

Ontologias para Interoperabilidade de Modelos e Sistemas de Informação de Biodiversidade

José L. Campos dos Santos¹, José F. de Magalhães Netto², Alberto Nogueira de Castro², Andréa C. F. Albuquerque^{1,2}, Edilson Ferneda³, Luiza Alonso³, Ricardo L. da C. Rocha¹, Daniel T. de Mendonça¹

¹ Laboratório de Interoperabilidade Semântica (LIS)
Núcleo de BioGeo Informática (NBGI)
Coordenação de Tecnologia da Informação (CTI)
Instituto Nacional de Pesquisas da Amazônia (INPA)
Av. André Araújo, 2936 – Aleixo
CEP.: 69060-001 - Manaus-AM, Brasil

² Programa de Pós-Graduação em Informática (PPGI)
Departamento de Ciência da Computação (DCC)
Universidade Federal do Amazonas (UFAM)
Av. Gen. Rodrigo Otávio Jordão Ramos, 3000 – Coroado
CEP.: 69077-000 - Manaus-AM, Brasil

{lcampos, andreaa, rlc, daniel.trusman}@inpa.gov.br, {jnetto.alberto}@dcc.ufam.br, {eferneda, lualonso}@ucb.edu.br

Resumo. As pesquisas no INPA são apoiadas por sistemas computacionais que precisam dispor de recursos para analisar, sintetizar e comunicar informações e conhecimentos de forma precisa. As comunicações devem ser interpretadas corretamente pelos sistemas com que interagem, de forma a garantir implicações lógicas deriváveis simétricas. O LIS objetiva aplicar tecnologias da Web Semântica, como metadados, modelos e ontologias, aos problemas de interoperabilidades de dados e sistemas sobre a biosfera e atmosfera da Amazônia. O estudo desses componentes produz grande volume de dados sobre a biota da Amazônia e o seu gerenciamento demanda ferramentas automáticas inseridas no contexto da Web Semântica, oferecendo assim suporte para integração e disseminação de resultados. Recuperar informações utilizando um simples mecanismo de busca de palavras chaves em textos na Web não é eficaz devido ao conjunto de atividades e processos existente na geração de conhecimento científico. Tecnologias da Web Semântica indicam oferecer uma oportunidade de adicionar conteúdo semântico aos dados online permitindo que computadores possam inferir a cerca dos mesmos, integrá-los e gerar conhecimento.

1 Posição de Pesquisa

Ao longo dos anos, instituições de pesquisas, como o Instituto Nacional de Pesquisas da Amazônia (INPA), o Museu Emílio Goeldi (MPEG), têm produzido grande quantidade de dados sobre a biodiversidade, provenientes de coletas de campo, experimentos científicos, inventários, entre outras atividades.

Dados científicos, de maneira geral, encontram-se geralmente dispersos em diferentes fontes, algumas de fácil acesso (via Web), periódicos e livros científicos, relatórios técnico-científicos, dissertações e teses, e outras de difícil localização (metadados) e acesso, como arquivos, pastas e cadernos de campo, etc. Algumas dificuldades devem-se também à falta de uma política para gestão destes dados e do conhecimento. É visível a crescente demanda por estes dados em diversas aplicações consideradas importantes, como avaliação de impacto ambiental, definição de áreas de preservação ambiental, proteção de espécies ameaçadas, recuperação de áreas degradadas, bioprospecção, estabelecimento de políticas públicas, legislação ambiental, entre outras.

Os dados e conhecimentos científicos sobre biodiversidade exercem um importante papel no atendimento a demandas deste tipo, pois acumulam investimentos de anos em exploração e pesquisa. No entanto, tornar isso cada vez mais acessível ao público de forma adequada, rápida e confiável, impõe o desenvolvimento de sistemas de informações capazes de extrair, armazenar, gerenciar, analisar, integrar e disseminar os diferentes dados das diversas fontes de dados de biodiversidade [1,2].

