.

Uma abordagem típica é registrar e codificar as atividades em sala de aula como documentos audiovisuais, como nos formatos de áudio e vídeo (Abowd et al., 2000; Deshpande & Hwang, 2001). Os documentos audiovisuais, juntamente com os *slides* eletrônicos associados, poderiam então, ser transmitidos através da rede para a apresentação on-line.

Além da apresentação eficaz, uma forma mais sofisticada de ensino à distância

é apoiar a consulta semântica de documentos audiovisuais (He et al,1999; Mahmood, 2000). Por exemplo, para proporcionar a possibilidade de apoiar a consulta de temas de interesse de um conjunto de vídeos de aulas que foram gravadas por um semestre. Esta aplicação requer especificamente que os “índices de ligação” possam associar explicitamente a cada segmento de vídeo os seus *slides* eletrônicos associados. Por esses índices de ligação, o problema da consulta semântica pode ser transformado no tradicional problema de recuperação da informação.

Em outras palavras, o passo seguinte é permitir que o índice de palavras-chave encontradas em *slides* eletrônicos e o áudio estejam contidos em um banco de dados para recuperação, para que quando fosse realizada uma consulta, um conjunto de *slides* relevantes fossem recuperados. Quando um *slide* relevante é selecionado para exibição, o segmento de vídeo associado será exibido.

Um dos desafios na recuperação da informação em palestras, no entanto, é a modelagem da relação (ou encontrar os índices de ligação) entre os documentos audiovisuais registrados, áudio, imagem/*slides* eletrônicos e seu contexto. Por isso, nesta pesquisa propomos uma indexação multimodal que utilize palavras-chave advindas de áudio, imagem/*slides* e faça uso de ontologias para contextualizar o domínio consultado.

2.3 Indexação

Com o excesso de informação já existente, somada às milhares que são produzidas diariamente no contexto da web, enfrentamos questões sérias relativas a organização e recuperação do conteúdo, e a indexação é um fator determinante para que um usuário que busca por informação recupere documentos que satisfaçam suas necessidades.

A indexação, de acordo com Lancaster (2004), implica na preparação de uma representação do conteúdo temático dos documentos, na qual os termos utilizados servirão como pontos de acesso para a localização e recuperação dos mesmos.

As etapas da indexação são denominadas de análise conceitual e

tradução, na qual a primeira consiste na identificação dos conceitos que melhor representam o assunto do documento, envolvendo, segundo Fujita (2003), a compreensão do conteúdo do documento, identificação dos conceitos que representam este conteúdo e a seleção dos conceitos válidos para recuperação. Já a tradução consiste na conversão da análise conceitual de um documento num determinado conjunto de termos de indexação.

A indexação pode ser realizada de forma manual ou automática. Na indexação automática a tendência é que a especificidade seja maior, pois é utilizada indexação por extração, que possibilita buscas em linguagem natural. Já a indexação manual é feita por atribuição, o que pode gerar uma precisão maior na descrição da informação, já que são atribuídos conceitos que vão além do documento, utilizando-se vocabulários controlados, e baseando-se na subjetividade do indexador. Segundo Fujita (2003):

O bom ou mau desempenho da indexação reflete-se na recuperação da informação feita através de índices. Isso nos leva a considerar que a recuperação do documento mais pertinente à questão de busca é aquele cuja indexação proporcionou a identificação de conceitos mais pertinentes ao seu conteúdo, produzindo uma correspondência precisa com o assunto pesquisado em índices (FUJITA, 2003, p. 62).

Por ser a indexação fator chave para a recuperação da informação, como evidencia a autora, a informação audiovisual possui particularidades tais que exigem uma indexação específica e detalhada, o que será aprofundado a seguir.

2.3.1 Indexação da informação audiovisual

A informação audiovisual, de acordo com Barreto (2007), apresenta grandes desafios relativos a armazenamento, indexação, formulação de consultas e recuperação de conteúdo semântico. Para o autor, um caminho seria a construção de sistemas híbridos, com indexação automática e análise de conteúdo supervisionada, pois existem sérias limitações ao uso de indexadores manuais, uma vez que requerem anotação individual, dificultando seu uso em grandes repositórios de informação, e que sofrem influência tanto do domínio de aplicação quanto do

conhecimento da pessoa que realiza a tarefa. Segundo Lancaster (2004):

A capacidade de armazenar, em formato digital, em bases de dados, qualquer tipo de imagem, e especialmente de poder acessar milhões delas na Rede, causou impressionante ressurgimento do interesse por imagens em geral, e, em particular, por modos de indexá-las (LANCASTER, 2004, p. 214).

Barreto (2007) afirma que o conteúdo visual de imagens pode ser classificado em dois tipos principais:

- conteúdo primitivo de imagens: refere-se aos elementos básicos que compõem a imagem; são características visuais que podem ser reconhecidas e extraídas automaticamente pelo computador com reconhecimento de padrões e visão computacional. Conteúdos primitivos são em geral de natureza quantitativa;
- conteúdo complexo de imagens: refere-se aos padrões de uma imagem que são percebidos por seres humanos como fontes de significados. Dificilmente podem ser identificados por máquinas e são principalmente de natureza qualitativa.

Para Lancaster (2004, p. 285), “o problema da recuperação da informação consiste essencialmente em procurar cotejar aproximações de necessidades de informação com aproximações de mensagens”. Barreto (2007, p. 20) possui visão semelhante, atestando que a recuperação eficaz do conteúdo visual e sonoro só é possível com uma indexação significativa e discriminante, e que deve estar relacionada com intenções e procedimentos do usuário quando faz a consulta no ambiente real. Ambas afirmações nos indicam que o usuário é fator determinante para a indexação, e que esta deverá ser baseada em suas necessidades de informação.

A precisão da busca no processo de recuperação da informação vai demonstrar a utilidade do sistema.

O papel da precisão em um processo de busca e recuperação da

informação é dar a noção exata se o que está sendo recuperado na base de dados é útil ao usuário. Desta conclusão uma série de decisões poderão mudar os rumos de qualidade da resposta que se obtém nos sistemas de recuperação da informação (ARAÚJO JÚNIOR, 2005).

2.3.2 Indexação multimodal

Um dos fatores cruciais na construção de um novo caminho para a recuperação de informação visual é, obviamente, a possibilidade de extrair as informações reais por imagem, áudio etc. Muitas das propostas de recuperação da informação apresentadas por pesquisadores e desenvolvedores de web semântica, é nunca usar dois ou mais canais. Isso está para mudar, pois a análise multimodal tornou-se mais popular, e os benefícios de sua utilização estão cada vez mais evidentes.

Analisando a partir da perspectiva geral do sistema centralizado, a modalidade é a capacidade do sistema para se comunicar com o usuário ao longo de diferentes tipos de canais de comunicação e extrair e transmitir um significado automaticamente (O'Halloran, 2004). Outra perspectiva é a visão que a multimodalidade dos meios de comunicação de vídeo é a capacidade de um autor para expressar uma ideia pré-semântica, através da combinação de um *layout* com um conteúdo específico, usando pelo menos dois canais de informação, onde os canais podem ser visuais, auditivos ou textuais (Snoek e Worring, 2005).

Como o sistema está utilizando várias técnicas de indexação de vídeos, a qualidade dessas técnicas irá certamente afetar a qualidade das informações de saída. Os usuários podem querer limitar o regime apenas às técnicas que passar por testes de qualidade determinados, ou são prestados por organizações específicas ou autores. A aplicação de técnicas multimodais e processos de avaliação podem ajudar a reduzir este problema.

Com a utilização de indexação multimodal, pretendemos compensar os ruídos existentes na indexação de um canal. Nesta pesquisa, o uso de vários canais de comunicação e o auxílio de uma ontologia de domínio irá visar a diminuição de falhas na indexação com resultados favoráveis ao usuário.

2.3.2.1 Indexação do conteúdo imagético em vídeos

Como mostrado na seção 2.2.1, várias abordagens para reconhecimento de imagem têm sido investigadas e descritas pela literatura. O tipo de reconhecimento exigido pelo nosso sistema proposto poderia ser muito simplista, como a recuperação de imagens baseada em conteúdo discutido por Flickner et al. (1995), que evita o uso de descritores textuais e, ao invés disso recupera imagens com base em semelhanças de seus conteúdos (texturas, cores, formas, etc) com uma imagem de consulta fornecida pelo usuário ou características da imagem especificada pelo usuário.

Abordagens mais complexas, que levam a recuperação da imagem em consideração, também podem ser implantadas, como aquelas discutidas por Baeza-Yates (1999) e Smith e Chang (1997).

Esta pesquisa irá indexar vídeos de palestras. Com isso, acredita-se que conseguiremos extrair um conteúdo significativo dos *slides* contidos nos vídeos. Para tanto, é importante conhecer uma forma de indexação imagética específica para tal aplicação.

Quando é preciso converter texto em imagens para texto codificado em caracteres, uma das tecnologias que podem ser utilizadas é o Reconhecimento Óptico de Caracteres, que em inglês é denominado OCR – *Optical Character Recognition*. Nada mais são do que *softwares* que extraem texto através da digitalização do documento. É um procedimento que pode transformar textos impressos, desde que tenham qualidade visual, em material editável. Se o documento possuir figuras, por exemplo, pode tornar o processo de reconhecimento mais lento.

Mello e Lins (1999) coloca que uma transposição de imagem para texto não-automática é inaceitável devido aos custos envolvidos e às baixas velocidade e confiabilidade do processo. Há duas dificuldades básica no uso de OCR's. O bom resultado depende da melhor escolha na hora de transpor de imagem para texto e também da qualidade do documento a ser reconhecido.

Quando os OCR's surgiram em 1959, a leitura dos programas era limitada a uma única fonte, tipo e tamanho fixos. Com o passar do tempo, surgiram novos

tipos de fontes e tamanhos, levando ao desenvolvimento dos programas para identificar de forma mais precisa os caracteres. Para uma margem maior de acertos, fez-se necessário ampliar o que se pode chamar de biblioteca de fontes, ou seja, foi preciso ampliar o número de fontes reconhecíveis pelos OCR's. Para personalizar ainda mais a transcrição, foram criados dois padrões: americano (OCR-A) e europeu (OCR-B), em meados dos anos sessenta.

No final dos anos setenta, um novo sistema prometia identificar qualquer tipo de fonte criado por Kurzweil Computer Products. Foram determinados padrões para reconhecer o desenho das letras mas ainda persistiu a dificuldade de se transcrever imagens de baixa qualidade.

As redes neurais foram introduzidas no final dos anos oitenta. Utilizaram a base de vocabulário para diminuir os erros. O sistema era alimentado com um grande número de caracteres e o computador realizava a pesquisa e identificação por generalização. Pela primeira vez, o programa de OCR foi capaz também de realizar a transcrição de documentos em mal estado. O mais alto nível de desenvolvimento do sistema de OCR foi alcançado nos anos noventa com aperfeiçoamento e uso aprimorado das redes neurais.

Mello e Lins (1999) encontrou vários erros nas pesquisas feitas com OCR. Os textos transcritos tinham problemas quanto a substituição de um caractere por outro, quando estes eram parecidos graficamente ou até mesmo substituir uma letra por duas ou o inverso. Entre os erros mais comuns, também foram encontrados supressão de letras ou de espaços em branco. Até mesmo, o desaparecimento de frases completas, com menor incidência. Todos esses erros são comuns na digitalização de imagens de baixa qualidade.

O protótipo utilizado nesta pesquisa indexa o texto das imagens por meio de reconhecimento óptico de caracteres. Apesar da precisão ser acima de 80%, ruídos podem interferir no desempenho da busca. Por isso, acredita-se que a multimodalidade poderá influenciar significativamente na avaliação final do modelo proposto.

2.3.2.2 Transcrição automática do discurso

As várias formas de reconhecimento de padrões de transcrição de fala têm sido utilizadas para separar o fluxo sonoro e discriminar a voz, a música e os ruídos. Atualmente, alguns sistemas, como o Shazam¹⁰, separam música por ritmo e timbres, e ainda determinam o gênero musical.

Um sistema de reconhecimento de voz independente de locutor pode, por exemplo, extrair informação falada em vídeos. Este é um dos métodos utilizados para auxiliar na recuperação da informação de documentos em áudio, mas ainda está em fase experimental. Segundo Neto, Silva e Sousa (2005), a base de dados para o português brasileiro é insuficiente e atrasa tanto a pesquisa como a colaboração no desenvolvimento dessa área. Nesta busca por pesquisas e ferramentas que reconheçam as especificidades da linguagem falada no Brasil, destaca-se o FalaBrasil.

O FalaBrasil é um grupo de pesquisa criado pelo Laboratório de Processamento de Sinais (LaPS) da UFPA cujo o objetivo é a criação e disponibilização de ferramentas e recursos para reconhecimento de voz em Português Brasileiro.

O grupo concentra-se prioritariamente em ações visando desenvolver e disponibilizar recursos para a construção de sistemas de reconhecimento automático de voz para o Português Brasileiro. Vários recursos já foram desenvolvidos: dicionários fonéticos, modelos de linguagem, modelos acústicos, novos *corpora* de voz e texto, entre outros.

Novos recursos e ferramentas estão sendo desenvolvidos. Dentre os quais destacam-se:

- *Coruja*: um sistema de reconhecimento de voz com suporte a “ditado”. Em contraste com sistemas mais simples, baseados em gramáticas, tal sistema permitirá o desenvolvimento de aplicativos como transcrição forense, criação automática de legenda, entre outros.

