Esta comunicação está sendo submetida ao

GT 8 – Informação e Tecnologia

RECUPERAÇÃO DA INFORMAÇÃO DISPONIBILIZADA EM VÍDEOS: UMA ABORDAGEM UTILIZANDO ONTOLOGIA, PROCESSAMENTO DE LINGUAGEM NATURAL E DESCRITORES MULTIMODAIS

Esta comunicação está sendo submetida para apresentação na modalidade:

[X] Comunicação oral [ ] Pôster

Emilio Evaristo de Sousa – UnB

Cláudio Gottschalg Duque – UnB

emilioevaristo@hotmail.com

Resumo: A necessidade de recuperar a informação disponibilizada em vídeos em grandes repositórios, têm levado a esforços para aumentar a precisão das respostas aos usuários finais. A recuperação da informação em vídeos longos apresenta alguns problemas, pois é realizada, principalmente, através das descrições textuais subjetivas determinadas para o vídeo inteiro. O objetivo da 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. Foi desenvolvida uma ontologia de domínio, associada a 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 descritores multimodais e uso de processamento de linguagem natural, em um sistema de recuperação de trechos de vídeos de palestras sobre esse tópico, permitiu melhorar a eficiência do sistema em termos de precisão.

Palavras-chave: Recuperação da Informação Audiovisual. Ontologia. Processamento de Linguagem Natural. Multimodalidade.

1 INTRODUÇÃO

Uma definição para Ciência da Informação (CI), diz que essa ciência tem como objeto a produção, seleção, organização, interpretação, armazenamento, recuperação, disseminação, transformação e uso da informação (BORKO, 1968). Apesar de ser uma definição amplamente citada e aceita na área, existe o problema de que não há um consenso sobre o significado do termo informação. Devido ao grande número de influências e de aplicações para as pesquisas em CI, às dificuldades encontradas na definição do conceito básico para o campo – a informação (CAPURRO; HJØRLAND, 2007) – e a interdisciplinaridade que caracterizam a área, mesmo temas simples podem tornar-se complexos e controversos, dificultando a sedimentação teórica.

Le Coadic (2004), considera que informação é um inscrito em forma escrita, oral ou audiovisual, em um suporte. Segundo Le Coadic (1996), com o advento da indústria da informação eletrônica o alvo da ciência da informação não é mais a biblioteca, o livro, o centro de documentação, o centro de documentação e o documento, o museu, e o objeto, mas a informação. Para Rijsbergen (1975), no contexto da Recuperação da Informação, o conceito de informação não pode ser prontamente medido. Entretanto, Rijsbergen destaca que o problema da recuperação eficiente da informação continua largamente sem solução.

No atual estágio da internet, a quantidade de informação e conteúdo disponibilizado por meio de vídeos vem crescendo a cada dia. 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*[1]. Diante de tal realidade, encontrar a informação precisa no *YouTube* é uma tarefa desafiante devido ao tamanho do repositório e à sua grande base de dados textuais (Sureka, 2010). 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.

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. Em Saracevic (1995),

---

[1] Os dados da Alexa são atualizados diariamente. Acesso em 23/05/2011.

é defendido que, na ciência da informação, deve-se considerar que informação é usada em um contexto e em decorrência de razões.

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.

Um método para melhorar a recuperação seria 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 palavras, ontologias subjacentes e a análise audiovisual.

Nesta pesquisa iremos utilizar a análise de vários descritores para auxiliar na definição dos relacionamentos entre os conceitos. A conceituação do domínio completo será formalizada em uma ontologia. As ontologias são estruturas de tal maneira que permitem um considerável ganho de qualidade quando empregadas num sistema de classificação. Elas oferecem maiores possibilidades estruturais das que são oferecidas por outros sistemas, como por exemplo, thesauri (Duque, 2006).

Avançar para além da recuperação de texto requer a análise de segmentos de áudio e vídeo. De que forma a utilização de descritores audiovisuais, conjuntamente com ontologias de domínio, poderá auxiliar na recuperação da informação disponibilizada em videos?

> 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).

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 FUNDAMENTAÇÃO TEÓRICA

Nesta seção são apresentados os conceitos teóricos sobre indexação audiovisual, multimodalidade e ontologia, estabelecendo uma base para a verificação proposta neste trabalho.

2.1 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).

2.2 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).

A linguagem é parte de um conjunto multimodal. A linguagem é amplamente considerada como o modo mais significativo de comunicação, particularmente nos contextos da aprendizagem e do ensino. Entretanto, a multimodalidade parte do pressuposto que representação e comunicação são feitas de múltiplos modos, os quais tem potencial para contribuir igualmente para o significado.