Biodiversidade e Cenários de Integração de Dados Biodiversidade é hoje um termo científico citado e propagado em todo o mundo. O conceito de biodiversidade procura referir e integrar toda a variedade que encontramos em organismos vivos, nos mais diferentes níveis. A situação da biodiversidade neste início de século XXI pode ser caracterizada pela combinação de processos acelerados de destruição de ecossistemas primários associados a esforços mobilizadores para a conservação e uso sustentável da biodiversidade e de grandes avanços em tecnologia de informação e comunicação de dados. Esta combinação de fatores está conduzindo à emergência de uma nova área de desenvolvimento científico e tecnológico, denominada informática para biodiversidade [3].

A importância de projetos integradores de informações sobre a biodiversidade já está sendo sentida há algum tempo. Investimentos para o desenvolvimento de sistemas de informação, ou mesmo formação de redes de informação sobre biodiversidade já vêm sendo tomados em escala institucional, regional, nacional e internacional, seja no âmbito de governos, organizações ou áreas temáticas. Projetos como Experimento da Grande Escala da Biosfera e Atmosfera na Amazônia (LBA), Global Biodiversity Information Facility (GBIF), Sistemas de Bases Compartilhadas de Dados sobre a Amazônia (BCDAM), Programa de Pesquisas em Caracterização, Conservação e Uso Sustentável da Biodiversidade do Estado de São Paulo (BIOTA-FAPESP), Programa de Pesquisa em Biodiversidade na Amazônia (PPBio), Species Link entre outros, representam tais investimentos.

Um dos problemas mais frequentes é como integrar dados das múltiplas fontes disponíveis. A integração de dados requer alguns cuidados, tais como, integração semântica de dados, interoperabilidade, etc.

Ressalta-se que, sistemas que manipulam dados de uma forma geral, mas especificamente sobre biodiversidade requerem interoperabilidade (capacidade de um sistema se comunicar com outro sistema) das diversas fontes utilizadas [4, 5].

Desta forma, heterogeneidade semântica, onde um único fato pode ter mais de uma descrição, dependendo de quem vê/descreve, deve ser considerada. Por exemplo, tipo de vegetação também pode ser entendido como bioma. Uma estratégia para tratar disto é o uso de ontologia que explicitamente define termos do esquema de dados e vocabulário para auxiliar na eliminação dos conflitos semânticos [6].

2 Principais Tópicos de Pesquisa

2.1 Desenvolvimento de Ontologias de Domínio de Biodiversidade

A pesquisa em biodiversidade é um campo multidisciplinar que requer a cooperação de vários tipos de pesquisadores. Biólogos realizam diferentes tipos de atividades, incluindo coletas de campo, análises e síntese de dados sobre os espécimes coletados, seus habitats e correlações com outros seres vivos, construindo modelos capazes de descrever essas interações. Os dados disponíveis vêm sendo coletados em vários lugares do mundo, sendo publicados em formatos heterogêneos, que são especificados para o uso de inúmeros padrões. Este cenário é caracterizado por sua heterogeneidade intrínseca – não apenas de dados e modelos conceituais utilizados, como também de necessidades e perfis dos especialistas que coletam e analisam os dados [1].

O grande volume de dados coletados e a diversidade de espécimes atuam como fatores intensificadores da complexidade deste cenário. As estimativas sobre o número de espécies vegetais e animais existentes no mundo variam entre 10 e 50 milhões, das quais apenas 1,5 milhões de espécies são atualmente classificadas pelos cientistas (estimativa realizada pelo WWF – Fundo Mundial para a Natureza). Entre os especialistas, o Brasil é considerado um país megadiverso – abrangendo cerca de 20% das espécies conhecidas em todo o mundo. Essa diversidade motiva esforços na coleta de dados, dando origem, por consequência, a um grande volume de informações. Isto gera uma demanda por mecanismos sofisticados de armazenamento, gerenciamento, compartilhamento, processamento e mineração, que permitam uma análise integrada e correlacionada desses dados¹.

Muitos ramos da biologia possuem seus domínios descritos por ontologias consensuais (reflete um conhecimento consensual aceito por um grupo, comunidade), como a Gene Ontology [7] (em genética) e TAMBIS Ontology (*Transparent Access to Multiple Bioinformatics Information Sources*) [8] (em biologia molecular). Em biodiversidade, entretanto, ainda não existe uma ontologia consensual, embora existam vários especialistas envolvidos em iniciativas multinacionais, como o GBIF [9].

