

# Busca multimodal para apoio à pesquisa em biodiversidade

Gabriel de Souza Fedel<sup>1</sup>, Claudia M. Bauzer Medeiros<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Estadual de Campinas (UNICAMP)  
Caixa Postal 6.176 – 13.084-971 – Campinas – SP – Brasil

`gabriel.fedel@students.ic.unicamp.br, cmbm@ic.unicamp.br`

**Nível:** Mestrado

**Ano de ingresso:** 2009

**Época esperada de conclusão:** Março/2011

**Etapas já concluídas:** Defesa da proposta e Seminário de andamento

**Palavras-chave:** Biodiversidade, Bancos de dados multimeios, Consultas multimodais, Ontologias

**Resumo.** *A pesquisa em computação aplicada à biodiversidade apresenta muitos desafios, que vão desde o grande volume de dados altamente heterogêneos até a variedade de tipos de usuários. Isto gera a necessidade de ferramentas versáteis de buscas na Web. As ferramentas disponíveis ainda são limitadas e normalmente só consideram dados textuais, deixando de explorar a potencialidade da busca por dados de outra natureza, como imagens ou sons. O objetivo deste projeto é analisar os problemas de realizar consultas multimodais com texto e imagem para o domínio de biodiversidade, propondo um conjunto de ferramentas para processar tais consultas. Espera-se que com esta busca integrada a recuperação dos dados de biodiversidade se torne mais abrangente, auxiliando os pesquisadores em biodiversidade em suas tarefas, além de incentivar usuários leigos a acessar esses dados. Este trabalho está inserido no projeto BioCORE, uma parceria entre pesquisadores de computação e biologia para criar ferramentas computacionais para dar apoio à pesquisa em biodiversidade.*

## **1. Introdução e Motivação**

Biodiversidade descreve a riqueza e a variedade do mundo natural [WWF 2010]. A pesquisa em biodiversidade, entre outras coisas, resulta em parâmetros para conservação, preservação e proteção das espécies em nosso planeta. A quantidade e heterogeneidade de dados coletados vem motivando o desenvolvimento de pesquisas em computação para seu gerenciamento.

Tais dados podem ser textuais (como nome da espécie ou tipo de habitat), imagens (como fotos da espécie), sons (gravações de ruídos de animais), croquis, entre outros. A abordagem mais comum para recuperar esses dados, a busca textual sobre texto e metadados, não considera particularidades como o conteúdo de uma imagem ou de um arquivo de áudio, o que poderia refinar o resultado.

O objetivo deste trabalho é analisar os problemas de realizar consultas multimodais combinando palavras-chave, imagem e informações taxonômicas para recuperar dados de coleta (observações na natureza) e de coleções (acervos de museus), propondo e implementando uma ferramenta para processar tais consultas. A proposta é uma continuação de trabalhos prévios e será implementada como Serviços *Web* [Wang et al. 2004].

Embora haja várias propostas de sistemas para buscas multimodais, estas costumam ser restritas a um tipo específico de usuário. A variedade de dados e necessidades de usuários em biodiversidade apresenta vários desafios que precisam ser enfrentados. Como exemplo, grupos distintos de espécies são caracterizados por aspectos visuais diferentes (e portanto não é possível usar um único método de recuperação de imagens). Nomes vulgares - o que leigos usam para identificar espécies - variam com a região geográfica ou cultura do usuário, dificultando seu uso como parâmetro de busca. Mesmo taxonomias de identificação de espécies não são consensuais, de novo criando obstáculos no processamento de consultas. Em casos de múltiplas taxonomias, por exemplo, informações adicionais multimodais podem facilitar a busca.

## **2. Conceitos Básicos**

Esta seção apresenta os conceitos básicos que são usados neste trabalho. Estes conceitos provém a fundamentação teórica da proposta apresentada na seção 3.

### **2.1. Projeto BioCORE e o Museu Virtual**

Os resultados deste trabalho fazem parte do projeto interdisciplinar BioCORE [Bio 2007] que visa desenvolver ferramentas de gerenciamento de dados para dar apoio à pesquisa em biodiversidade. Tais ferramentas visam facilitar a construção de modelos complexos e a modelagem de ecossistemas, incluindo a descoberta de relacionamentos e interações entre as espécies.

Uma parte da pesquisa setá sendo dedicada à especificação e implementação do museu virtual de zoologia da UNICAMP, construído a partir da coleção existente não-digital. O museu virtual espera atingir inicialmente dois grupos de usuários: pesquisadores em biodiversidade e visitantes. O primeiro grupo busca informações no acervo com intuito de encontrar dados para sua pesquisa. O segundo utiliza o museu para conhecer o acervo, e buscar informações sobre espécies, porém sem interesse em informações especializadas. Isto configurou a necessidade de criar ferramentas de consulta online que explorassem características específicas dos dados e que atendessem estes dois tipos de usuários.