O pressuposto básico da multimodalidade é que significados são feitos, distribuídos, recebidos, interpretados e refeitos sua interpretação através de vários modos comunicativos e representativos – mesmo falados ou escritos. Norris (2004) define que todas as interações são multimodais. Ainda coloca que deve-se afastar a noção que a linguagem sempre tem o papel fundamental na interação, todavia coloca que isso ocorre de forma frequente.

Olhar, gestos e postura tendem a ser considerados um suporte para a fala; reforçando ou

modificando a fala mas não provendo comunicação por si próprio. Enquanto isso, a imagem é colocada com uma relação de suporte para a escrita. Pesquisas no campo da multimodalidade trazem dúvidas quanto à essa suposição.

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 descritores de comunicação multimodais e o auxílio de uma ontologia de domínio irá visar a diminuição de falhas na indexação com resultados favoráveis ao usuário.

2.3 Ontologia

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).

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 surgiram para a representação de ontologias que podem ser utilizadas (Corcho e Gómez-Pérez, 2000). A Web Ontology Language, OWL (McGuinness e Harmelen, 2003), é uma recomendação apresentada pela W3C como uma linguagem de ontologias para a

web. OWL aumenta a expressividade da linguagem pois permite identificar restrições de propriedade, de equivalência, e quantificadores.

Para os fins desta pesquisa, 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 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.

Russell e 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.

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. Neste trabalho, iremos fazer uso do analisador sintático PALAVRAS em algumas das diferentes configurações apresentadas.

3 METODOLOGIA

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.

Para a pesquisa, foi desenvolvido um protótipo. O protótipo foi baseado em uma solução aberta (Schulte et al., 2008) desenvolvida em linguagem *Java* 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 1 mostra a tela do usuário ao acessar um dos vídeos indexados pelo protótipo.

Figura 1 – Tela do protótipo

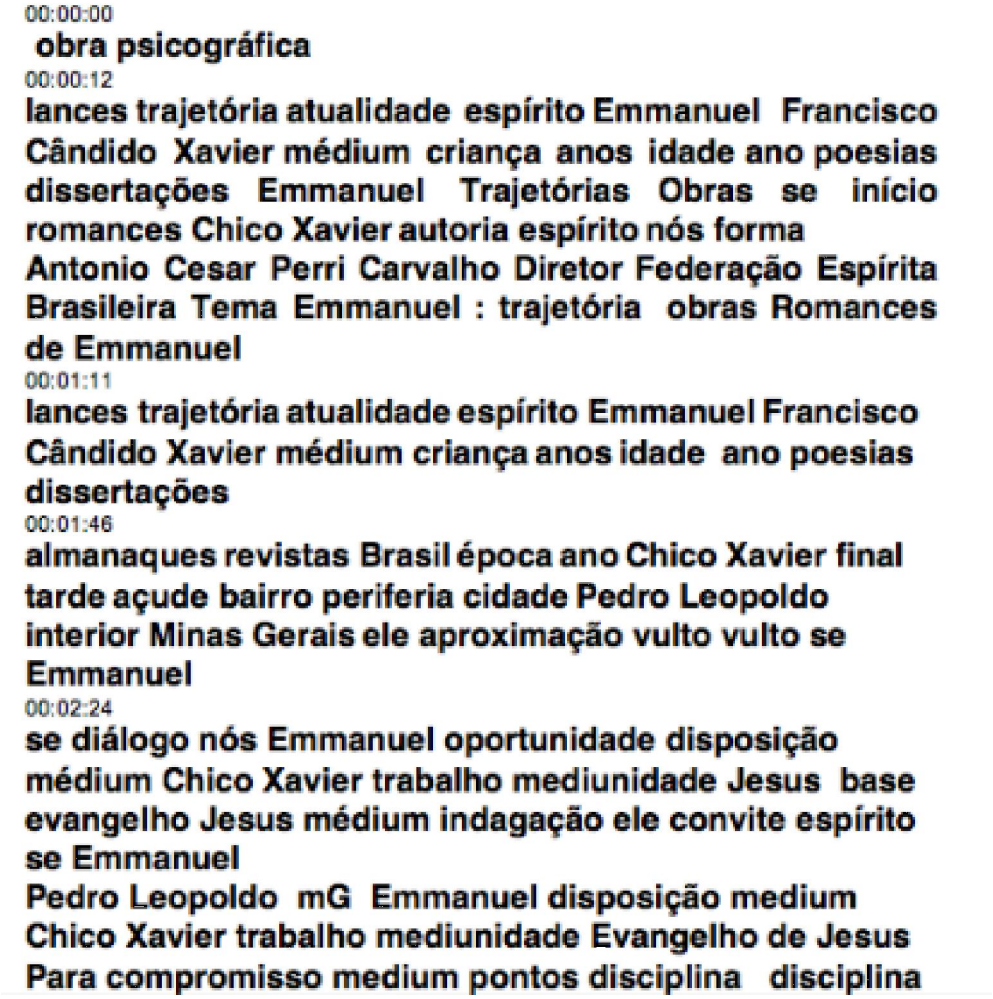
Fonte: Dados da pesquisa

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 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.