Ainda há muito trabalho a ser realizado na especificação de uma ontologia para esse domínio – resultado da variedade de características dos dados e perfis de especialistas. Esta é uma das razões que contribui para que a integração de dados de biodiversidade e de estudos ecológicos não seja considerada trivial. Soluções para interoperabilidade são necessidades reais para pesquisa nesse domínio. Até mesmo as classificações taxonômicas de espécies são alvo de discussão entre pesquisadores. Algumas árvores taxonômicas são claramente definidas – como em zoologia, para mamíferos. Entretanto, ainda existem divergências de autores em vários domínios além de muitas espécies para serem classificadas ou reclassificadas – como os insetos, por exemplo. Problemas similares afetam também descrições de habitats de espécies, como ocorrem com as várias classificações de solo existentes no mundo. Somente no Brasil, existem dois sistemas de classificação de solos considerados oficiais. Com isso, uma mesma amostra de solo pode não apenas ser representada por diferentes identificadores e nomes, como também pertencer a classes de ontologias distintas, de acordo com o sistema de classificação adotado [10].

Metadados (dados sobre dados) e ontologias são complementares e constituem os blocos de construção da Web Semântica. Não permitem ambigüidade de significados e fornecem respostas mais precisas. Ontologia combina esquemas de metadados, fornecendo um vocabulário controlado de conceitos, semanticamente definidos e processáveis por

¹ World Wide Fund For Nature, fonte <http://www.wwf.org.br>

máquina. Por definir teorias compartilhadas e de domínio comum, ontologia auxilia pessoas e máquinas a se comunicar concisamente, suportando a troca de semântica e não somente sintaxe.

Ontologias foram desenvolvidas na filosofia desde os anos 60 e, recentemente, algumas teorias foram propostas na área de Ontologia Aplicada à Ciência da Computação com o nome de Ontologias de Fundamentação [11].

A necessidade da utilização de linguagens de modelagem conceitual ontologicamente fundamentadas e ontologias pertencentes a um domínio específico é apresentada na literatura [12, 13, 14]. Isto resulta da preocupação com os aspectos de interoperabilidade e da falta de adequação das linguagens de representação para mitigar estas questões. Apesar disso, estas linguagens ainda não são largamente utilizadas na prática. Uma das principais razões é a necessidade de um especialista de alto nível para manipular os conceitos filosóficos que as baseiam.

Uma ontologia de domínio, no sentido usado pelas demais comunidades em computação, é um tipo particular de modelo conceitual. Em particular, é um modelo conceitual que deve satisfazer o requisito adicional de servir como uma representação de consenso (ou modelo de referência) de uma conceituação compartilhada por uma determinada comunidade. Portanto, se uma ontologia de domínio é, antes de qualquer coisa, um modelo conceitual, uma linguagem adequada para representação de ontologias de domínio deve satisfazer os requisitos gerais de uma linguagem adequada para modelagem conceitual, ou seja, deve ter como teoria subjacente uma ontologia de fundamentação.

Tendo como cenário, a realidade da biodiversidade hoje e a tendência de se utilizar engenharia de ontologias para o desenvolvimento das mesmas, o LIS tem se dedicado ao desenvolvimento e à implementação de uma ontologia de domínio de biodiversidade, com recursos para validação e verificação da expressividade da linguagem para projeto de ontologias de domínio OntoUML [15].

Por se tratar de um domínio amplo e complexo, têm-se desenvolvido uma versão preliminar, baseada em uma delimitação dos requisitos do domínio. Esta área de pesquisa do LIS tem vislumbrado com clareza as disparidades do poder de expressividade semântica das linguagens que utilizam ontologias como modelo conceitual e como implementação.