10 Disponível em <http://www.shazam.com/>. Acesso em 24/05/2011.

- Tornar o sistema *Coruja* compatível com a interface de programação SAPI (*Speech API*) da Microsoft. Essa iniciativa visa facilitar a utilização do *Coruja* em aplicativos destinados a plataforma Windows.
- Criar um procedimento simples que permita ao usuário fazer o “down-sizing” do *Coruja*, criando versões customizadas para os aplicativos a serem desenvolvidos. Esse processo consiste em reduzir o vocabulário e o modelo acústico, baseando-se na gramática fornecida pelo usuário.
- Implementar um algoritmo para estimar a confiança (“confidence”) de reconhecimento. Permitir que o sistema rejeite hipóteses com baixa confiança. Isso evita que o sistema reaja a palavras que sequer pertençam ao seu vocabulário.
- Construir aplicativos como “prova-de-conceito” tendo em vista pessoas sem formação específica em processamento de voz e que precisem usar a tecnologia. Fomentar a formação de recursos humanos para o desenvolvimento tanto de sistemas de reconhecimento de voz, quanto para aplicativos com interface aural.
- Desenvolver o *SimonBR*¹¹, um sistema para comando e controle em Português Brasileiro que funcione tanto no Linux quanto no Windows. Esse sistema será baseado no projeto Simon¹², um *código aberto para reconhecimento de voz*, e permitirá que usuários possam utilizar a voz para navegar em aplicativos de ambos sistemas operacionais.
- Com base no *Coruja*, disponibilizar um sistema de criação automática de legendas para programas televisivos.

11 O SimonBR é uma versão para português do amplamente utilizado software Simon . O Simon é um programa *open source* idealizado por Peter Gräsch que auxilia pessoas com necessidades especiais a usufruir das funcionalidades computacionais usando somente comandos de voz. Já o pacote SimonBR disponibiliza modelo acústico base, dicionário fonético, interfaces traduzidas e cenários específicos para o Português Brasileiro.

12 Disponível em <http://sourceforge.net/projects/speech2text>. Acesso em 24/05/2011.

- Desenvolver novos algoritmos visando incrementar o desempenho do sistema de reconhecimento de voz. Uma das prioridades consiste no melhoramento do dicionário fonético através do estudo das variações da fala no Brasil e da modelagem acústica através do uso de técnicas para treinamento discriminativo das cadeias escondidas de Markov (HMMs).

Devido à complexidade para prover o uso de tal programa integrado ao protótipo optou-se, portanto, por realizar a transcrição manualmente, com a contratação de um profissional da área. Porém, muito em breve, esta etapa poderá ser realizada de forma automática, dado a iniciativa de alinhamento deste projeto de pesquisa com o grupo de pesquisa FalaBrasil.

2.4 Ontologias

Dado um certo domínio do conhecimento humano, todo e qualquer conceito pertencente a este domínio, bem como suas relações e funções são uma ontologia. Alguns autores defendem que ontologia é uma descrição formal e explícita de uma conceitualização compartilhada (Gruber, 1996). Já outros autores acreditam que ontologia é uma teoria lógica que fornece um relato explícito e parcial de uma conceitualização (Guarino e Giaretta, 1995). Essa forma de organizar o conhecimento visa descrever estruturas conceituais de domínios específicos, tornar viável o uso do vocabulário compartilhado de uma maneira coerente e consistente.

Uma ontologia é criada por especialistas e define as regras que regulam a combinação entre termos e relações em um domínio do conhecimento. Os usuários formulam consultas usando conceitos definidos pela ontologia. O que se busca, em última instância, são melhorias nos processos de recuperação da informação (ALMEIDA; BAX, 2003, p.7).

De acordo com Neches (1991), uma ontologia define os termos básicos e as relações compreendendo o vocabulário de uma área de tópico, bem como as regras para a combinação de termos e as relações para definir as extensões do vocabulário. Esta definição fornece também as linhas gerais para a construção de

uma ontologia: identificar os termos básicos e as relações entre eles; identificar as regras para combiná-los; fornecer definições para tais termos e relações.

Segundo Gruber (1993), uma ontologia é uma especificação explícita de uma conceituação. Nesta definição: conceituação refere-se a um modelo abstrato de algum fenômeno, sendo identificados os conceitos relevantes desse fenômeno; explícita significa que o tipo de conceitos utilizados e as restrições a esse uso são explicitamente definidos; formal refere-se ao fato de que a ontologia deve ser legível por máquina; compartilhada reflete a noção de que uma ontologia captura um conhecimento consensual, isto é, não privativo de um indivíduo, mas aceito por um grupo.

Na visão de Borst (1997) as ontologias são definidas como uma especificação formal de uma conceituação compartilhada.

Para Swartout et al. (1996), uma ontologia é um conjunto hierarquicamente estruturado de termos para descrever um domínio que pode ser usado como um esqueleto fundamental para uma base de conhecimentos.

Segundo Bernaras (1996), uma ontologia fornece significado para descrever explicitamente uma conceituação atrás de um conhecimento representado em uma base de conhecimento.

Feitosa (2005, p. 209), afirma que

Desde o início dos anos 90, as ontologias tornaram-se um tópico de pesquisa popular nas áreas de inteligência artificial, engenharia do conhecimento, processamento da linguagem natural e representação do conhecimento. A razão de tal popularidade é acreditar-se que, por seu intermédio, um entendimento comum e compartilhado sobre um determinado domínio de conhecimento poderá ser comunicado tanto entre pessoas como entre computadores (FEITOSA, 2005, p. 209).

Uma das características principais das ontologias é que elas permitem compartilhar informações de um domínio, podendo ser utilizada por diversas aplicações. Um requisito para compartilhamento é a padronização da linguagem de representação. Nesse sentido, existem diversas linguagens que foram surgindo para a representação de ontologias que serão apresentadas (Corcho e Gómez-Pérez, 2000).

O Resource Definition Framework (RDF), foi desenvolvido para a criação de metadados que permitam a descrição de recursos na Web (World Wide Web

Consortium - W3C, 1999). Um recurso é qualquer coisa que possa ser identificada, física (um país ou um continente) ou lógica (um item de informação em uma página web). RDF é uma linguagem de especificação que permite definir grafos que estabelecem, através de propriedades e restrições de valor, relações entre recursos ou entre recursos e valores.

A Web Ontology Language, OWL (Welty et al., 2004), é uma recomendação apresentada pela W3C como uma linguagem de ontologias para a web. OWL estende, e pode ser vista como uma aplicação de RDF e RDF *Schema* adicionando uma camada de semântica formal. RDF *Schema* nos permite expressar as classes e propriedades, bem como as relações hierárquicas. OWL aumenta a expressividade da linguagem pois permite identificar restrições de propriedade, de equivalência, e quantificadores.

Há três versões de OWL que são de diferentes complexidades:

- OWL Lite é projetado para suportar as tarefas que requerem uma hierarquia de classificação ou taxonomia. Ele suporta apenas um subconjunto das construções da linguagem OWL, e é projetado para permitir o apoio a ferramenta mais simples;
- OWL DL suporta toda a construção de linguagem OWL mas coloca algumas restrições sobre como eles podem ser usados. Isso é para garantir completude computacional e decidibilidade, o que significa que todas as conclusões OWL DL não só são computáveis, mas também irão terminar em um tempo finito. Toda ontologia OWL Lite válida é também uma ontologia OWL DL;
- OWL Full permite a expressão completa da linguagem OWL. Não existem restrições sobre construções (por exemplo, as classes podem ser instâncias de outras classes). Ao contrário da OWL Lite e DL, isso não garante integridade computacional ou decidibilidade. Toda ontologia OWL DL válida é também uma ontologia OWL Full.

Na prática, OWL Lite ou OWL DL são usualmente empregados em função da tarefa e finalidade, enquanto o uso da OWL Full é para usuários que querem a máxima expressividade e a liberdade sintática do RDF sem nenhuma garantia computacional em suas implementações. Para os fins desta dissertação, estamos mais interessados em OWL Lite como uma linguagem para ontologias modelo leve. Embora a linguagem forneça suporte para construções que vão além de simples redes conceituais e semânticas, a disponibilidade de editores como o Protegé (Noy, Sintek et al, 2001) e ferramentas, tais como Jena (Carroll et al, 2003) torna o seu uso atraente.

2.4.1 Tipos de ontologia

Uma tipologia exaustiva de ontologias é apresentada em (VanHeijst et al., 1997) e (Mizoguchi et al., 1995). No entanto, esta seção apresenta os tipos mais usados de ontologias para exibir uma ideia dos conhecimentos a serem incluídos em cada tipo de ontologia. Basicamente, as seguintes categorias são identificadas: uma ontologia de representação do conhecimento, meta-ontologias, as ontologias de domínio, ontologias de tarefas, ontologias de tarefas de domínio, ontologias de aplicação, índice de ontologias, as ontologias “dizer e pedir”, etc:

- Ontologias de Representação do Conhecimento (vanHeijst et al., 1997) explicam as conceitualizações que estão por trás dos formalismos de representação do conhecimento. O exemplo mais representativo deste tipo de ontologia é a *Frame* Ontology (Gruber, 1993), que capta as primitivas de representação (classes, instâncias, slots, facetas, etc) usadas em linguagens baseadas em *frames*;
- Geral / Ontologias Comuns (Mizoguchi et al., 1995) incluem o vocabulário relacionado a coisas, eventos, tempo, espaço, causalidade, comportamento, função, etc;
- Meta-ontologias, também chamadas de ontologias genéricas ou Ontologias

Core (vanHeijst et al., 1997), são reutilizáveis em todos os domínios. O exemplo mais representativo pode ser uma ontologia mereologia¹³ (Borst, 1997), que incluem o termo parte de;

- Ontologias de domínio (Mizoguchi et al., 1995) (vanHeijst et al., 1997) são reutilizáveis em um determinado domínio. Elas fornecem vocabulários sobre os conceitos de um domínio (bisturi, raio-x) e seus relacionamentos, sobre as atividades que ocorrem nesse domínio (anestesiari, dar à luz), e sobre as teorias e princípios elementares que regem esse domínio;
- Ontologias de tarefas (Mizoguchi et al., 1995) fornecem um vocabulário sistematizado de termos utilizados para resolver problemas relacionados com as tarefas que podem ou não ser do mesmo domínio. Essas ontologias fornecem um conjunto de condições por meio do qual pode-se descrever genericamente como resolver um tipo de problema. Incluem nomes genéricos (plano, meta, restrição), os verbos genéricos (atribuir, classificar, selecionar), adjetivos genéricos e outros no agendamento de tarefas;
- Ontologias de domínio de tarefas são ontologias de tarefa reutilizáveis em um determinado domínio, mas não entre domínios. Por exemplo, uma ontologia de tarefas no domínio da saúde podem incluir os termos relacionados com o agendamento de uma cirurgia: plano de cirurgia;
- Ontologias de aplicação (vanHeijst et al., 1997) contêm o conhecimento necessário para a modelagem de um domínio específico.

2.4.2 Métodos de desenvolvimento de ontologia

2.4.2.1 Cyc

Em 1984, o projeto CYC foi desenvolvido por Lenat e seus colaboradores. A

¹³ Teoria ou estudo lógico-matemático das relações entre as partes e o todo e das relações entre as partes no interior de um todo.

intenção era apresentar um sistema capaz de interpretar o senso comum e torná-lo compreensível às máquinas. Criada a partir da experiência CYC KB, era como se o sistema comportasse uma imensa quantidade de informações para usá-las como base para novas interpretações, nem sempre pontuais. A codificação era feita manualmente, representando o conhecimento, seja ele implícito ou explícito, com base na linguagem natural (Lenat, Guha, 1990; Lenat, 1995; Reed, Lenat, 2002).

A linguagem de representação do CYC, a CYCL, foi desenvolvida por Lenat e Guha em 1990. Esse tipo de linguagem ampliou os resultados de buscas mas ainda houve falha quanto a categorização, por exemplo. O sistema trabalha de forma dedutiva, o que pode provocar erros de interpretação.

2.4.2.2 Uschold e King

O método foi proposto inicialmente por Mike Uschold e Martin King em 1995 (Uschold e King, 1995) e estendido em 1996 por Mike Uschold e Michael Gruninger (Uschold e Gruninger, 1996). A Enterprise Ontology usa métodos de integração e ferramentas com base em uma modelagem organizacional. O glossário de termos em linguagem natural é disposto em classes.

O Método de Uschold & King tem como base para a construção da ontologia: a identificação dos conceitos-chave e dos relacionamentos; a definição textual e não ambígua dos conceitos e relacionamentos; a codificação em linguagem formal e, por fim, a integração a outras ontologias, caso existam.

2.4.2.3 Gruninger e Fox

A metodologia foi proposta por Michael Gruninger e Mark Fox em 1995 (Gruninger; Fox, 1995), tendo como base para o seu desenvolvimento a experiência obtida no projeto Toronto Virtual Enterprise - conhecido como projeto Tove (Fox, 1992).

O projeto tem por objetivo criar ontologias, para modelar organizações públicas e privadas, levando em consideração as seguintes características: criação de cenários de motivação, formulação de questões de competência, especificação

da ontologia em uma linguagem formal, formulação das questões formais de competência, especificação de axiomas utilizando a terminologia da ontologia, verificações de teoremas completos.

2.4.2.4 Methontology

A Methontology é baseada nos padrões de desenvolvimento de softwares pelo IEEE (IEEE, 1990). Desenvolvida entre 1996 e 1997 em Madri, a Methontology respeita os seguintes estágios de desenvolvimento: especificação, conceitualização, formalização, integração, implementação e manutenção. Esse método também leva em consideração o ciclo de vida da ontologia.

2.4.2.5 Método 101

A principal característica do Método 101 é o enfoque iterativo. Noy e McGuinness determinam sete passos do método: 1) determinar o domínio e escopo da ontologia, 2) considerar o reuso de ontologias existentes, 3) enumerar os termos importantes da ontologia, 4) definir classes e hierarquias de classes, 5) definir as propriedades das classes, 6) definir os valores das propriedades e 7) criar instâncias individuais das classes de hierarquia.

O método 101 foi concebido por Natalya F. Noy e Deborah L. McGuinness (Noy; McGuinness, 2001) a partir da experiência no desenvolvimento da ontologia de vinhos e alimentos, utilizando o ambiente de edição de ontologias Protégé-2000 (Horridge et al., 2004).

Neste trabalho utilizamos este método por apresentar passos bem definidos e explicados, de forma que é possível segui-los com facilidade.