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 2, foram resultantes do processamento pelo *parser* PALAVRAS.

Figura 2 – Palavras-chave após indexação multimodal e processamento pelo parser PALAVRAS



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.

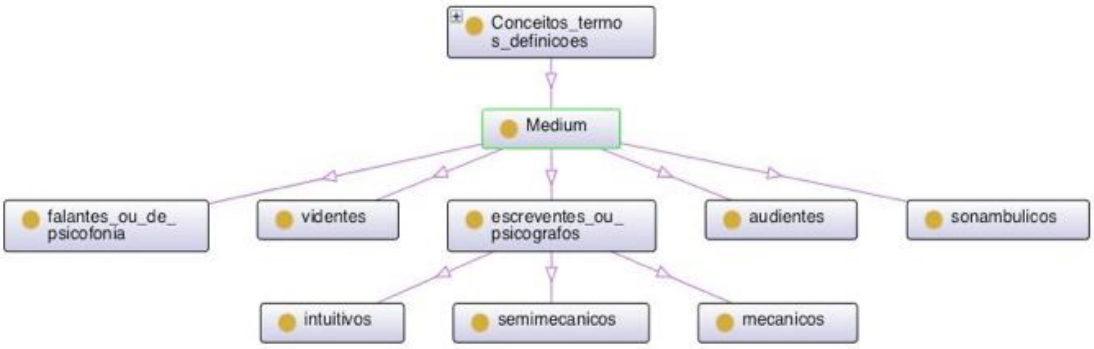
A lista de termos foi obtida por meio de pesquisas em glossários *online*, textos da área e principalmente a obra "Doutrina Espírita para principiantes", de Luis Hu Rivas, 2007. O entendimento dos termos relativos ao domínio foi facilitado pelo acompanhamento de um especialista do domínio, o próprio 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 3.