Quando se utiliza engenharia de ontologias onde há linguagens para ontologia de referência (ou ontologia como linguagem de modelo conceitual) no nível de análise e, linguagens para criação de ontologias leves (ontologias como linguagem de implementação) no nível de implementação, observa-se: conceitualmente a completude de uma especialização é semanticamente relevante, o mesmo é indiferente quando tratamos da ontologia como implementação; algumas estruturas/estereótipos encontrados nas linguagens no nível de análise não conseguem ser mapeados para uma linguagem no nível de implementação. Têm-se uma ontologia bem modelada conceitualmente, mas no momento da implementação, não são encontrados recursos equivalentes para representar aquela situação específica. Não há preciosismo no nível de implementação. Nestes casos, faz-se necessário compreender conceitualmente a estrutura/estereótipo em questão e implementar o conceito utilizando a linguagem do nível de implementação.

2.2 Framework para Integração de Informações Biológicas

Com o desenvolvimento da ontologia de biodiversidade, surge a necessidade de ferramentas que possam fazer uso dessa ontologia. O desenvolvimento de um framework para integração de dados biológicos e aquisição de conhecimento baseado na ontologia de biodiversidade está sendo desenvolvido. Este framework consiste em um protocolo orientado pela ontologia de biodiversidade, com o objetivo de extrair informações relevantes ao domínio. O resultado do protocolo gera uma nova base de dados integrados, que será disponibilizado para uso dos pesquisadores, bem como se tornará parte da Web [17, 18, 19, 20, 21].

2.3 Visualizador de Ontologias Utilizando Árvore Hiperbólica

Ontologias são hierarquias de conceitos, com seus relacionamentos e até instâncias relacionadas a esses conceitos. Para visualizar todo esse esquema, é necessário uma estrutura que permita representar todos esses elementos. Portanto, árvores hiperbólicas se torna a opção mais relevante, devido a sua característica de hierarquização e a possibilidade de representação de relações e mesmo, atributos. Para este trabalho, está sendo usado uma suite java, chamada treebolic, essa ferramenta ajuda no processamento de dados hierárquicos em formato hiperbólico. No entanto, é necessário o mapeamento de um arquivo OWL para XML, já que a treebolic não suporta o formato OWL.

2.4 Web Semântica Aplicada às Coleções Biológicas do INPA

A Web Semântica permite a divulgação de dados na Internet através de um formato comum com o objetivo de integrar ou combinar bases de dados provenientes de diversas fontes. O INPA possui várias coleções de dados, principalmente científicos, que podem ser divulgadas e utilizadas na pesquisa e desenvolvimento da Amazônia [22]. Objetivamos investigar como a utilização das tecnologias da Web Semântica, dentre elas os recursos da ferramenta Rexplorator desenvolvida na Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio, podem trazer melhorias no trabalho de

pesquisa através do processamento da semântica das coleções de dados biológicos do instituto. A abordagem utilizada é de desenvolver casos de uso junto com os próprios pesquisadores, através de operações simples em cima dos modelos RDFS (*Resource Description Framework Schema*) das próprias bases. Os casos de uso poderão ser reutilizados por outros pesquisadores, inclusive de domínios de pesquisa diferentes. Neste processo de reutilização é possível que os casos de uso sejam customizados e evoluídos colaborativamente no próprio ambiente em que foram desenvolvidos. Espera-se que, ao final de um período de treinamento e utilização das tecnologias, possam ser geradas aplicações Web que abstraíam os modelos RDF nos quais os dados estão representados tornando possível o acesso às informações por outros pesquisadores que não conheçam o modelo RDF. Essa facilidade de acesso, além de permitir consultas a bases semânticas por usuários leigos em um dado domínio de pesquisa, também permitirá que pesquisadores possam realizar consultas transdisciplinares enriquecendo sua visão no desenvolvimento da pesquisa, bem como seu poder nas tomadas de decisões políticas, econômicas e sociais, e, conseqüentemente, uma melhor gestão do conhecimento.

3 Breve Histórico do Grupo e Membros

A motivação proporcionada pela missão do INPA, que é “Gerar e disseminar conhecimentos e tecnologias e capacitar recursos humanos para o desenvolvimento da Amazônia”, levou a Coordenação de Ações Estratégicas criar o Núcleo de BioGeo Informática, cujo objetivo é de dar apoio as ações de pesquisas no domínio das pesquisas e prover soluções utilizando o estado-da-arte da computação.