### **2.2. Ontologias**

Uma das principais inovações introduzidas no desenvolvimento do museu é a possibilidade de expandir consultas utilizando vários tipos de ontologias [Malaverri et al. 2009]. Uma ontologia é uma especificação de uma conceitualização [Gruber 1993], a visão e o entendimento que

se tem da realidade. Ontologias procuram capturar a semântica de um domínio pelo desenvolvimento de primitivas de representação do conhecimento, habilitando máquinas a entender parcialmente o significado dos relacionamentos entre os conceitos em um domínio.

A capacidade de descrever formalmente um conhecimento e de raciocinar sobre ele torna possível adotar uma ontologia como forma de representar um domínio, padrão ou consenso de um grupo e, posteriormente, utilizar a mesma representação por meio de mecanismos e ferramentas que a processem. Já há alguns trabalhos de consultas em biodiversidade que utilizam ontologias para processar consultas, restringindo-se no entanto à busca textual.

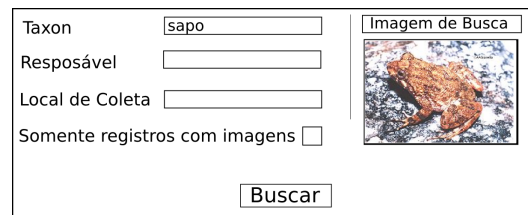
### 2.3. Consultas Multimodais

Uma facilidade não disponível em sites de museus virtuais e de sistemas de biodiversidade online é o suporte a consultas multimodais - aquelas que envolvem mais de um tipo de dado, como som ou imagens. A utilização de múltiplas modalidades permite ao sistema se beneficiar das vantagens de cada uma delas. Segundo [Yan and Hauptmann 2007], a combinação de resultados de múltiplas modalidades pode consistentemente melhorar o desempenho da recuperação quando comparada com a utilização isolada de cada modalidade.

As técnicas para fusão de modalidades se dividem em dois grandes grupos: *early fusion* e *late fusion*. Tais técnicas diferem pela maneira como são combinadas as características da modalidade. A abordagem *Early Fusion* extrai as características de cada modalidade, então as combina para resultar em uma única representação. Já na *Late Fusion*, além de extrair as características separadamente, o aprendizado também é separado, e por fim os resultados de cada modalidade são combinados para produzir uma classificação final.

## 3. Caracterização da Contribuição

### 3.1. Visão Geral e Arquitetura Proposta

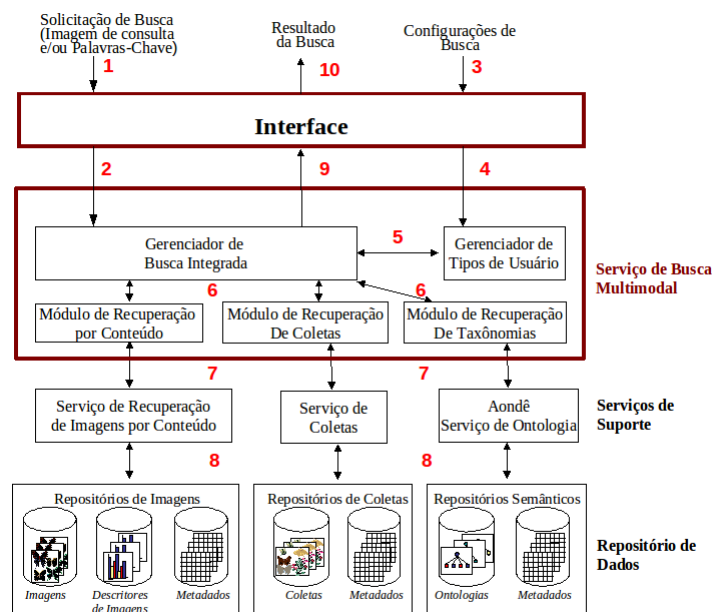


O protótipo da tela de consulta apresenta um formulário com os seguintes campos: 'Taxon' com o texto 'sapo' inserido; 'Responsável' com um campo vazio; 'Local de Coleta' com um campo vazio; e 'Somente registros com imagens' com uma caixa de seleção desmarcada. À direita do formulário, há uma seção intitulada 'Imagem de Busca' que contém uma fotografia de um sapo. Abaixo do formulário, há um botão 'Buscar'.