2.4.3 Ferramentas de manipulação de ontologia

Dada a complexidade na elaboração de ontologias, são necessárias ferramentas que auxiliem nos critérios de construção das mesmas. Para construir ontologias, há diversas ferramentas que podem ser utilizadas:

ONTOLÍNGUA

Utiliza paradigmas das linguagens baseadas em *frames* e lógica de primeira ordem. Possibilita a construção de ontologias compartilhadas entre grupos (Chaudri et al., 1998). Pode ser usado para criação, edição e uso de ontologias. Este sistema foi desenvolvido e é mantido pelo Knowledge Systems Laboratory da Universidade de Stanford. É citado por Fonseca e Egenhofer (1999), Novello (2002), Guizzardi (2000). O sistema *Ontolingua* está disponível em: <http://www.ksl.stanford.edu/software/ontolingua/>.

Para permitir definições muito completas, qualquer ontologia criada na Ontolingua deve incluir e referenciar uma ontologia complexa para lidar com frames, a Frame-Ontology (Gruber, 1995), e o código gerado ou traduzido pode tornar-se complexo devido às constantes referências às classes desta ontologia.

ONTOSAURUS

Ontosaurus (Swartout et al, 1997) possui interface web acessível de qualquer navegador e representa o conhecimento na linguagem de programação LOOM1 (Ding, 2001), permitindo tradução para Ontolingua, KIF e C++. Trata-se de um servidor de ontologias e um servidor de navegação por ontologias.

ONTOEDIT

Trata-se de uma ferramenta para edição de ontologias via Internet, que implementam as visualizações arbórea (folder-tree), hiperbólica (Hyperbolic Tree) e de grafos (usando o TouchGraph). Essa ferramenta permite a criação de ontologias a partir de arquivos texto (com extensão TXT) tabulados.

O OntoEdit foi construído sobre várias tecnologias de desenvolvimento web, como por exemplo: Navegador da Árvore Hiperbólica⁵, Applets Java, JavaScript e PHP em conjunto com banco de dados MySQL.

Esse tipo de ferramenta concentra a vantagem de trabalhar no ambiente web, assim como possibilita a visualização da estrutura da ontologia no formato de árvore hiperbólica e de grafos. Outro ponto a se destacar é a capacidade que esse

programa tem de converter uma ontologia em TXT em visualização *folder-tree* a partir de *upload*.

PROTÉGÉ

É um editor de ontologias, usado para o desenvolvimento de sistemas baseados em conhecimento. O software está disponível gratuitamente no site <http://protege.stanford.edu>. Foi desenvolvido pelo departamento de informática médica da Universidade de Stanford, dos EUA, e modernizado gradativamente para acompanhar a evolução da tecnologia de SBC (Sistemas Baseados em Conhecimento). O Protégé suporta classes que em uma ontologia descrevem conceitos pertencentes a um determinado domínio, além disso tem por objetivo ser uma ferramenta de aquisição de conhecimento extensível de fácil configuração. É a ferramenta mais utilizada mundialmente.

O Protégé 2000 foi desenvolvido em Java e inclui uma API (interface de programação) para estender o ambiente e suportar aplicações em domínios específicos (Santos, 2001). É uma ferramenta representada por meio de classes, instâncias dessas classes, *slots*, *facet*as e axiomas. Usa a arquitetura de metaclasses, ou seja, um *template* que é usado para definir novas classes em uma ontologia. Possibilita a especificação de herança múltipla e de classes abstratas.

Essa ferramenta foi desenvolvida em diferentes linguagens para a web semântica, que permite a estruturação de informações de um domínio específico e possibilita a comunicação, por meio de um vocabulário comum, entre agentes de software e páginas da web.

Por sua facilidade e extensibilidade de formatos e *plugins* utilizaremos o Protégé para desenvolver a ontologia nesta pesquisa.

2.5 Processamento de Linguagem Natural

O processamento de linguagem natural (PLN), também conhecido como lingüística computacional, é o estudo da interface da informática e da lingüística. Ou seja, trata-se do desenvolvimento de sistemas com capacidade para produzir informação em linguagem natural (Gagnon, 2000).

Geralmente, esse tipo de processamento é aplicado às linguagens escrita e falada. Deve-se porém dar mais atenção à linguagem escrita devido a ausência de entonação, caso da fala, por exemplo. O processamento automático da linguagem natural requer conhecimentos da Computação e da Lingüística.

Na década de sessenta, podemos identificar quatro categorias de programas em linguagem natural. Alguns programas, como o ELIZA (Weizenbaum, 1966), geravam um número reduzido de resultados em domínios específicos. Já em alguns dos primeiros sistemas, como o PROTO-SYNTAX1, utilizava-se indexação para auxiliar na recuperação de determinados termos. Havia ainda aqueles sistemas de lógica limitada, como o CONVERSE, com o objetivo de converter frases digitadas em notação formal para o banco de dados. E, por último, os sistemas com base de conhecimento, como o LUNAR (Woods, 1973), que usavam informação de um determinado assunto para representação semântica no banco de dados.

Segundo Sag & Wason (1999), ainda que as tecnologias existentes funcionem de forma eficiente, o conhecimento mais aprofundado da linguagem natural pode trazer benefícios para o aprimoramento desses sistemas. Nesse sentido, faz-se necessário um tratamento conjunto e complementar da computação e da lingüística.

Russell and Norving (2003) descrevem a linguagem informal como um conjunto possivelmente infinito de cadeias que seria um dos princípios fundamentais do PLN. Seriam ainda princípios fundamentais adotados por esse tipo de processamento a cadeia, o símbolo terminal e a gramática. Todos esses princípios são também utilizados no tratamento da linguagem formal.

É possível dar às linguagens naturais o mesmo tratamento dado às linguagens formais. Para tanto, precisamos de uma gramática, ou seja, uma especificação da estrutura das sentenças da linguagem.

É a partir dos formalismos gramaticais que é possível entender melhor como funciona a busca de sentenças, por exemplo. Esses formalismos são utilizados para o desenvolvimento de softwares. Segundo Chomsky (1957), as gramáticas podem ser regulares, livres de contexto, sensíveis ao contexto e recursivamente enumeráveis. Esta última seria a mais ampla de todas por permitir reconhecer um número maior e mais completo de expressões com menos restrições.

As estratégias de reconhecimento dos softwares podem ser adotadas de diferentes formas. Segundo Luger (2005), todos esses programas utilizam algoritmos no caso de buscas quanto à linguagem escrita.

É importante considerar os tipos de análises de texto para compreensão dos comandos solicitados à máquina. São as seguintes:

- Análise Morfológica

Cada palavra é analisada e classificada de forma isolada, levando em consideração a sua própria estrutura.

- Análise Sintática

Analisa-se uma sequência de palavras de forma linear a fim de analisar o seu relacionamento e emprego na frase.

- Análise Semântica

Trata-se de um processo mais complexo. Cada estrutura analisada recebe um significado. É nessa etapa que incorre o risco de ambigüidade. Uma frase escrita com estrutura sintática correta pode não ter equivalência semântica nenhuma. Para que isso não aconteça, Rich & Knight (1993) sugerem que seja feito um processamento léxico para identificar o significado de cada palavra.

- Integração de Discurso

É quando a frase só passa a ter um sentido completo a partir da relação com outras frases anteriormente citadas. Assim, cada parte poderá ser entendida.

- Análise Pragmática

Os enunciados são analisados dentro de um determinado contexto, perpassando o sentido literal afim de encontrar o que está implícito na comunicação. Assim, as palavras podem ser melhor interpretadas para se ter uma resposta do programa. Existem alguns problemas na recuperação de informação comumente enfrentados pelos sistemas de buscas atuais. Entre eles, pode-se destacar a ambigüidade.

Quando se fala de ambigüidade, trata-se de uma frase ou um termo que possui mais de uma interpretação do seu significado. Por exemplo, quando uma pessoa diz que “O dinheiro estava no banco”. Há ambigüidade, causando dupla interpretação do termo “banco”. Pode-se entender como: “assento” ou “instituição

financeira”.

É nesse ponto que pode ocorrer um ruído na comunicação. Ou seja, uma interferência capaz de alterar a compreensão desejada e até mesmo impedir essa compreensão.

Segundo Fuchs (1988), a ambigüidade pode ser classificada da seguinte forma: morfológica, lexical, sintática, predicativa, semântica e pragmática. No caso dos sistemas de buscas, precisa-se ter um cuidado especial quanto à ambigüidade lexical bem como na interpretação de dados no contexto. Quanto ao léxico, a ambigüidade pode ser provocada por homonímia ou polissemia.

A polissemia ocorre quando uma palavra pode ter mais de um significado, com a mesma forma ortográfica e fonológica. À medida que uma significação nova é dada à palavra, parece multiplicar-se e produzir exemplares novos, semelhantes na forma, mas diferentes no valor. A esse fenômeno de multiplicação chamamos de polissemia. Todas as línguas das nações civilizadas participam desse fenômeno; quanto mais um termo acumulou significações, mais se deve supor que ele represente aspectos diversos da atividade intelectual e social (Bréal, 1992).

Segundo Lyons (1987, p. 143), “enquanto a identidade entre as formas é uma questão de sim ou não, a relação entre os significados é um problema de mais e ou menos”. Para Ullmann (1964), a polissemia pode levar à ambigüidade no contato entre as línguas, no uso técnico e científico e na fala vulgar.

Quanto à ambigüidade, Berruto (1979) indica que a homonímia deveria se diferenciar segundo a realização oral (fônica) e escrita (gráfica) dos significantes, ou seja, entre homofonia (sons iguais) e homografia (grafia igual). “Homônimos são palavras que têm formas idênticas, mas que expressam conteúdos distintos. Em outras palavras: significantes idênticos se referem a significados diferentes”. (Biderman, 1978).

Silva (1989) buscou diferenciar polissemia e homonímia. A pesquisa realizou testes com falantes de língua portuguesa e atribuiu valores para similaridade das palavras. O resultado levou à conclusão de que os falantes da língua têm consciência da diferença entre os dois fenômenos e possuem mecanismos “desambiguadores” para sua própria compreensão.

2.5.1 O analisador sintático PALAVRAS

O PALAVRAS é um analisador automático (*tagger-parser*) para português que foi desenvolvido por Eckhard Bick no contexto de um projeto de doutoramento (1994-2000) na Universidade de Århus (Dinamarca). O sistema apoia-se num léxico de 50.000 lemas e milhares de regras gramaticais para fornecer uma análise completa, tanto morfológica como sintática, de qualquer texto. O formalismo aplicado integra-se na tradição da Constraint Grammar, introduzido por Fred Karlsson (Universidade de Helsínquia, Finlândia) em 1992. Embora usando um conjunto de etiquetas gramaticais bastante diversificado, o *Parser* alcança um nível de correção de 99% em termos de morfologia (classe de palavras e flexão), e 97-98% em termos de sintaxe. Para trabalhar com o PALAVRAS, é possível utilizá-lo por meio do *síte* do projeto VISL (<http://visl.sdu.dk>). Na figura 4 é possível ver uma tela de exemplo, onde é representada a análise sintática de um pequeno trecho da pesquisa utilizando o PALAVRAS.

Figura 4 – Tela do analisador sintático PALAVRAS

Flat structure

Enter Portuguese text to parse:

E passou a meditar, quando, repentinamente, percebeu a aproximação de um vulto espiritual. Esse vulto se identificou e era Emmanuel.

Go!

Reset

Parser: Visualization:

e [e] **KC**
passou [passar] <ve> <vt> **V PS 3S IND VFIN**
a [a] **PRP**
meditar [meditar] <vt> <vi> **V INF 0/1/3S**
,
quando [quando] <rel> <ks> **ADV**
,
repentinamente [repentinamente] **ADV**
,
percebeu [perceber] <vt> **V PS 3S IND VFIN**
a [o] <artd> <dem> **DET F S**
aproximação [aproximação] <event> **N F S**
de [de] **PRP**
um [um] <quant> <arti> **DET M S**
vulto [vulto] <H> <f-an> <am> **N M S**
espiritual [espiritual] **ADJ M/F S**
.
esse [esse] <dem> **DET M S**
vulto [vulto] <H> <f-an> <am> **N M S**
se [se] <refl> **PERS M/F 3S/P ACC/DAT**
identificou [identificar] <vt> **V PS 3S IND VFIN**
e [e] **KC**
era [ser] <vK> **V IMPF 1/3S IND VFIN**
Emmanuel [Emmanuel] **PROP M S**

Fonte: <http://visl.sdu.dk>

O VISL (*Visual Interactive Language Learning*) foi lançado como um projeto de pesquisa e ensino na Universidade do Sul da Dinamarca em 1996, e está atraindo apoio financeiro de várias fontes externas. Partindo do sistema português PALAVRAS como modelo para outras línguas, a equipe VISL construiu um núcleo de ferramentas e bancos de dados linguísticos para usar através da internet. Trabalha-se hoje com a gramática, e especificamente a sintaxe, de 14 línguas, entre elas 6 com análise automática. As áreas mais recentes de atividade são semântica e

tradução automática, como também a coleção e etiquetagem de corpora.

Ao lado do sistema aberto, foi estabelecida uma base de orações controladas para todas as línguas VISL, cobrindo vários fenômenos sintáticos de uma maneira mais sistemática. Na interface de ensino, usuários podem escolher entre diversos filtros notacionais, apoiando diferentes paradigmas descritivos da língua. Exemplos são exercícios nos quais as palavras são coloridas para marcar sua classe, e árvores de sintaxe gráficas construídas pelo estudante e controladas pelo computador, com etiquetas de forma e função em cada nó.

Neste trabalho, iremos fazer uso do analisador sintático PALAVRAS por meio do *site* do projeto VISL.

3 METODOLOGIA

3.1 Delimitação do estudo

Esta pesquisa procurou analisar a precisão na recuperação da informação disponibilizada em vídeos por meio do uso de ontologia e indexação multimodal.

A pesquisa limitou-se a tratar a questão da recuperação da informação audiovisual no que tange à precisão e por meio de indexação de imagem, áudio, análise sintática e uso de ontologia.

3.2 Caracterização do universo estudado

O universo em estudo é o repositório de vídeos da Televisão do Conselho Espírita Internacional (TVCEI). A TVCEI é uma iniciativa do Conselho Espírita internacional (CEI), instituição resultante da união, em âmbito mundial, das Associações Representativas dos Movimentos Espíritas Nacionais, de 32 países. O CEI tem como meta principal a divulgação do espiritismo no mundo e a TVCEI é um de seus órgãos de comunicação. Iniciou suas atividades como WebTV em agosto de 2006. Em junho de 2009, iniciou seu sinal no Satélite (Estrela do Sul) para todo o Brasil e agora está presente em mais de 20 operadoras de TV a cabo. Foram realizadas parcerias com produtoras de programas de TV Espíritas, Entidades Federativas, Instituições de grande porte, além de renomados expositores do movimento espírita.