Para tratar da complexidade das atividades de desenvolvimento de aplicativos da biociência, buscamos parcerias com grupos de pesquisas no contexto da Web Semântica e como resultante, sentiu-se a necessidade da implantação de um Laboratório de Interoperabilidade Semântica (LIS), que ocorreu no final de 2010. O grupo é formado por 8 alunos colaboradores e 7 pesquisadores, cuja biografia resumida é apresentada a seguir:

O Laboratório de Interoperabilidade Semântica (LIS) é formado por 8 alunos colaboradores e 7 pesquisadores, cuja biografia resumida é apresentada a seguir:

José Laurindo Campos dos Santos, graduado em Engenharia Modalidade Civil pelo Instituto de Tecnologia do Amazonas (1984). Mestre em Ciência da Computação pela Universidade Federal da Paraíba (1988) - área de concentração: Banco de Dados e Inteligência Artificial. Doutorado em Ciência da Computação pela Universidade de Twente Holanda (2003) - área de concentração: Banco de Dados Não Convencionais. Atualmente é o Coordenador de Tecnologia da Informação, do Núcleo de BioGeo Informática, o Grupo de Dados e Sistemas do Programa LBA (Experimento de Grande Escala da Biosfera-Atmosfera na Amazônia), e do Laboratório de Interoperabilidade Semântica do Instituto Nacional de Pesquisas da Amazônia - INPA.

José Francisco de Magalhães Netto é professor da Universidade Federal do Amazonas e trabalha com Informática na Educação desde 1992, tendo orientado projetos de conclusão de graduação na área. Mais recentemente, a partir de 2006, tem orientado, também, projetos de Mestrado na área de Inteligência Artificial aplicada em Informática na Educação, Web Semântica, além de projetos de iniciação científica. Atua em cursos de Educação a Distância, tendo participado de cursos de capacitação de professores, de 2003 a 2004, em projetos coordenados pela Universidade Federal do Espírito Santo, mantidos pelo MEC. Tem experiência na elaboração de material didático para cursos nessa modalidade. Atualmente, também, desenvolve trabalhos visando a difusão de Novas Tecnologias direcionados para escolas, como, por exemplo, coordenando projetos baseados em Robótica Educacional e em Laboratórios Virtuais, acessáveis via Web, possibilitando acesso mais amplo a recursos ainda escassos, como robôs e equipamentos básicos de laboratórios.

Alberto Nogueira de Castro Junior, graduado em Engenharia Civil pela Universidade Federal do Amazonas (1986), graduação em Tecnologia Eletrônica pelo Instituto de Tecnologia da Amazônia (1984), mestrado em Automação Industrial pela Universidade Federal do Espírito Santo (1993), doutorado em Computer Science - Artificial Intelligence pela University of Edinburgh (1998) e pós-doutorado pela University of Edinburgh (2010). Atualmente é Professor adjunto da Universidade Federal do Amazonas e Membro de corpo editorial da Revista Brasileira de Informática na Educação (1414-5685). Tem experiência na área de Ciência da Computação, com ênfase em Metodologia e Técnicas da Computação. Atuando principalmente nos seguintes temas: programação e lógica, interfaces inteligentes, e sistemas baseados em conhecimento.

Luiza Beth Nunes Alonso, graduada em Ciências Sociais pela Universidade de São Paulo (1975), Mestre (1981) e Doutora (1985) em Educação pela Universidade de Harvard (EUA), atualmente é professora, pesquisadora e diretora do Mestrado em Gestão do Conhecimento e da Tecnologia da Informação da Universidade Católica de Brasília. Vem desenvolvendo trabalhos relacionados a temas tais como Complexidade e Transdisciplinaridade. É membro participante do CIRET (Centre International de Recherches et Études Transdisciplinaires) e membro fundadora do CETRANS (Centro de Educação Transdisciplinar).