Figura 1. Protótipo da Tela de Consulta

A proposta deste trabalho é definir e implementar uma arquitetura de software que possibilite fazer consultas a dados de biodiversidade utilizando combinação de busca textual, por conteúdo de imagens e ontologias. A Figura 1 mostra um exemplo de uma possível consulta ao museu, combinando imagens e palavras-chave. A entrada são fotos da espécie desejada, fornecidas usuário, e as palavras-chave, algum termo que ajude a identificar a espécie procurada. Este termo pode ser científico (como uma classificação taxonômica, por exemplo a Família *Hylodidae*), ou então um nome vulgar (por exemplo, *sapo*, um dos nomes vulgares da espécie *Hylodes otavioi*). Tais dados de entrada serão comparados com três tipos de fontes de dados: a) registros de coleções do museu e observações feitas em campo sobre determinadas espécies; b) fotos de espécies; c) ontologias com informações taxonômicas e nomes vulgares. O resultado exibe ao usuário os dados encontrados, ordenados pela semelhança com os dados de entrada.

Além dos dados de entrada, também será considerado o tipo de usuário em questão, a fim de tratar suas diferença (pesquisador e leigo) e níveis de segurança, como regras de acesso. Com isto, o sistema pode refinar a consulta tornando o resultado mais apurado.



**Figura 2. Arquitetura proposta**

A Figura 2 mostra a arquitetura proposta para o sistema de consulta, que estende a proposta de [Freitas and da S. Torres 2005], aproveitando os serviços de suporte já desenvolvidos [Daltio and Medeiros 2008]. Nela estão presentes 4 camadas: Repositório de Dados, Serviços de Suporte, Serviço de Busca Multimodal e Interface. A proposta de projeto, destacada na figura, abrange a implementação do Serviço de Busca Multimodal e um protótipo de interface.

O fluxo de execução é o seguinte: a interface recebe uma solicitação de busca (1), com os respectivos parâmetros (3). Tal solicitação é passada ao serviço de busca multimodal (2), que processa a busca a partir dos serviços de suporte (6,7,8), e retorna os resultados para a interface (9,10). As configurações de busca são repassadas ao gerenciador de tipos de usuário (4) que determina os parâmetros para a busca e os encaminha para o Gerenciador de Busca Integrada (5) que determina a integração dos resultados. A figura também mostra os 3 tipos de repositórios considerados: de imagens, de coletas e semântico (ontologias).

Devido às características dos dados utilizados e os tipos de usuários que utilizarão o sistema, pretende-se utilizar a *Late Fusion* para permitir filtragens e classificações separadas por modalidade.

### 3.2. Museu Virtual e Ontologias em Desenvolvimento

A arquitetura proposta será implementada como um protótipo no site do museu de zoologia. Tal protótipo irá permitir a busca a partir de imagens e/ou palavras-chave, com acesso aos dados do acervo do museu. A versão online do museu, desenvolvida no BioCORE, conta apenas com busca textual, para pesquisadores, por meio de alguns atributos dos dados armazenados. Tal versão não permite consulta buscando nas fotos de espécies cadastradas e nem em relações taxonômicas.

O processo de consulta multimodal proposto deverá integrar ontologias com informações taxonômicas e nomes vulgares. A única ontologia disponível para as mais de 5 mil espécies do museu é obtida através de classificações taxonômicas. Não há associação com nomes vulgares. Portanto está sendo necessário constituir uma ontologia armazenando

informações taxonômicas e nomes vulgares associados aos nomes científicos. Os nomes vulgares serão associados não somente a espécies mas a qualquer nível taxonômico (como por exemplo o nome vulgar *aranha*, que é relativo a todos os artrópodes da ordem *araneae*). Esta associação, manual, está sendo feita com auxílio dos usuários.

A ontologia também será utilizada para expandir os resultados de uma consulta. Quando uma consulta retornar uma espécie, por exemplo, poderá subir um (ou vários) nível na taxonomia, retornando mais resultados que podem interessar ao usuário.

### **3.3. Estado Atual**

Já foram finalizados a especificação da arquitetura proposta e os protótipos de tela para o museu virtual. Os esforços atuais do trabalho estão centrados na especificação de ontologias com informação taxonômica e nomes vulgares, e na implementação dos módulos de recuperação (por conteúdo, de coletas e de ontologias). A recuperação por conteúdo irá utilizar trabalhos já desenvolvidos no Laboratório de Sistemas de Informação (LIS) para pré-processamento das imagens. A avaliação será feita pela comparação com um conjunto de controle definido pelos usuários.

Além desses passos em execução, ainda resta definir as regras de integração dos resultados, do gerenciador de busca integrada, e realizar testes iniciais para definir as configurações do gerenciador de tipos de usuário. Como há partes da arquitetura que precisam ser implementadas ainda não foi possível realizar testes com o sistema proposto, apenas testes parciais.

Desde o início da pesquisa foram encontradas dificuldades, principalmente referente às imagens de espécies. Elas têm uma grande variedade de características, o que vem dificultado o seu processamento para recuperação por conteúdo. A ausência de ontologias taxonômicas e informações organizadas sobre nomes vulgares também dificultou o andamento da pesquisa.