3.3 Caracterização da amostra

Na programação da TVCEI, diversos tipos de vídeos são encontrados, tais como palestras, desenhos animados, eventos, biografias, filmes e etc. A amostra foi selecionada obedecendo aos seguintes critérios: somente farão parte da pesquisa vídeos-palestras que utilizaram *slides*, os demais vídeos integrantes da programação não serão estudados. Sendo assim, utilizaremos uma amostra de 6 horas de vídeos do universo descrito.

3.4 Métodos, técnicas e instrumentos

Para a pesquisa, foi desenvolvido um protótipo. O protótipo foi baseado em uma solução aberta (Schulte et al., 2008) desenvolvida em linguagem *Java*¹⁴ para integrar, posicionar, arquivar, e distribuir gravações audiovisuais em grande escala. Desta forma, um dos módulos do protótipo criado possui a característica de segmentar os vídeos a cada novo slide identificado. O usuário tem por meio do protótipo a visualização do vídeo completo e dos segmentos de vídeo. A figura 5 mostra a tela do usuário ao acessar um dos vídeos indexados pelo protótipo.

Figura 5 – Tela do protótipo



Fonte: Dados da pesquisa

A partir da segmentação do vídeo e identificação dos *slides* foi possível utilizar o OCR para extrair o conteúdo imagético. O protótipo utiliza a tecnologia de OCR alinhada a um dicionário pré-definido do Português Brasileiro para aumentar a

¹⁴ <http://www.java.com>

taxa de acerto no reconhecimento das palavras. Assim, além de segmentar em *clips* de vídeo, todas as palavras constantes nos *slides* dos palestrantes nos vídeos foram indexadas sincronicamente. A figura 6 mostra um trecho de palavras indexadas relacionadas com a linha do tempo do vídeo.

Figura 6 – Trecho indexado relacionado ao tempo do segmento

00:56:16

"Dessa feita, 'O Livro dos Médiuns', que justamente agora, em 1961, está celebrando o primeiro centenário, foi objeto de nossa especial atenção" e considera que "os campos da Ciência e da Filosofia, nos domínios doutrinários do Espiritismo, são continentes de trabalho a se perderem de vista..."

00:57:28

"diante do Centenário do 'O Evangelho Segundo o Espiritismo' (...) este livro desprezioso de servidor reconhecido, como sendo Livro da Esperança". O autor espiritual extravasa sua emoção: "Oh! Jesus! No luminoso centenário de 'O Evangelho Segundo o Espiritismo', em vão tentamos articular, diante de ti, a nossa gratidão jubilosa! (...) – Obrigado, Senhor!..."

XAVIER, Francisco Cândido. Livro da Esperança. Pelo Espírito Emmanuel. 4.ed. Uberaba: Edição CEC, 1973, p.11-12.

00:58:26

00:58:34

"... traçamos os despreziosos comentários contidos neste volume, em torno das instruções relacionadas no livro 'O Céu e o Inferno' (...) dando continuidade à tarefa de consultar a essência religiosa da Codificação Kardequiana..."

XAVIER, Francisco Cândido. Justiça Divina. Pelo Espírito Emmanuel. 10. Ed. Rio de Janeiro: FEB. 2002, p.12.

00:59:29

00:59:40

A fidelidade de Emmanuel ao Codificador está explicitada no texto "O

Fonte: Dados da pesquisa

O áudio foi outro canal de comunicação indexado nesta pesquisa, e apesar de existirem algumas propostas, modelos, projetos e até soluções abertas para indexação automática do áudio (Silva et al, 2009; Lee e Kawahara, 2009; Silva, 2005; Teruszkin e Vianna, 2006), utilizamos a indexação manual. Esta escolha deve-

se pela necessidade de adaptação computacional com o protótipo, o que demanda maior complexidade computacional. Assim, fizemos a indexação manual e adicionamos as palavras indexadas ao protótipo. O protótipo contém as palavras resultantes da indexação multimodal. É importante salientar que as palavras indexadas foram relacionadas em sincronia com a ordem na qual aparecem nos *slides* ou foram pronunciadas nos vídeos.

Após a indexação multimodal das palestras, foi realizado um processamento de linguagem natural. O conjunto de frases indexadas foi analisado por meio do *parser* PALAVRAS. Os termos mais significativos para esta pesquisa foram determinados como sendo a principal palavra do sujeito e os substantivos contidos no predicado. As palavras extraídas, conforme mostra a figura 7, foram resultantes do processamento pelo *parser* PALAVRAS.

Figura 7 – Palavras-chave após indexação multimodal e processamento pelo parser PALAVRAS

00:00:00
obra psicográfica

00:00:12
**lances trajetória atualidade espírito Emmanuel Francisco
 Cândido Xavier médium criança anos idade ano poesias
 dissertações Emmanuel Trajetórias Obras se início
 romances Chico Xavier autoria espírito nós forma
 Antonio Cesar Perri Carvalho Diretor Federação Espírita
 Brasileira Tema Emmanuel : trajetória obras Romances
 de Emmanuel**

00:01:11
**lances trajetória atualidade espírito Emmanuel Francisco
 Cândido Xavier médium criança anos idade ano poesias
 dissertações**

00:01:46
**almanaques revistas Brasil época ano Chico Xavier final
 tarde açude bairro periferia cidade Pedro Leopoldo
 interior Minas Gerais ele aproximação vulto vulto se
 Emmanuel**

00:02:24
**se diálogo nós Emmanuel oportunidade disposição
 médium Chico Xavier trabalho mediunidade Jesus base
 evangelho Jesus médium indagação ele convite espírito
 se Emmanuel
 Pedro Leopoldo mG Emmanuel disposição medium
 Chico Xavier trabalho mediunidade Evangelho de Jesus
 Para compromisso medium pontos disciplina disciplina**

Fonte: Dados da pesquisa

Parte dos objetivos desta pesquisa é melhorar de forma significativa a recuperação da informação utilizando-se recursos automáticos ou semiautomáticos disponíveis, através do uso de uma ontologia. Para tanto, utilizamos o Método 101 (Noy; McGuinness, 2001) no desenvolvimento da ontologia deste trabalho. A ontologia aborda conceitos relativos à Doutrina Espírita, com foco nos conceitos utilizados e difundidos atualmente. O propósito da ontologia é realizar o mapeamento dos conceitos da Doutrina Espírita para integração com o protótipo, que foi implementado para auxiliar na recuperação da informação. Na elaboração de

uma ontologia utilizada na recuperação da informação, devem-se contemplar os principais conceitos (palavras-chave) utilizados nos documentos da coleção, e conceitos que, em geral, ajudam a distinguir a relevância dos documentos. Deste modo, a ontologia deve definir, de forma não ambígua, todos os conceitos e relações deste domínio. A ontologia desenvolvida foi usada com o intuito de expandir a rede semântica e diminuir ambiguidades.

Neste trabalho consideramos a hipótese de trabalhar com reuso de ontologias porém, uma ontologia com o domínio específico “Doutrina Espírita” não foi encontrada.

A lista de termos foi obtida por meio de pesquisas em glossários *online*, textos da área e principalmente a obra “Doutrina Espírita para principiantes”, de Luis Hu Rivas, 2007. O entendimento dos termos relativos ao domínio foi facilitado pelo acompanhamento de um especialista do domínio, o próprio autor Luis Rivas, na elaboração da ontologia.

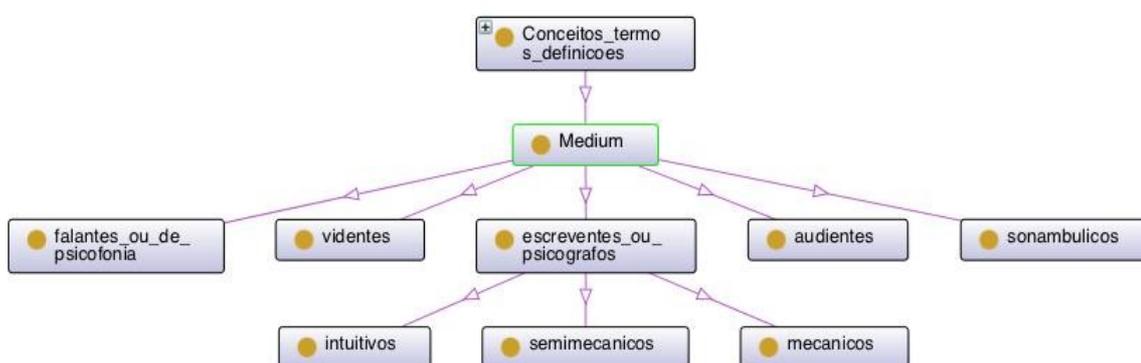
Na definição da hierarquia de classes utilizamos inicialmente a abordagem *top-down* e em seguida, uma combinação das abordagens *top-down* e *bottom-up*. Por exemplo, primeiro foram definidas as subclasses da classe *Principios fundamentais* e da classe *Ser*. A partir da definição das condições destas classes foram definidas as subclasses da classe *Leis Morais*. Durante este processo algumas classes foram reagrupadas; é o caso da classe *lei causa e efeito*, subclasse da classe *Reencarnacao*.

Para limitar o domínio da Doutrina Espírita a um sistema de conceitos e categorias compreensíveis pelo computador, o processo de desenvolvimento começou pela modelagem do domínio da Doutrina com termos e definições, como: “*Passe*”, “*Medium*”, “*Leis Morais*”, “*Ser*” e detalhando cada um desses blocos em sub-classes em uma relação *e_um*. Assim, um Medium poderia ser um vidente, um psicógrafo, um audiente, etc. Após isso e da análise de palavras-chave extraídas da indexação multimodal dos vídeos, essas palavras-chave foram analisadas para verificar se apareciam como conceitos ou relações na ontologia. Dessa forma, surgiu a necessidade de criar super-classes adicionais tais como “*Obras*” e “*Nomes Proprios*”. A seguir fez-se uma listagem dos sinônimos para as palavras presentes na ontologia, considerando os termos relevantes, e que poderiam ser utilizados nas

consultas do sistema. Os sinônimos foram transformados em classes equivalentes.

O desenvolvimento dos relacionamentos entre classes e a adição de instâncias ao modelo foram os últimos passos do desenvolvimento dessa primeira versão da ontologia. A especificação de instâncias na ontologia foi um passo importante, pois muitas palavras ditas durante as palestras tratam de indivíduos e não das classes propriamente ditas. Uma parte da hierarquia presente abaixo da classe “Medium” na ontologia é mostrada na Figura 8.

Figura 8 – Parte da hierarquia abaixo da classe “Medium” na ontologia do domínio



Fonte: Dados da pesquisa

Esse processo de desenvolvimento foi iterativo, e gerou uma ontologia reutilizável no domínio da Doutrina Espírita, que possa ser ampliada à medida em que mais palestras forem utilizadas como referência.

Na pesquisa, foi necessária a aplicação OnAir v1.0 (2010) que permitisse ao usuário realizar consultas em diferentes configurações, inclusive fazendo o uso da ontologia de domínio criada. Podemos visualizar a aplicação na figura 9 a seguir:

Figura 9 – Representação da área administrativa para configuração das consultas

Arquivo

Personagem: Palestras – Congresso Espirita Brasileiro
 Titulo: Doutrina Espirita

Clips Recursos de apoio Associações de recursos

Nome	Arquivo	Duracao	Palavras Chave
Jovem1	anateixeira/clips/clip5...	6	equivocos orientacao f...
Jovem2	anateixeira/clips/Clip...	0	energias
Jovem3	anateixeira/clips/Clip...	0	Alan eu hora parceiro...
Jovem4	anateixeira/clips/Clip...	0	cigarro filho a meditac...
Jovem5	anateixeira/clips/Clip...	0	hipocrisia pontos situa...
Jovem6	anateixeira/clips/Clip...	0	unificacao bandeira u...
Jovem7	anateixeira/clips/Clip...	0	variantes me processo...
Jovem8	anateixeira/clips/Clip...	0	sociedade Agostinho...
Jovem9	anateixeira/clips/Clip...	0	vídeos comprometime...
Jovem10	anateixeira/clips/Clip...	0	x
Emmanuel1	anateixeira/clips/Clip...	0	obra psicografica
Emmanuel2	anateixeira/clips/Clip...	0	C7ndido mEdium Emm...
Emmanuel3	anateixeira/clips/Clip...	0	anos lances poesias E...
Emmanuel4	anateixeira/clips/Clip...	0	ele periferia se Minas...
Emmanuel5	anateixeira/clips/Clip...	0	mediunidade medium...
Emmanuel6	anateixeira/clips/Clip...	0	Chico compromisso p...
Emmanuel7	anateixeira/clips/anat...	0	disciplina disposia?o s...
Emmanuel8	anateixeira/clips/anat...	0	aspas tentativas culto...
Emmanuel37	anateixeira/clips/anat...	0	tramas sÉrie condia?e...
Emmanuel9	anateixeira/clips/anat...	0	Emmanuel Chico coisa...

Nome: Emmanuel6
 Arquivo: emmanuel/clips/Clip24.mov
 Duracao(s): 0
 Palavras Chave: Pedro Emmanuel episódio Para coisa de Kardec mg Xavier ele lugar disposição Alan Leopoldo o disciplina evangelho Jesus mediunidade medium dia trabalho
 Texto do clip: ps/ftfiles/emmanuel/Clip_6.txt

Novo Clip Apaga clip

Fonte: Dados da pesquisa

A figura 9 representa a área administrativa que permite associar as palavras-chave (ou a indexação multimodal completa) a seu respectivo arquivo de vídeo. Além disso, proporciona a opção de inserir uma ontologia para expandir a busca, utilizando o motor de inferência *Jena*. Finalmente, o protótipo utiliza cada configuração para realizar a busca e apresenta como resposta para o usuário uma listagem dos vídeos que possuem maior relevância com a consulta. A visão geral de todo o processo poder ser vista na figura 10. A utilização do protótipo para avaliação da pesquisa e as quatro diferentes configurações são descritas na próxima seção.