Edilson Ferneda, Professor Titular do Mestrado em Gestão do Conhecimento e da Tecnologia da Informação da Universidade Católica de Brasília, graduado em Tecnologia da Computação pelo Instituto Tecnológico da Aeronáutica (ITA, 1979), Mestre em Sistemas e Computação pela Universidade Federal da Paraíba (UFPB, 1988) e Doutor em Ciência da Computação pelo LIRMM / Universidade de Montpellier II (França, 1992). Trabalha na problemática da

mediação da máquina na atividade humana de modelagem e aquisição do conhecimento desde 1988, com diversos trabalhos em Sistemas de Apoio à Descoberta Científica. Também desenvolve trabalhos relacionados à Inteligência Artificial Aplicada (Data Mining, Text Mining, Descoberta de Conhecimento em Bancos de Dados), Informática na Educação (Sistemas Tutores Inteligentes).

Andréa Corrêa Flôres Albuquerque, graduada em Processamentos de Dados pela Universidade Federal do Amazonas - UFAM (1996) e em Ciências Econômicas pelo Centro Integrado de Ensino Superior do Amazonas - CIESA (1997), Pós-Graduação Lato Sensu em TeleInformática na UFAM (1999). Mestrado em Informática na UFAM (cursando). Atualmente é pesquisador colaborador do Instituto Nacional de Pesquisas da Amazônia no Núcleo de BioGeo Informática / Laboratório de Interoperabilidade Semântica, além de prestar consultoria nas áreas de banco de dados para biodiversidade, bioinformática, sistemas de informações geográficas, modelagem de dados de biodiversidade, atuando principalmente nos seguintes temas: ontologias, extração semântica de dados, modelagem conceitual.

Ricardo Luís da Costa Rocha, graduado em Processamento de Dados pela Universidade Federal do Amazonas (2000) e especialização em Tecnologia Web também pela Universidade Federal do Amazonas (2007). Atua como tecnologista pleno no Instituto Nacional de Pesquisas da Amazônia. Tem experiência na área de Ciência da Computação, com ênfase em Redes de Computadores, atuando principalmente nos seguintes temas: gerência de redes de computadores, serviços de redes em software livre, segurança em redes, tecnologia web. Atualmente, cursa o Mestrado em Ciências em Informática da Pontifícia Universidade Católica do Rio de Janeiro - PUC-Rio.

Daniel Trusman de Mendonça, graduando em ciência da computação na Fundação Centro de Análise, Pesquisa e Inovação Tecnológica (FUCAPI) com o trabalho sobre a Implementação de uma Ontologia de domínio sobre Biodiversidade. Desde 2009, faz parte do Núcleo de BioGeo Informática (NBGI) do INPA.