Figura 3 – Parte da hierarquia abaixo da classe "Medium" na ontologia do domínio

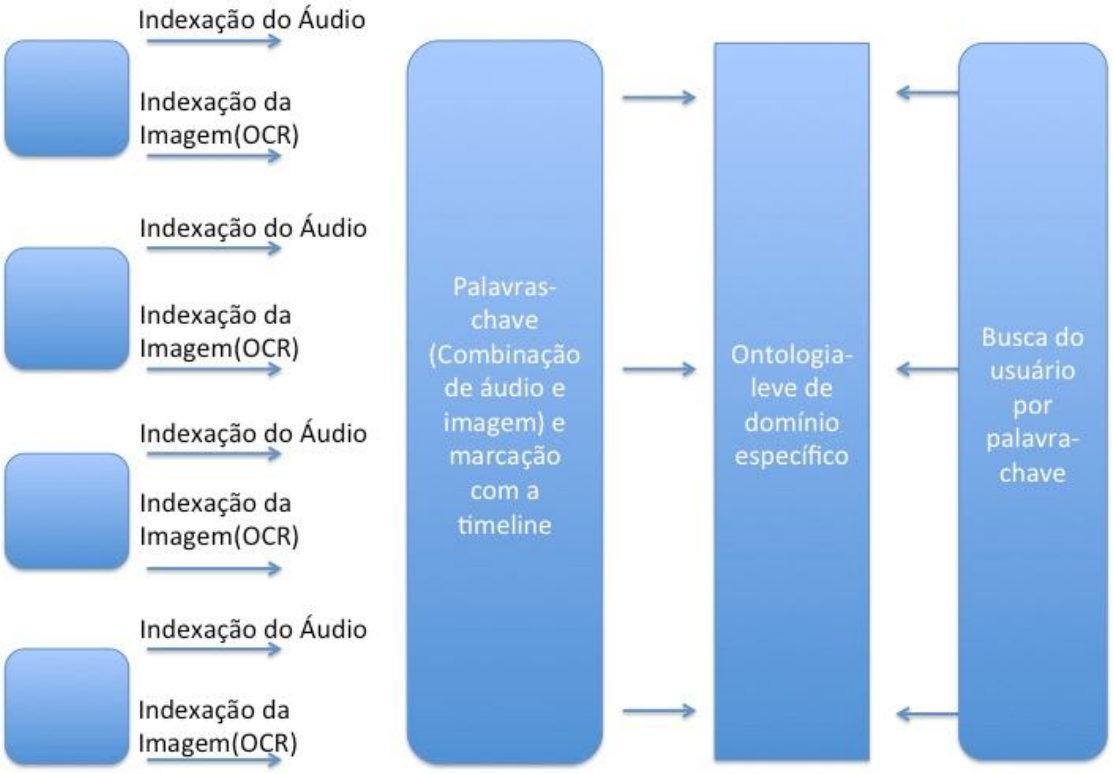


Fonte: Dados da pesquisa

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 necessário incorporar ao protótipo uma aplicação que permitisse ao usuário realizar consultas em diferentes configurações, inclusive fazendo o uso da ontologia de domínio criada. Esta aplicação permite associar as palavras-chave (ou a indexação multimodal completa) a seu respectivo trecho 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 4.

Figura 4 – Representação generalizada do método utilizado.



Fonte: Dados da pesquisa

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 dentre 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.

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).

Tabela 1 – Configurações de execução do experimento

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

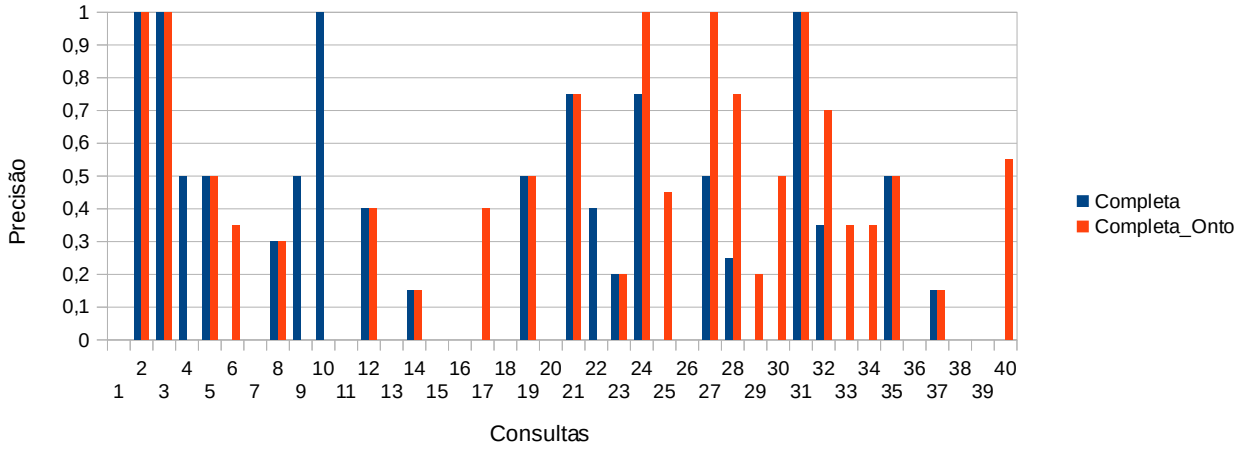
Fonte: Dados da pesquisa

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.

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.

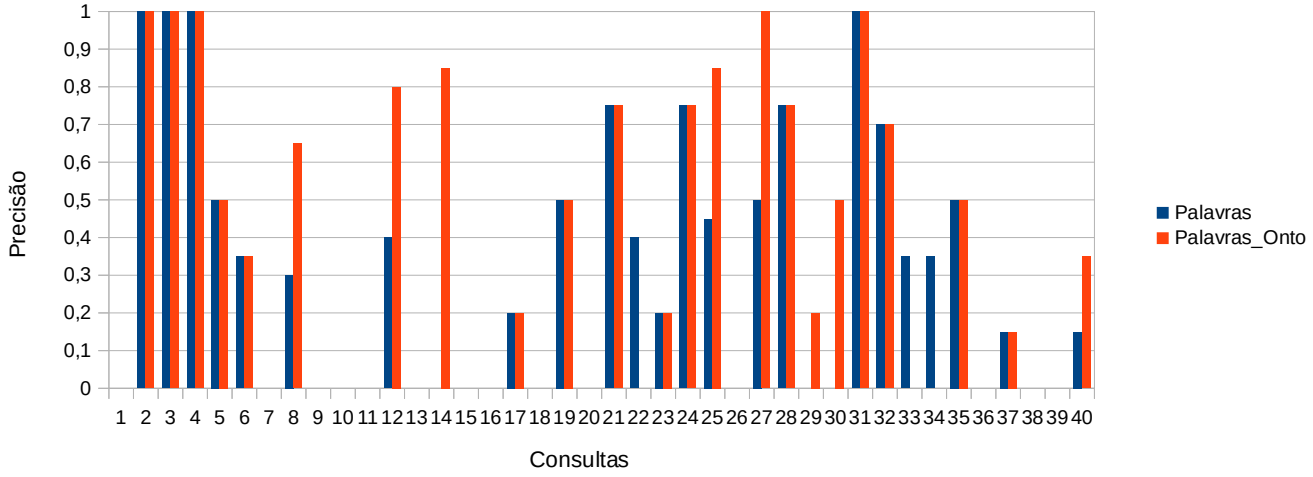
Gráfico 1 – Resultado de Precisão do experimento comparando as configurações Completa e Completa_Onto