## **4. Trabalhos relacionados**

O trabalho proposto em [Freitas and da S. Torres 2005] é um sistema desktop monousuário que utiliza conteúdo, palavras-chave e a ontologia WordNet para recuperação multimodal de imagens, mostrando como realizar um processo de anotação combinando essas informações. Além disso, propõe uma arquitetura flexível para uso de múltiplas ontologias e múltiplos descritores de imagens. A solução apresentada, o ambiente OntoSAIA, sistematiza a anotação de imagens (para que o resultado da abordagem baseado em palavras-chave funcione). Tal anotação funciona com base em sugestões, ligadas a anotações feitas em outras imagens, e anotações relacionadas ao conteúdo da imagem. O ambiente OntoSAIA foi ponto de partida para nossa pesquisa, principalmente no que tange o uso combinado de imagens, palavras-chave e ontologias.

A ferramenta C-DEM [Guo et al. 2008] manipula dados biológicos e tem como finalidade minerar dados de moscas de frutas. A mineração/busca ocorre a partir de uma imagem de consulta e/ou genes e/ou palavras chaves. A base de dados utilizada nessa ferramenta (BDGP) possui mais de 70.000 fotografias digitais documentando a expressão de padrões de mais de 3000 genes, anotados com um conjunto padrão de termos. Este trabalho apresenta uma solução para uso combinado de modalidades diferentes de dados no contexto de biodiversidade, porém para um domínio muito específico, sendo as imagens referentes a padrões genéticos, enquanto nosso trabalho é dificultado por ser genérico.

O trabalho de [Kesorn and Poslad 2009] utiliza um modelo baseado em ontologias para reestruturar os conceitos semânticos presentes em legendas de imagens, buscando atingir um

resultado mais eficaz de recuperação destas imagens. O processamento semântico visa superar os problemas de busca textual e adicionar conceitos indiretamente relevantes. Fotos de esportes retiradas da Web são utilizadas nos experimentos, que mostram que técnicas baseadas em ontologias podem melhorar significativamente a efetividade dos sistemas de recuperação de imagens. Apesar deste trabalho não utilizar dados de biodiversidade, apresenta soluções para o uso de ontologias para melhorar a recuperação de imagens.

## 5. Conclusão

A possibilidade de combinar, em consultas, imagens, palavras-chave e ontologias pode resolver os problemas apresentados por buscas baseadas somente em texto ou somente em conteúdo. A busca integrada também apresenta uma possibilidade nova para que os pesquisadores em biodiversidade analisem os seus dados. A proposta também traz novidades para leigos que utilizarem o sistema, permitindo a busca combinando nomes vulgares e fotos.

As principais contribuições do trabalho são: (1) explorar a multimodalidade dos dados de biodiversidade para processar consultas, criando ferramentas para combinar resultados; (2) utilizar o tipo do usuário para dirigir a busca; (3) criar e utilizar ontologias taxonômicas e com nomes vulgares para processar a consulta. O banco de dados atualmente tem mais de 42000 dados de espécies e imagens.

## Referências

- (2007). BioCORE: Biodiversity and Computing Research. <http://www.lis.ic.unicamp.br/projects/bio-core>.
- Daltio, J. and Medeiros, C. B. (2008). Aondê: An ontology web service for interoperability across biodiversity applications. *Inf. Syst.*, 33(7-8):724–753.
- Freitas, R. B. and da S. Torres, R. (2005). Ontosaia: Um ambiente baseado em ontologias para recuperação e anotação semi-automática de imagens. *Primeiro Workshop de Bibliotecas Digitais, Simpósio Brasileiro de Banco de Dados*, pages 60–79.
- Gruber, T. (1993). Toward principles for the design for ontologies used for knowledge sharing. In *Formal Ontology in Conceptual Analysis and Knowledge Representation*.
- Guo, F., Li, L., Faloutsos, C., and Xing, E. P. (2008). C-dem: a multi-modal query system for drosophila embryo databases. *Proc. VLDB*, 1(2):1508–1511.
- Kesorn, K. and Poslad, S. (2009). Semantic restructuring of natural language image captions to enhance image retrieval. *Journal of Multimedia*, 4(5):284–297.
- Malaverri, J. G., Vilar, B., and Medeiros, C. B. (2009). A tool based on web services to query biodiversity information. In *5th Webist*.
- Wang, H., Huang, J. Z., Qu, Y., and Xie, J. (2004). Web services: problems and future directions. *J. Web Sem.*, 1(3):309–320.
- WWF (2010). Site do WWF Brasil. [http://www.wwf.org.br/informacoes/questoes\\_ambientais/biodiversidade/](http://www.wwf.org.br/informacoes/questoes_ambientais/biodiversidade/) acessada em 18/06/2010.
- Yan, R. and Hauptmann, A. (2007). A review of text and image retrieval approaches for broadcast news video. *Information Retrieval*, 10:445–484.