Figura 10 – Representação generalizada do método utilizado.



Fonte: Dados da pesquisa

3.5 Coleta de dados

A avaliação dos Sistemas de Recuperação da Informação é fundamental para o estabelecimento de um ponto de comparação entre os sistemas já existentes e os que estão sendo desenvolvidos. A intenção é medir a capacidade do sistema para satisfazer a necessidade de informação do usuário em termos da relevância dos documentos recuperados.

O usuário, a partir de suas necessidades de informação, dá início ao processo de busca e recuperação da informação. Além disso, a qualidade da estratégia de busca e o vocabulário são fatores importantes para a atividade (Lancaster, 1998). De acordo com Araújo Júnior (2007), os resultados só poderão ser validados por intermédio da avaliação dos usuários. Isto significa dizer que os sistemas de recuperação da informação, além de buscar atender às demandas informacionais

dos usuários, dependem destes para que a qualidade dos seus serviços seja reconhecida.

Nesta pesquisa, utilizamos 40 sujeitos, que possuem conhecimento sobre o domínio, e solicitamos que fizessem uma consulta por qualquer assunto interessante ao usuário dentro o domínio estudado. De todos os documentos recuperados pelo sistema, o entrevistado definiu o grau de relevância para o documento conforme uma escala pré-definida.

Tabela 1 – Configurações de execução do experimento

Nome das Configurações	Indexação Multimodal	Analizador Sintático PALAVRAS	Ontologia do domínio
Completa	SIM	NÃO	NÃO
Palavras	SIM	SIM	NÃO
Palavras_Onto	SIM	SIM	SIM
Completa_Onto	SIM	NÃO	SIM

Fonte: Dados da pesquisa

O experimento contou com 40 sujeitos, que realizaram 1 (uma) consulta cada. Essa consulta foi repetida igualmente nas diferentes configurações. Os participantes do experimento são alunos dos cursos de evangelização da Federação Espírita Brasileira. Os fatores que levaram a essa escolha foram: a facilidade e necessidade de utilização dos equipamentos de informática e a disponibilidade dos alunos para a realização da tarefa. A média de idade dos sujeitos é de 38 anos. Estes colaboradores realizaram as tarefas individualmente, nas diferentes configurações propostas nessa pesquisa.

3.6 Tratamento dos dados

Os dados foram tabulados com contagem simples de relevância e cálculo de índice de precisão de cada documento e contagem cruzada com dois ou mais conjuntos de consultas. Para a análise estatística dos dados, foi utilizado o Broffice.org calc (2011), o qual permitiu a montagem dos gráficos comparativos que

ilustraram e apoiaram a análise dos dados. Para cada consulta analisada, foi determinado o índice de precisão do sistema, apresentados em gráficos e/ou tabelas. Foram realizadas consultas à coleção, utilizando palavras-chave desejadas pelo usuário, e avaliação dos cinco primeiros documentos retornados pela busca das quatro configurações, de acordo com a relevância do documento para a consulta. Foram utilizados os níveis: Muito Relevante, Relevante, Satisfatório, Irrelevante e Muito Irrelevante. Para viabilizar esta comparação, utilizamos como métrica avaliativa o índice de precisão, uma medida objetiva e consagrada desde a sua proposta por Cleverdon (1962). Para tanto, a fórmula usada foi a seguinte:

• **Precisão** (*precision*): É a razão entre o número de documentos recuperados pelo sistema que são relevantes para a consulta e o número total de documentos recuperados.

$$P = \frac{\text{Número de documentos relevantes recuperados}}{\text{Número total de documentos recuperados}}$$

Por exemplo, se o sistema recuperou 4 documentos para uma consulta, em que 2 deles foram relevantes, a medida de precisão do sistema é de 0,5 ou 50%.

Este cálculo do índice de precisão foi efetuado para cada um dos cinco primeiros documentos retornados pela busca, de acordo com a relevância determinada pelo usuário. Ao final, foi calculado o índice de precisão médio da consulta.

As 40 consultas foram executadas sob todas as configurações, e foi medida a variável de Precisão. Foi realizada avaliação dos cinco primeiros documentos retornados na busca, de acordo com a relevância do documento para a consulta. Os resultados são apresentados a seguir.

4 RESULTADOS

Esse experimento contou com 6 horas de vídeos extraídos da coleção do Conselho Espírita Internacional. Os vídeos são palestras ministradas por especialistas de um domínio. Resultaram, após a segmentação da coleção de vídeos, 147 (cento e quarenta e sete) trechos de vídeos.

Tabela 2 – Resultado da Precisão nas diferentes configurações realizadas no experimento

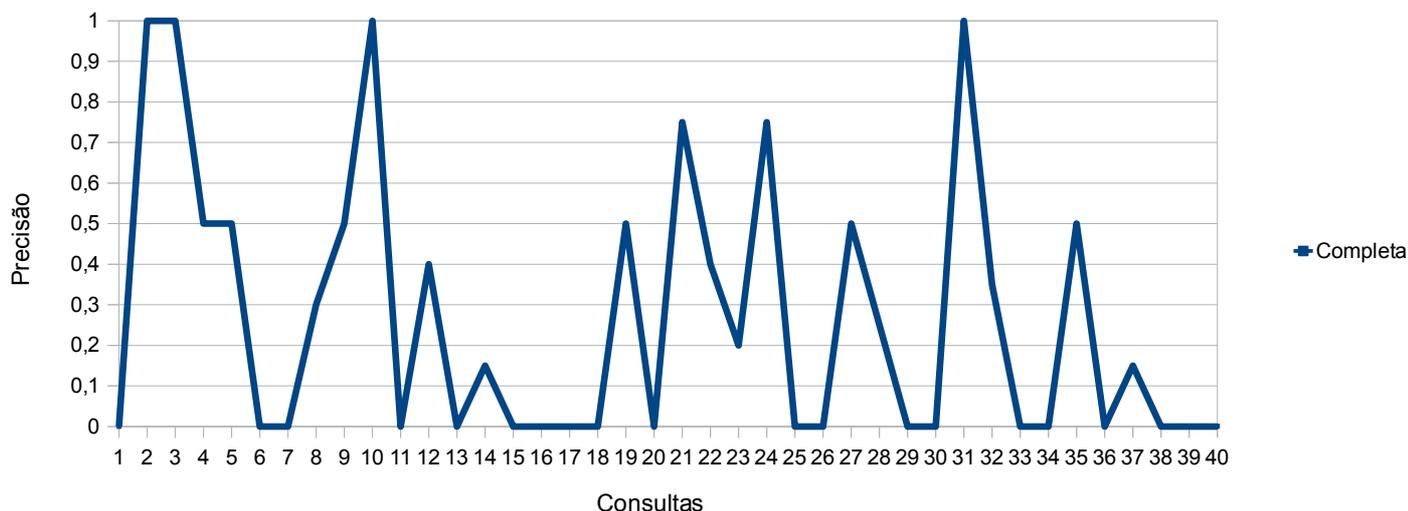
Consulta	Completa	Palavras	Completa_Onto	Palavras_Onto
1	0	0	0	0
2	1	1	1	1
3	1	1	1	1
4	0,5	1	0	1
5	0,5	0,5	0,5	0,5
6	0	0,35	0,35	0,35
7	0	0	0	0
8	0,3	0,3	0,3	0,65
9	0,5	0	0	0
10	1	0	0	0
11	0	0	0	0
12	0,4	0,4	0,4	0,8
13	0	0	0	0
14	0,15	0	0,15	0,85
15	0	0	0	0
16	0	0	0	0
17	0	0,2	0,4	0,2
18	0	0	0	0
19	0,5	0,5	0,5	0,5
20	0	0	0	0
21	0,75	0,75	0,75	0,75
22	0,4	0,4	0	0

23	0,2	0,2	0,2	0,2
24	0,75	0,75	1	0,75
25	0	0,45	0,45	0,85
26	0	0	0	0
27	0,5	0,5	1	1
28	0,25	0,75	0,75	0,75
29	0	0	0,2	0,2
30	0	0	0,5	0,5
31	1	1	1	1
32	0,35	0,7	0,7	0,7
33	0	0,35	0,35	0
34	0	0,35	0,35	0
35	0,5	0,5	0,5	0,5
36	0	0	0	0
37	0,15	0,15	0,15	0,15
38	0	0	0	0
39	0	0	0	0
40	0	0,15	0,55	0,35

Fonte: Dados da Pesquisa

Na tabela 2 apresentou-se os resultados obtidos para as consultas utilizando as configurações: Completa, Palavras, Completa_Onto e Palavras_Onto. O uso de ontologia, que serve para expandir a busca semântica e diminuir ambiguidades, otimiza a precisão do Sistema de Recuperação de Informação em relação ao sistema sem ontologia, que efetua uma busca baseada apenas em termos e palavras-chave utilizados na consulta.

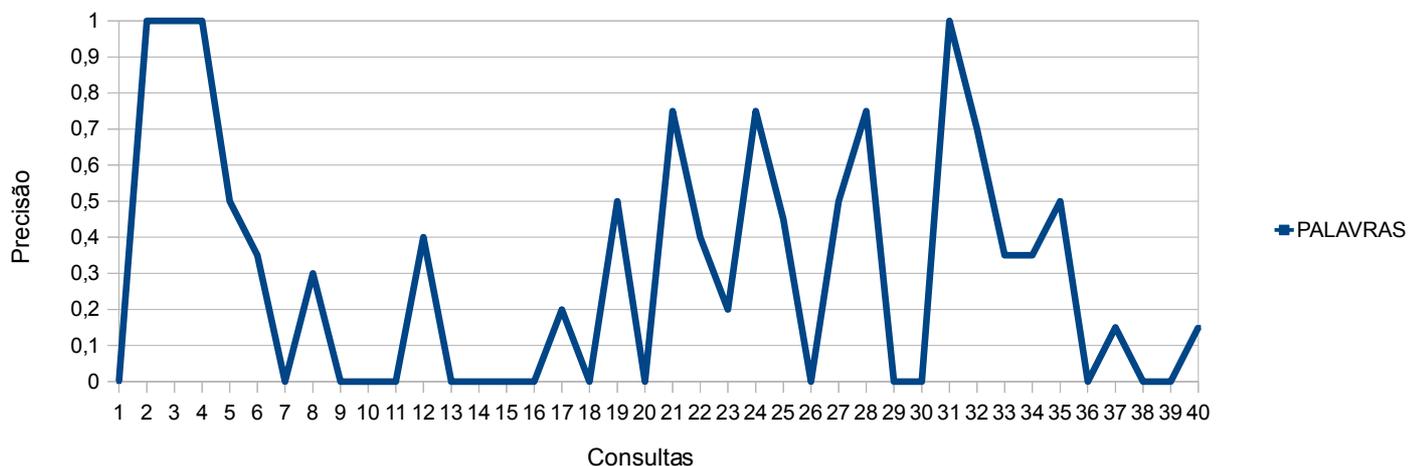
Gráfico 1 – Resultado da Precisão utilizando a configuração Completa



Fonte: Dados da Pesquisa

No Gráfico 1 tem-se os resultados de Precisão obtidos nas consultas com a configuração Completa. Como citado anteriormente, para a Recuperação da Informação, Precisão é o número total de documentos relevantes retornados dividido pelo número total de documentos retornados. A configuração utilizada nestas consultas é a chamada “Completa”, pois utiliza-se da multimodalidade para indexar os documentos a partir de todos os termos contidos no áudio e na imagem. Esta configuração não despreza termos muito frequentes nos textos e considerados vazios de conteúdo semântico, sem valor para indexação, como artigos por exemplo. Além disso, essa configuração não utiliza uma ontologia para expansão da busca. A configuração Completa apresentou precisão média de 0,27 para as consultas realizadas no experimento.

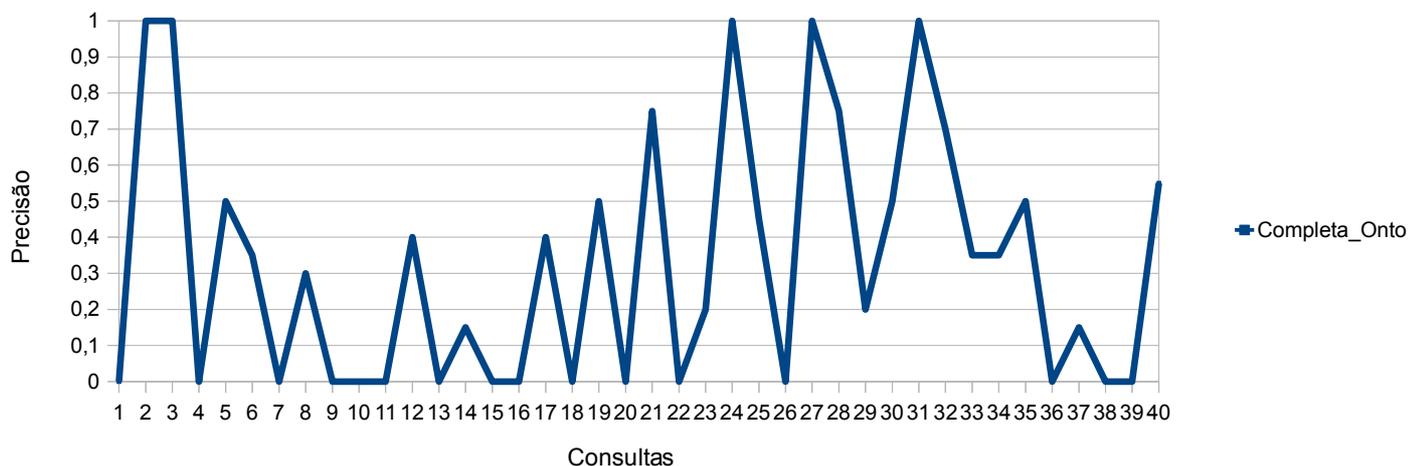
Gráfico 2 – Resultado da Precisão utilizando a configuração Palavras



Fonte: Dados da Pesquisa

No Gráfico 2 apresentam-se os resultados de Precisão obtidos nas consultas realizadas com a configuração Palavras. Esta configuração utiliza o analisador sintático PALAVRAS nas transcrições multimodais dos documentos da coleção indexada. Essa configuração apresentou precisão média de 0,31 para as consultas realizadas no experimento.