Fonte: Dados da Pesquisa

No gráfico 1 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 2 – Resultado de Precisão do experimento comparando as configurações Palavras e Palavras_Onto



Fonte: Dados da Pesquisa

No Gráfico 2 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 E CONCLUSÃO

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.

Abstract: The necessity of retrieve the information available in videos in large repositories, have led to efforts to increase the accuracy of searches made by users. Information retrieval in long videos presents some problems because it is mainly done through subjective textual descriptions given for the entire video. The objective of this research was limited to deal of the issue of audiovisual information retrieval regarding the precision by indexing image, audio, syntax analysis and use of ontology. We developed a domain ontology, associated with an multimodal index for evaluation of precision in a audiovisual information retrieval system. There was a significant increase in the experiments of the precision of the configurations that used ontology to audiovisual information retrieval. The use of a domain ontology of the Spiritist Doctrine aligned multimodal descriptors and use of natural language processing in a lectures video-clips retrieval system on this topic, has improved the efficiency of the system in terms of accuracy.

Keywords: Audiovisual Information Retrieval. Ontology. Natural Language Processing. Multimodality.

REFERÊNCIAS

ALMEIDA, M; BAX, M. 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, R. H. **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, R. H. **Precisão no processo de busca e recuperação da informação.** Brasília: Thesaurus, 2007.

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.

BORKO, Harold. Information science: what is it? **American documentation**, v. 19, n. 1, 1968.

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

CAPURRO, R.; HjØrland, B. O conceito de informação. ***Perspectivas em Ciência da Informação***, Belo Horizonte, v.12, n.1, p. 148-207, abr. 2007.

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

CHOMSKY, N. **Syntactic structures**. 2.ed. [S.I]: Walter de Gruyter, 1957.

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.

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 abril de 2009.

GAGNON, Michel. Processamento da Linguagem Natural, 2000. In.: http://www.professeurs.polymtl.ca/michel.gagnon/Disciplinas/Bac/IA/PLN/pln.html. Acesso em outubro de 2011.

GRUBER, T. **What is an Ontology?** Disponível em:< http://www-ksl.stanford.edu/kst/what-is-an-ontology.html> 1996. Acesso em julho de 2012.

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.

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.

LE COADIC, Y. F. **A Ciência da informação**. 2a ed. Brasília: Briquet de Lemos, 2004.

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.

McGUINNESS, D. L.; HARMELEN, F. V. **OWL web ontology language overview:** W3C Candidate Recommendation, 2003. Disponível em : <http://www.w3.org/TR/2003/CR-owl-features-20030818/>. Acesso em fevereiro de 2012.

NORRIS, S. Analysing Multimodal Interaction. London, UK: Routledge, 2004.

NOY, F. N.; GUINNESS, D. L. **Ontology development 101:** a guide to create your first ontology, 2001. Disponível em: <http://ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.doc>. Acesso em abril de 2012.

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

O'HALLORAN, Kay L. **Multimodal discourse analysis:** systemic functional perspectives. London: Continuum, 2004.

RIJSBERGEN, C. J. V**. Information retrieval,** 1975. Disponível em: <http://www.dcs.gla.ac.uk/Keith/Preface.html> . Acesso em 08 jul. 2012.

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

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

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

SARACEVIC, T. Interdisciplinary nature of Information Science. ***Ciência da Informação***, Brasília, v.24, n.1, p.36-41, 1995.

SCHULTE,O.; WUNDEN,T.; BRUNNER,A. **REPLAY – An integrated lecture recording infrastructure to interactively and collectively generate learning objects,** 2008. Disponível em <http://eunis.dk/papers/p13.pdf> Acesso em junho de 2012.

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.

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, 2005.

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

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.