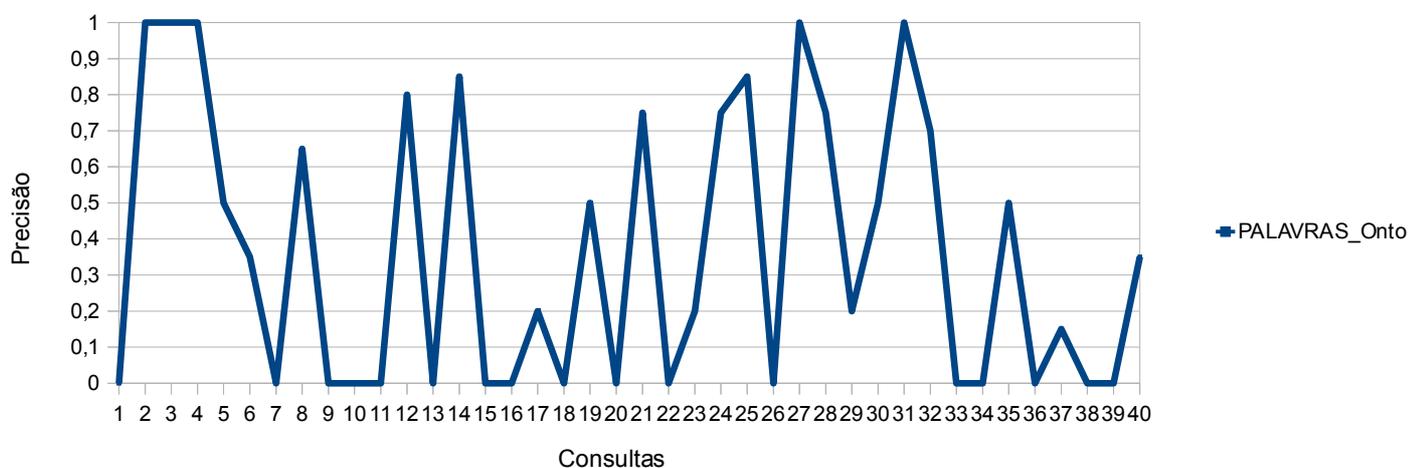
Gráfico 3 – Resultado da Precisão utilizando a configuração Completa_Onto



Fonte: Dados da Pesquisa

No gráfico 3 apresentam-se os resultados de Precisão obtidos nas consultas realizadas com a configuração Completa_Onto. A configuração Completa_Onto possui as características da configuração Completa associada a uma ontologia. A configuração Completa_Onto apresentou precisão média de 0,33 para as consultas realizadas no experimento.

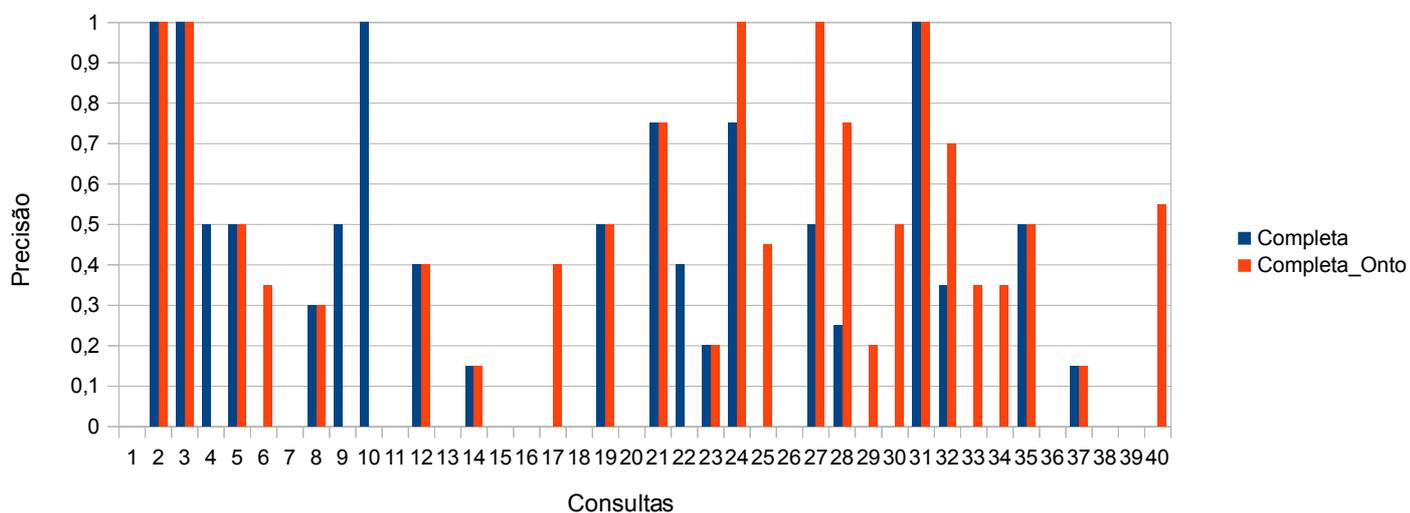
Gráfico 4 – Resultado da Precisão utilizando a configuração Palavras_Onto



Fonte: Dados da Pesquisa

No gráfico 4 apresentam-se os resultados de Precisão obtidos nas consultas realizadas com a configuração Palavras_Onto. A configuração Palavras_Onto possui as características da configuração Palavras associada a uma ontologia. A configuração Palavras_Onto apresentou precisão média de 0,36 para as consultas realizadas no experimento.

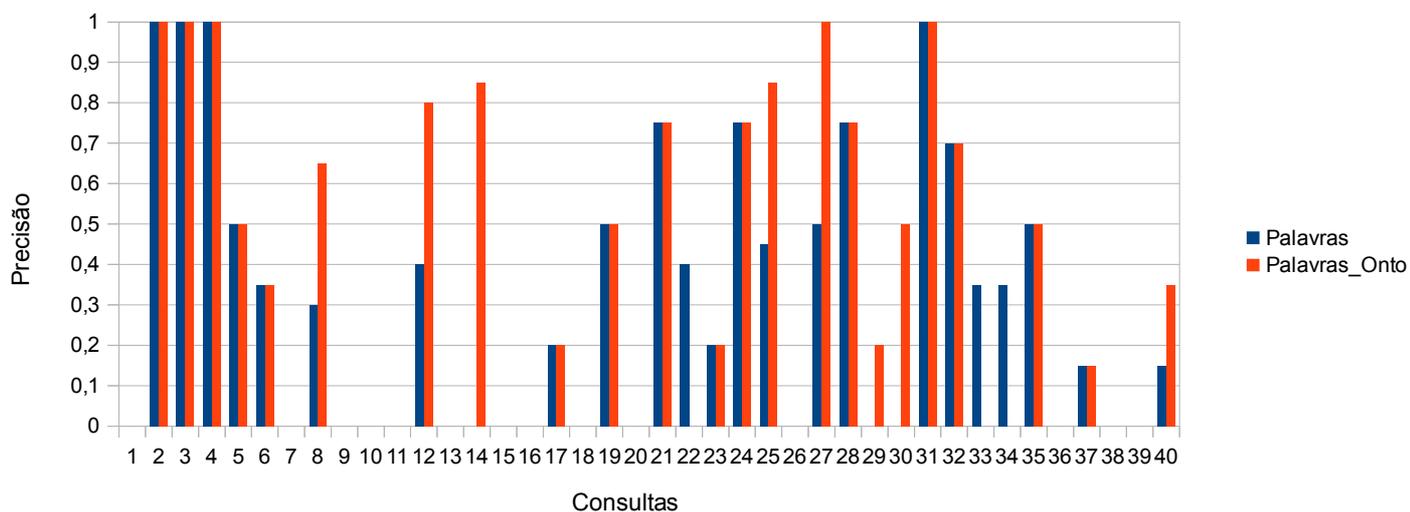
Gráfico 5 – Resultado de Precisão do experimento comparando as configurações Completa e Completa_Onto



Fonte: Dados da Pesquisa

No gráfico 5 observam-se os resultados de Precisão obtidos tanto para a configuração Completa quanto para a configuração Completa_Onto. Os resultados de Precisão apresentados pelo Completa_Onto foram em média 21,96 % (vinte e um, noventa e seis por cento) superiores aos resultados apresentados pela configuração Completa nas consultas realizadas no experimento.

Gráfico 6 – Resultado de Precisão do experimento comparando as configurações Palavras e Palavras_Onto



Fonte: Dados da Pesquisa

No Gráfico 6 observam-se os resultados de Precisão obtidos tanto para a configuração Palavras quanto para a configuração Palavras_Onto. Os resultados de Precisão apresentados pela configuração Palavras_Onto foram em média 18,78 % (dezoito, setenta e oito por cento) superiores aos resultados apresentados pela configuração Palavras nas consultas realizadas no experimento.

5 DISCUSSÃO, CONCLUSÃO E CONSIDERAÇÕES FINAIS

5.1 DISCUSSÃO

Este trabalho desenvolveu-se com base na abordagem de que a Indexação Multimodal (Snoek e Worring, 2005; O'Halloran, 2004) e a Ontologia (Corazzon, 2003) podem oferecer grandes contribuições para as áreas de disseminação e recuperação de informação, especialmente para a análise automática de vídeos. Este estudo descreveu a estruturação de um Sistema de Recuperação de Informação aplicado a uma coleção de vídeos em língua portuguesa que versam sobre Doutrina Espírita. Para o auxílio na busca utilizou-se de conceitos extraídos dos vídeos da própria coleção, por meio de indexação de áudio e imagem, e que foram associados a uma ontologia leve desenvolvida sobre o domínio contido nos vídeos.

Com base na proposta da análise do uso de ontologia para recuperar informação disponibilizada em vídeos por meio de indexação multimodal, a investigação concentrou-se na avaliação da resposta obtida no processo de busca e recuperação da informação, por meio de uma medida objetiva, o índice de precisão. Desde a década de 70, a questão da precisão foi amplamente discutida em associação com a análise de desempenho de um sistema de recuperação da informação. Para Lancaster e Fayen (1973), ao considerar os fatores que interferem no desempenho destes sistemas, será necessário conhecer anteriormente os pré-requisitos do usuário em relação aos resultados de busca e recuperação da informação.

A escolha do índice de precisão permitiu avaliar o desempenho de um protótipo com indexação multimodal, confeccionado para ser utilizado em diferentes configurações, na qual uma ontologia pode ser associada. O resultado da avaliação dos desempenhos do protótipo atingiu um dos objetivos do estudo, qual seja, avaliar se a recuperação da informação disponibilizada em vídeos por meio do uso da ontologia traz ganho no índice de precisão, se comparada à outras configurações utilizadas sem ontologia.

A precisão média obtida pela configuração Completa no experimento foi de

0,27, enquanto que a precisão média obtida pela configuração Completa_Onto foi de 0,33. Existem 4 consultas em que a configuração Completa sem o uso da ontologia obteve maior precisão que a configuração Completa_Onto. Nesse caso específico, o uso da expansão semântica apresentou documentos irrelevantes devido ao sentido da consulta estar diferente do determinado nas relações da ontologia.

A precisão média obtida pela configuração Palavras no experimento foi de 0,31, enquanto que a precisão média obtida pela configuração Palavras_Onto foi de 0,36. Houve um ganho de 0,05 a favor da configuração Palavras_Onto. Na maioria das vezes, a consulta foi expandida por meio da ontologia com termos que permitiram recuperar com alta similaridade os documentos relevantes.

Em termos de Precisão os valores obtidos são baixos para um Sistema de Recuperação de Informação. Isso é devido fundamentalmente ao fato do tamanho da coleção ser bastante reduzido. O conjunto de vídeos utilizado no teste ainda é pequeno. Isso faz com que um único documento irrelevante recuperado ou, um único documento relevante não recuperado, afete consideravelmente a precisão do sistema.

Uma medida de precisão 0, indica que a configuração não recuperou documentos ou que houve documentos recuperados mas nenhum foi relevante. Como a coleção é pequena, alguns termos das consultas apresentaram apenas um documento relevante. Isso justifica, em parte, a presença dos valores “0” e “1” encontrados nas respostas. As configurações com o uso de ontologia, na maioria das consultas que retornaram respostas, obtiveram precisão maior ou igual às demais configurações sem o uso da mesma. A realização de experimentos como este, utilizando sujeitos nativos falantes de Português do Brasil e coleções em Português do Brasil, contribuem para a definição e validação de critérios avaliativos adotados na Recuperação de Informação de documentos nesse idioma (Santos, 2004).

5.2 CONCLUSÃO

A utilização de uma ontologia de domínio da Doutrina Espírita alinhada a indexação multimodal, em um sistema de recuperação de trechos de vídeos de

palestras sobre esse tópico, permite melhorar a eficiência do mesmo em termos de precisão.

Este trabalho consistiu na elaboração de um sistema de recuperação de informação (RI) que utiliza ontologias, linguística computacional e indexação multimodal. O uso de ontologias para recuperação da informação disponibilizada em vídeos mostrou-se promissor. Alguns dos termos utilizados na versão da ontologia descrita são dependentes do contexto no qual será feito uso das informações, e isto deve ser melhor avaliado em futuras versões da ontologia.

O tamanho da coleção é um fator importante no desenvolvimento e, principalmente, na avaliação de sistemas de RI. A variedade de consultas que podem ser feitas é muito grande, e o sistema termina, em muitos casos, trazendo respostas não muito relevantes, pois não conta com respostas adequadas para a consulta do usuário. Além disso, a avaliação é muito sensível, sendo que um único documento irrelevante recuperado ou um único documento relevante não recuperado, faz com que a precisão seja fortemente afetada.

5.3 CONSIDERAÇÕES FINAIS

O presente trabalho descreveu a arquitetura de um sistema de recuperação de informação disponibilizada em vídeos, testado em palestras do domínio da Doutrina Espírita.

A arquitetura foi projetada para possibilitar a melhor indexação de palestras em vídeos. As contribuições acadêmicas expostas neste trabalho poderiam auxiliar a área de Educação a Distância (*eLearning*) para permitir o melhor aproveitamento de cursos e seminários, os quais geralmente tem longa duração.

Foi comprovado que a precisão média do sistema foi melhorada, para o protótipo, quando foi utilizado um mecanismo de expansão de consulta que utiliza uma ontologia do domínio.

Como parte do trabalho foi desenvolvida uma ontologia do domínio da Doutrina Espírita, contendo principalmente os termos tratados na obra *Doutrina Espírita para Principiantes*.

5.4 TRABALHOS FUTUROS

Como trabalhos futuros sugerem-se temas para o desenvolvimento de pesquisas nas áreas de Recuperação de Informação, através de novos experimentos envolvendo o protótipo utilizado. Em particular trabalhos que envolvam a expansão de consultas e a multimodalidade para SRIs. Esta abordagem teórica, explicitada na pesquisa, pode contribuir para as possíveis soluções do problema de *gap* semântico, por meio da aproximação usuário-sistema.

Pode-se também utilizar essa abordagem para desenvolver pesquisas na área de geração automática de vídeos, e/ou mineração de vídeos. A indexação automática de vídeos e o desenvolvimento de material didático-pedagógico para a Educação a Distância (*eLearning*) também são áreas que podem ser auxiliadas pelas contribuições acadêmicas expostas neste trabalho.

Outra questão a ser tratada é o desenvolvimento automático de ontologias. Esta é, sob a ótica da Recuperação da Informação, um desafio típico para esta aplicação.

REFERÊNCIAS

- ABOWD, G.; ATKESON, C.; FEINSTEIN, A.; HMELO, C.; KOOPER, R.; LONG, S.; SAWHNEY, N.; TANI, M. **Teaching and learning as multimedia authoring: the classroom 2000 project.** Disponível em:<<http://www.cc.gatech.edu/fce/c2000/pubs/mm96/c2000.html>>. Acesso em 25 jan. 2010.
- ALATAN, A. A.; AKANSU, A. N.; WOLF, W. Multimodal dialogue scene detection using hidden markov models for content-based multimedia indexing. **Multimedia Tools and Applications**, [S.l], v. 14, n. 2, p. 137–151, 2001.
- ALMEIDA, Maurício; BAX, Marcello. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, v.32, n.3, p.7-20, set./dez. 2003.
- ARAÚJO JÚNIOR, Rogério Henrique de. **Precisão no processo de busca e recuperação da informação.** Tese (Doutorado em Ciência da Informação) – Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2005.
- ARAÚJO JÚNIOR, Rogério Henrique de. **Precisão no processo de busca e recuperação da informação.** Brasília: Thesaurus, 2007.
- BABAGUCHI,N.; KAWAI,Y.; KITAHASHI,T. Event based indexing of broadcasted sports video by intermodal collaboration. **IEEE Transactions on Multimedia**, Los Angeles, v 4, n. 1, p. 68–75, 2002.

BACH, J. et al. The virage image search engine: an open framework for image management. In: STORAGE AND RETRIEVAL FOR IMAGE AND VIDEO DATABASES, 4., 1996 , San Diego. **Proceedings...** San Diego, 1996. p.76-87.

BAEZA-YATES, Ricardo. An extended model for full text databases. **Journal of the Brazilian Computer Society**, Porto Alegre, v.2, n.3, 1996.

BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. New York: ACM Press, 1999. 499 p.

BALLARD, D. H.; BROWN, C. M. **Computer Vision**. New Jersey: Prentice-Hall, 1982.

BARRETO, Juliano Serra. Desafios e avanços na recuperação automática da informação audiovisual. **Ciência da Informação**, Brasília, v.36, n. 3, p.17-28, set./dez, 2007.

BERRUTO, G. **La semántica**. México: Nueva Imagen, 1979.

BERNARAS, A. Building and reusing ontologies for electrical network applications. In: EUPOREAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, 12., 1996, Budapest. **Proceedings...** Budapest, 1996. p.298-302

BICK, Eckhard. **The parsing system palavras**: automatic grammatical analysis of portuguese in a constraint grammar framework. [S.I]: Aarhus University Press, 2000. 412p.

BIDERMAN, Maria Tereza Camargo. **Teoria lingüística**: lingüística quantitativa e computacional. Rio de Janeiro: LTC, 1978.

BLIUJUTE, R. et al. Developing a datablade for a new index. In: INTERNATIONAL CONFERENCE ON DATA ENGINEERING, 15., 1999, Sydney. **Proceedings...** Sydney, 1999. p.314-323.

BORST, W. **Construction of engineering ontologies for knowledge sharing and reuse**. Ph.D. Dissertation, University of Twente, 1997.

BOSSON, A. et al. Non-retrieval: blocking pornographic images. In: INTERNATIONAL CONFERENCE ON IMAGE AND VIDEO RETRIEVAL, 2002, London. **Proceedings...** London, 2002. p.50-60.

BRASCHER, M. A ambigüidade na recuperação da informação. **DataGramZero**, Rio de Janeiro, v.3, n.1, fev. 2002.

BRÉAL, Michel. **Ensaio de Semântica**. São Paulo: Pontes/Educ, 1992.

CARROLL, J.J. et al. Jena: implementing the semantic web recommendations. **HP Labs Technical Reports** , 2003.

CHAUDHRI, V. K. et al. **Open knowledge base connectivity 2.0.3**: proposed. [S. l. : s. n.], 1998. Disponível em: <<http://www.ai.sri.com/~okbc/okbc-2-0-3.pdf>>. Acesso em jan. 2011.

CHOMSKY, N. **Syntactic structures**. 2.ed. [S.I]: Walter de Gruyter, 2002.

CLEVERDON, C. W. **Report on testing and analysis of investigation into comparative efficiency of indexing systems**. Cranfield: Aslib, 1962.

CORAZZON, R. **Descriptive and Formal Ontology**.
<http://www.formalontology.it/index.htm>. 2003.

CORCHO, O; GOMEZ-PEREZ, A. A roadmap to ontology specification languages. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT METHODS, MODELS AND TOOLS, 12., 2000, France. **Proceedings**. France, 2000. p.80-97.

DESHPANDE, S.G.; HWANG, J.N. A real-time interactive virtual classroom multimedia distance learning system. **Trans on Multimedia**, [S.I], v. 3, n. 4, p. 432-444, 2001.

DING, Z. Global output regulation of uncertain nonlinear systems with exogenous signals. **Automatica**, [S.I], v.37, n.1, p.113-119, 2001.

DUQUE, C. G. **SIRILICO**: Uma proposta para um sistema de recuperação de informação baseado em teorias da lingüística computacional e ontologia. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

DUQUE, C. G. **Uma abordagem ontológica para a indexação de documentos eletrônicos**. Brasília, 2006. Disponível em:
<<http://www.asocarchi.cl/DOCS/53.PDF>>. Acesso em 07 abr. 2009.

EAKINS, J. P. ; RILEY, K. J. ; EDWARDS, J. D. Shape Feature Matching for Trademark Image Retrieval. In: INTERNATIONAL CONFERENCE ON IMAGE AND VIDEO RETRIEVAL, 2., 2003, Urbana. **Proceedings...** Urbana, 2003.

EGAS, R. et al. Adapting k-d trees to visual retrieval. In: INTERNATIONAL CONFERENCE ON VISUAL INFORMATION SYSTEMS, 3., 1999, Amsterdam. **Proceedings...** Amsterdam, 1999..

FEITOSA, Ailton. **A integração entre sistemas legislativos, terminologia e web semântica na organização e representação da informação legislativa.** Tese (Doutorado em Ciência da Informação) – Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2005.

FLICKNER, M. et al. Query by image and video content: the QBIC system, **IEEE Computer**, [S.I.], v.28, p.23-32, 1995.

FONSECA, F. ; EGENHOFER, M. Sistemas de informação geográficos baseados em ontologias. **Informática Pública**, [S.I.], v.1, n.2, p. 47-65, 1999.

FOOTE, J. An Overview of audio information retrieval. **ACM Multimedia Systems**, [S.I.], v.7, p.2-10, 1999.

FORSYTH, D. A. ; FLECK, M. M. Automatic detection of human nudes, **International Journal of Computer Vision**, [S.I.], v.32, p.63-77, 1999.

FRANKEL, C. ; SWAIN, M.J. ; ATHITSOS, V. WebSeer: an image search engine for the world wide web. **World Wide Web Internet And Web Information Systems**, [S.l], v. 96, p.1-24, 1996.

FUCHS, C. **Ambigüité, paraphrase et langage en acte. Moléles Linguistiques**, n. 19, 1988.

FUJITA, Mariângela. A identificação de conceitos no processo de análise de assunto para indexação. **Revista digital de Biblioteconomia e Ciência da Informação**, Campinas, v.1,n.1, p.69-90, jul./dez. 2003.

GAGNON, Michel. Processamento da Linguagem Natural, 2000. In.: <http://www.professeurs.polymtl.ca/michel.gagnon/Disciplinas/Bac/IA/PLN/pln.html>. Acesso em 21 de outubro de 2010.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, London, v. 5, p. 199-220, 1993.

GRUBER. T. R. Towards principles for the design of ontologies used for knowledge sharing. **International Journal of Human-Computer Studies**, [S.l], v. 43, p.907-928, 1995.

GRUBER, T. **What is an Ontology?** Disponível em:< <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>> 1996. Acesso em agosto de 2010.

GRUNINGER, M.; FOX, M. S. The logic of enterprise modelling. In: Industrial engineering research conference, 1996, Nashville. **Proceedings...** Nashville, 1996, p.83-98.

Disponível em: <<http://citeseer.ist.psu.edu/523974.html>>. Acesso em 10 nov. 2010.

GUARINO, N. & GIARETTA, P. Ontologies and knowledge bases: Towards a terminological clarification. In.: Mars, N., **Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing**, 1995, Amsterdam, NL. IOS Press.

GUIZZARDI, Giancarlo. **Desenvolvimento para e com reuso: um estudo de caso no domínio de vídeo sob demanda**. Dissertação (Mestrado em Informática) – Centro Tecnológico, Universidade Federal do Espírito Santo, Vitória, 2000.

HAAS, M. ; LEW, M. S. ; HUIJSMANS, D. P. ; A new method for key frame based video content representation. In.: **Image Databases and Multimedia Search**, 1997.

HANJALIC, A.; LAGENDIJK, R. L.; BIEMOND, J. A new method for key frame based video content representation. **Image Databases and Multimedia Search**, p. 97-107, 1997.

HARALICK, R. M. ; SHAPIRO, L.G. **Computer and robot vision**. New York: Addison-Wesley, 1993.

HE, L., SANOCKI, E., GUPTA, A., GRUDIN, J. Auto-summarization of audio-video presentation. **ACM Multimedia**, p. 489-498, 1999. Disponível em: <<http://www.edocfind.com/download/ppt/Auto-Summarization%20of%20Audio-Video%20Presentations/aHR0cDovL3Jlc2VhcmNoLm1pY3Jvc29mdC5jb20vZW4tdXMvdW0vcGVvcGxIL2xoZS90YWxrcy9hdXRvc3VtLW1tOTkucHB0>>. Acesso em: set. de 2010.

HORRIDGE, M. et al. A practical guide to building OWL ontologies using the protégé-OWL plugin and CO-ODE tools. **The University of Manchester**, v. 27, p. 0-117, 2004. Disponível em: <<http://www.co-ode.org/resources/tutorials/ProtegeOWLTutorial.pdf>>. Acesso em 07 fev. 2011

HUANG,J. et al. Integration of multimodal features for video scene classification based on HMM. **Multimedia Signal Processing**, Copenhagen, n.3, p.53-58, 1999.

IEEE, IEEE standard glossary of software engineering terminology," IEEE Std 610.12 1990. Disponível em:<<http://ieeexplore.ieee.org/stamp/stamp.jsp>>. Acesso em 04 ago. 2011.

LANCASTER, F.W. **Indexação e resumos: teoria e prática**. 2.ed. Brasília: Briquet de Lemos, 2004.

LANCASTER, F. W.; **Indexing and abstracting in theory and practice**. 2nd ed. London: Library Association, 1998.

LANCASTER, F.W.; FAYEN, E. G. **Information retrieval on-line**. Los Angeles: Melville, 1973.

LEE,A. E KAWAHARA,T. **Recent development of open-source speech recognition engine julius**. Asia-Pacific and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2009.

LENAT, D. B. **Cycorp, inc**. 1995. Disponível em <<http://www.cyc.com/>> Acesso em 24 jan. 2001.

LENAT, D. B. ; GUHA, R. V. **Building large knowledge-based systems:** representation and inference in the Cyc project. Boston: Addison-Wesley, 1990.

LEVINE, M. **Vision in man and machine.** [S.I]: Mcgraw Hill, 1985.

LEW, M. S. ; HUIJSMANS, N. Information theory and face detection. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 13., 1996, Vienna. **Proceedings...** Vienna, 1996, p.601-605.

LEW, M.S. Next generation web searches for visual content. **Computer**, [S.I], v.33, n. 11, p. 46-53, 2000.

LIENHART, R. Reliable transition detection in videos: a survey and practitioner's guide. **International Journal of Image and Graphics**, [S.I], v.1, n.3, p. 469-486, 2001.

LUGER, G.F. **Artificial intelligence:** structures and strategies for complex problem solving 5.ed. London: Addison-Wesley, 2005.

LYONS, John. Linguagem e Cultura. In:_____. **Linguagem e lingüística:** uma introdução. Rio de Janeiro: Ed. Guanabara, 1987.

MAHMOOD, T. F. S. Indexing for topics in videos using foils. **Computer Vision and Pattern Recognition**, v. 2, p. 312-319, 2000.

McGUINNESS, D. L.; HARMELEN, F. V. **OWL web ontology language overview: W3C Candidate Recommendation**. Disponível em : <<http://www.w3.org/TR/2003/CR-owl-features-20030818/>>. Acesso em 05 fev. 2011.

MELLO, C. ; LINS, R. D. Segmentação de imagens de documentos históricos. SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES, 17., 1999, Vila Velha. **Anais...** Vila Velha, 1999.

MIZOGUCHI, R.; VANWELKENHUYSEN, J. ; IKEDA, M. Task ontology for reuse of problem solving knowledge. **Towards Very Large Knowledge Bases**, [S.l], p. 46-57, 1995.

MOORE, Michel G., KEARSLEY, Greg. **Distance education: a systems view**. Belmont(USA) : Wadsworth Publishing Company, 1996.

MORAN, José Manuel. **Ensino e aprendizagem inovadores com tecnologias**. Revista: Informática na Educação: Teoria & Prática. Porto Alegre, vol. 3, n.1. UFRGS. Programa de Pós-Graduação em Informática na Educação, set. 2000, pág 137-144.

NECHES, R. Enabling technology for knowledge sharing. **AI Magazine**, [S.l], v.12, n.3, p.36-56, 1991.

NEGROPONTE, N. **A Vida Digital**. 2.ed. São Paulo: Companhia das Letras, 1995.

NOY, N. F. et al. Creating semantic web contents with Protege-2000. **IEEE Intelligent Systems**, [S.l], v.16, n.2, p. 60-71, 2001.

NOY, F. N.; GUINNESS, D. L. **Ontology development 101**: a guide to create your first ontology. Disponível em: <<http://ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.doc>>. Acesso em 03 abr. 2011.

NOVELLO, T.C. **Ontologias**: sistema baseados em conhecimento e modelos de banco de dados. Seminário apresentado na disciplina de Tópicos Avançados em Banco de Dados. Universidade Federal do Rio Grande do Sul, 2002.

O'HALLORAN, Kay L. **Multimodal discourse analysis**: systemic functional perspectives. London: Continuum, 2006.

ONAIR – **Ontology Aided Information Retrieval**. Disponível em: <<http://www.ime.usp.br/~rmcobe/onair/>>. Acesso em 5 set. 2010.

PFEIFFER, S. ; LIENHART, R. ; EFFELSBURG, W. Scene determination based on video and audio features. **Multimedia Tools and Applications**, [S.l.], v. 15, n. 1, p. 59-81, 2001.

REED, S. L.; LENAT, D. B. **Mapping ontologies into cyc**. 2002. Disponível em: <http://www.cyc.com/doc/white_papers/mapping-ontologies-into-cyc_v31.pdf>. Acesso em 20 abr. 2011.

RIBEIRO-NETO, Berthier ; BAEZA-YATES, Ricardo. **Modern information retrieval**. New York: ACM Press Book, 1999.

RICH, E., & KNIGHT, K. **Inteligência Artificial** (2 ed.). (M. C. Santos Ribeiro Ratto, Trad.) São Paulo: Makron Books, 1993.

RIVAS, Luis Hu (org.). **Doutrina espírita para principiantes**: introdução ao estudo da doutrina que ilumina consciências e consola corações. Brasília: Conselho Espírita Internacional, 2007.

ROBREDO, J. **Documentação de hoje e de amanhã**. 4. ed. Brasília: Edição de autor, 2005.

ROWLEY, H.; BALUJA, S. ; KANADE, K. Human face detection in visual scenes. NEURAL INFORMATION PROCESSING SYSTEMS, 8., 1996, Denver. **Proceedings...** Denver, 1996.

RUSSELL, S. ; NORVING, P. **Artificial intelligence**: a modern approach. 2. ed. New Jersey: Prentice-Hall, 2003.

SAG, I.A. ; WASOW, T. **Syntactic theory**: a formal introduction. Stanford: CSLI Publications, 1999.

SANTOS, D. Processamento de linguagem natural: apresentação através das aplicações. In: Elisabete Ranchhod (ed.). **Tratamento das línguas por computador**: uma introdução à linguística computacional e suas aplicações. Lisboa: Caminho, 2001. p.229-259.

SANTOS, D. **Working with Portuguese corpora**. Presentation at the University of Oslo, 22 October 2004.

SANTOS, E. T.; BARROS, L. N.; VALENTE, V. C. P. Projetando uma ontologia de geometria descritiva. SIMPÓSIO NACIONAL DE GEOMETRIA DESCRITIVA E DESENHO TÉCNICO,15., 2001, São Paulo. **Anais...** São Paulo: USP, 2001. Disponível em: <<http://www.pcc.usp.br/toledo/pdf/ontologiafinal.pdf>>. Acesso em 18 fev. 2011. p.918-928.

SCHULTE,O.; WUNDEN,T.; BRUNNER,A. **REPLAY – An integrated lecture recording infrastructure to interactively and collectively generate learning objects.** Disponível em <<http://eunis.dk/papers/p13.pdf>> Acesso em junho de 2009.

SILVA, A. S. Homonímia e Polissemia: análise sêmica e teoria do campo léxico. In: CONGRESSO INTERNACIONAL DE FILOGIA E LINGÜÍSTICA ROMÂNICA, 19., 1989, Santiago de Compostela. **Anais...** Santiago de Compostela, 1989.

SILVA, ÊNIO. **Desenvolvimento de um Reconhecedor Automático de Voz com Suporte a Grandes Vocabulários para o Português Brasileiro**, Tech. Rep., 2005.

SILVA,P.; NETO,N.; KLAUTAU,A. “Novos recursos e utilização de adaptação de locutor no desenvolvimento de um sistema de reconhecimento de voz para o Português Brasileiro” In XXVII Simpósio Brasileiro de Telecomunicações, 2009.

SMEULDERS, A. Content based image retrieval at the end of the early years. **Pattern analysis and machine intelligence**, [S.i], v.22, n.12, p. 1349-1380, 2000.

SMITH, J. R. ; CHANG, S.F. Visually searching the web for content. **Multimedia**, [S.I], v.4, n.3, p. 12-20, 1997.

SNOEK, C.G.M.; WORRING, M. Multimodal video indexing: a review of the state-of-the-art. **Multimedia tools and applications**, [S.I], v. 25, n. 1, p. 5-35, 2003.

SNOEK, C.G.M. et al. The MediaMill TRECVID 2004 semantic video search engine proc. trecvid workshop, 2001.

SUREKA, A. Mining youtube to discover extremist videos, users and hidden communities. **AIRS**, v. 6458, p. 13–24, 2010.

SWARTOUT, B. et al. **Toward distributed use of large-scale ontologies**. 1996. Disponível em: <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/Banff_96_final_2.html>. Acesso em 10 abr. 2011.

TERUSZKIN,R., VIANNA,F. **Implementation of a large vocabulary continuous speech recognition system for Brazilian Portuguese**, Journal of Communication and Information Systems, vol. 21, pp. 204-218, 2006.

THE **Syntax of CycL**. Disponível em: <<http://www.cyc.com/cycdoc/ref/cycl-syntax.html>>. Acesso em 19 jun. 2009.

ULLMANN, S. **Semântica**: uma introdução a ciência do significado. [S.l.]: Calouste Gulbenkian, 1964.

USCHOLD, M. ; GRUNINGER, M. Ontologies: principles, methods and applications. **Knowledge Engineering Review**, v.11, n.2, 1996. Disponível em: <<http://www.l2f.inesc-id.pt/~joana/prc/artigos/03c%20Ontologies%20-%20principles,%20methods%20and%20applications%20-%20Uschold,%20Gruninger%20-%201996.pdf>>. Acesso em 02 fev. 2011.

USCHOLD, M. ; KING, M. **Towards a methodology for building ontologies**. Disponível em: <<http://citeseer.ist.psu.edu/uschold95toward.html>>. Acesso em 10 nov. 2010.

VAN HEIJST, G.; SCHREIBER , A. ; WIELINGA, B. Using explicit ontologies in KBS development. **Computer Studies**, [S.l], v. 45, p. 183-292, 1997.

WEIZENBAUM, Joseph. **ELIZA** — A Computer Program For the Study of Natural Language Communication Between Man And Machine, Communications of the ACM 9 (1): 36–45. Janeiro, 1966.

WOODS, W. A. **Progress in natural language understanding**: An application to Lunar geology. AFIPS Natl. Comput. Conj: Expo.. Con/.. Proc. 42, 441-450, 1973.

WORLD WIDE WEB CONSORTIUM. **RDF Resource Definition Framework**. Disponível em: < <http://www.w3.org/RDF/>>. Acesso em 20 nov. 2010.

APÊNDICES

APÊNDICE 1 - MODELO DE INSTRUMENTO UTILIZADO NOS TESTES

**Universidade de Brasília
Faculdade de Ciência da Informação
Programa de pós-graduação em Ciência da Informação**

**USO DE ONTOLOGIA PARA RECUPERAÇÃO DA INFORMAÇÃO
DISPONIBILIZADA EM VÍDEOS POR MEIO DE INDEXAÇÃO
MULTIMODAL**

A Pesquisa tem como propósito promover, com maior precisão, a recuperação de uma informação específica existente em um repositório de vídeos por meio da indexação multimodal e uso de ontologia.

**AS INFORMAÇÕES LEVANTADAS PELA PESQUISA SÃO DE CARÁTER
ESTRITAMENTE CONFIDENCIAL, ESTANDO VEDADA A DIVULGAÇÃO OU
ACESSO AOS DADOS INDIVIDUAIS DA FONTE INFORMANTE PARA
QUALQUER EMPRESA, ÓRGÃO PÚBLICO OU PESSOA FÍSICA.**

1 – Identificação

Nome: _____

Idade: _____ anos

Sexo: () Masculino () Feminino

Inscrito no Curso de Evangelização da Federação Espírita Brasileira: ()SIM ()NÃO

2 – Dados da Consulta

Termo da Consulta: _____

Configuração 1

1º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante

2º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante

3º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante

4º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante

5º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante

Configuração 2**1º Vídeo -**

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

2º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

3º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

4º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

5º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

Configuração 3**1º Vídeo -**

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

2º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

3º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

4º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

5º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

Configuração 4

1º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

2º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

3º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

4º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

5º Vídeo -

Muito Relevante Relevante Satisfatório Pouco Relevante Irrelevante

APÊNDICE 2 – INSTRUMENTO REFERENTE A CONSULTA 8

**Universidade de Brasília
Faculdade de Ciência da Informação
Programa de pós-graduação em Ciência da Informação**

**USO DE ONTOLOGIA PARA RECUPERAÇÃO DA INFORMAÇÃO
DISPONIBILIZADA EM VÍDEOS POR MEIO DE INDEXAÇÃO
MULTIMODAL**

A Pesquisa tem como propósito promover, com maior precisão, a recuperação de uma informação específica existente em um repositório de vídeos por meio da indexação multimodal e uso de ontologia.

**AS INFORMAÇÕES LEVANTADAS PELA PESQUISA SÃO DE CARÁTER
ESTRITAMENTE CONFIDENCIAL, ESTANDO VEDADA A DIVULGAÇÃO OU
ACESSO AOS DADOS INDIVIDUAIS DA FONTE INFORMANTE PARA
QUALQUER EMPRESA, ÓRGÃO PÚBLICO OU PESSOA FÍSICA.**

**Brasília
2011**

1 – Identificação

Nome: _____

Idade: 24 anos

Sexo: () Masculino (X) Feminino

Inscrito no Curso de Evangelização da Federação Espírita Brasileira: (X) SIM () NÃO

2 – Dados da Consulta

Termo da Consulta: mediumidade

Configuração 1

1º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

2º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

3º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

4º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

5º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

Configuração 2

1º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

2º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

3º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

4º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

5º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

Configuração 3

1º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

2º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

3º Vídeo -

() Muito Relevante () Relevante (X) Satisfatório () Pouco Relevante () Irrelevante

4º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

5º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

Configuração 4**1º Vídeo -** **Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante****2º Vídeo -** **Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante****3º Vídeo -****() Muito Relevante Relevante () Satisfatório () Pouco Relevante () Irrelevante****4º Vídeo -****() Muito Relevante () Relevante Satisfatório () Pouco Relevante () Irrelevante****5º Vídeo -****() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante**

APÊNDICE 3 – INSTRUMENTO REFERENTE A CONSULTA 30

**Universidade de Brasília
Faculdade de Ciência da Informação
Programa de pós-graduação em Ciência da Informação**

**USO DE ONTOLOGIA PARA RECUPERAÇÃO DA INFORMAÇÃO
DISPONIBILIZADA EM VÍDEOS POR MEIO DE INDEXAÇÃO
MULTIMODAL**

A Pesquisa tem como propósito promover, com maior precisão, a recuperação de uma informação específica existente em um repositório de vídeos por meio da indexação multimodal e uso de ontologia.

**AS INFORMAÇÕES LEVANTADAS PELA PESQUISA SÃO DE CARÁTER
ESTRITAMENTE CONFIDENCIAL, ESTANDO VEDADA A DIVULGAÇÃO OU
ACESSO AOS DADOS INDIVIDUAIS DA FONTE INFORMANTE PARA
QUALQUER EMPRESA, ÓRGÃO PÚBLICO OU PESSOA FÍSICA.**

**Brasília
2011**

1 – Identificação

Nome: _____

Idade: 26 anosSexo: Masculino () FemininoInscrito no Curso de Evangelização da Federação Espírita Brasileira: SIM () NÃO

2 – Dados da Consulta

Termo da Consulta: Divulgar lei causa e efeito

Configuração 1

1º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante

2º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante

3º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante

4º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante

5º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante

Configuração 2

1º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

2º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

3º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

4º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

5º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

Configuração 3

1º Vídeo -

(X) Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante

2º Vídeo -

() Muito Relevante (X) Relevante () Satisfatório () Pouco Relevante () Irrelevante

3º Vídeo -

() Muito Relevante (X) Relevante () Satisfatório () Pouco Relevante () Irrelevante

4º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

5º Vídeo -

() Muito Relevante () Relevante () Satisfatório () Pouco Relevante (X) Irrelevante

Configuração 4**1º Vídeo -** **Muito Relevante () Relevante () Satisfatório () Pouco Relevante () Irrelevante****2º Vídeo -****() Muito Relevante Relevante () Satisfatório () Pouco Relevante () Irrelevante****3º Vídeo -****() Muito Relevante Relevante () Satisfatório () Pouco Relevante () Irrelevante****4º Vídeo -****() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante****5º Vídeo -****() Muito Relevante () Relevante () Satisfatório () Pouco Relevante Irrelevante**