Referências

- [1]CAMPOS DOS SANTOS, J.; de BY, R.A. ; Magalhães, C. “A Case Study of INPA's Bio-DB and an Approach to Provide an Open Analytical Database Environment”. *International Archives of Photogrammetry and Remote Sensing*, 33 (B4): 155-163, 2000.
- [2]UMMINGER, B.;YOUNG, S. “Information Management for Biodiversity: a Proposed U.S. National Biodiversity Information Center”. In: Reaka-Kudla, M.L.; Wilson, D.E. & Wilson, E.O. (eds.), *Biodiversity II: Understanding and Protecting Our Biological Resources*. Washington, D.C., Joseph Henry Press. p. 491-504, 1997.
- [3]CANHOS, V.; “Informática para Biodiversidade:Padrões, Protocolos e Ferramentas”. *Ciência e Cultura*, Apr./June 2003Vol. 55, No. 2, p.45-47. ISSN 0009-6725.
- [4]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L. “Biological Ontology Modeling Supported by CLOSi Database Schemas”. In: *Proceedings of the 19th International Conference Informatics for Environmental Protection - ENVIROINFO 2005*. Sharing Environmental Information. By Hřebiček, J., Ráček, J. (Eds.). September 07-09, Masarykova universita, 2005. 1450 pp. Sborník konference. ISBN 80-210-3780-6.Brno, Czech Republic.
- [5]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L. “Applying Ontology for Amazon Biodiversity Data Extraction”. In: *Proceedings of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics - WMSCI 2005*. Vol.1; 20050710-13. July 10-13, 2005. Orlando,FL(US).
- [6]GUARINO, N. “Understanding, Building and Using Ontologies: A Commentary to Using Explicit Ontologies in KBS Development”, by van Heijst, Schreiber, and Wielinga. *International Journal of Human and Computer Studies*, v.46, n.2/3, p. 293-310, 1997.
- [7]ASHBURNER, M. *et al* “Gene ontology: tool for the unification of biology”. *The gene ontology consortium*. *Nature Genetics*, 25(1):25–29, May 2000.
- [8]BAKER, P. G. *et al*”TAMBIS–Transparent Access to Multiple Bioinformatics Information Sources”. In *Int Conf Intelligent Systems for Molecular Biology*, volume 6, pages 25–34, Montreal, Canada, June 1998.
- [9]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J., L. “Ontology Supported by CLOSi Data Schemas in the Semantic Web Context”. In *Proceedings of ITEE 2005, Second International ICSC Symposium on Information Technologies in Environmental Engineering*, By Walter Leal Filho, Jorge Marx Gomez, Claus Rautenstrauch (Editors). September 25-27, 2005 Otto-von-Guericke-Universität Magdeburg, Germany. ISBN 978-3832243623
- [10]PEDRAZA, J. I. S.; ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L. "Criação e Mapeamento de Ontologias de Domínio de Biodiversidade". Em *Anais da II Escola Regional de Informática – Informática e os Desafios Regionais (ERIN 2010)*. 6-8 Outubro, 2010. Manaus, Brasil, ISSN 9772178375006.
- [11]GUIZZARDI, G. “Ontological Foundations for Structural Conceptual Models”. PhD Thesis (CUM LAUDE), University of Twente, The Netherlands. Published as the same name book in Telematica Institut Fundamental Research. Series No. 15, ISBN 90-75176-81-3 ISSN 1388-1795; No. 015; CTIT PhD-thesis, ISSN 1381-3617; No. 05-74. Holanda, 2005
- [12]GUIZZARDI, G. “The Role of Foundational Ontology for Conceptual Modeling and Domain Ontology Representation”. *Proceedings of 7th DB&IS*, Vilnius, IEEE Press 2006
- [13]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L.; DE MAGALHÃES NETTO, J. F. "Modeling Complex Domain Ontology Based on the Unified Foundational Ontology". *Extended Proceedings of the 4th. Latin American Conference on Computer Human Interaction (CLIHC 2009) in conjunction with the 7th. Latin American Web Congress (LAWEB 2009) / Alberto L. Morán, comp. Ensenada, Baja Calif. : Universidad Autónoma de Baja California, 2009. November 9-11th, Mérida, Yucatán, México. ISBN: 978-607-7753-32-2*
- [14]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L. “Improving Textual Ontology Representation by Extending the Equivalent Programming Language for OSM”. In: *Proceedings of the International Symposium on Generalization of Information –*

- ISGI 2005, 20th Codata International Conference - CODATA 2005. Lecture Notes in Information Sciences, Horst Kremers (ed.). September 14-16, 2005, Berlin, Germany. Pp131-146. ISBN 3-00-016253-4.
- [15]BENEVIDES, A.; GUIZZARDI, G. "A Model-Based Tool for Conceptual Modeling and Domain Ontology Engineering in OntoUML". Lecture Notes in Business Information Processing, 2009
- [16]MORIN, E. *Método IV. As Idéias: Sua Natureza, Vida, Habitat e Organização*. Lisboa. Publicações Europa-América, 1998
- [17]MENDONÇA, D. T.; ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L. "Framework para Integração de Informação Biológica Orientado por uma Ontologia". Em Anais da II Escola Regional de Informática – Informática e os Desafios Regionais (ERIN 2010). 6-8 Outubro, 2010. Manaus, Brasil, ISSN 9772178375006.
- [18]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L.; MENDONÇA, D. T.; MAGALHÃES NETTO, J. F. de. "A Negotiation Protocol for Data Integration Driven by Ontology". Edited by Eduardo Tomé. In Proceedings of the 11th. European Conference on Knowledge Management (ECKM 2010). 2 -3 September, 2010. Universidade Lusíada de Nova Famalicão, Portugal, pp. 1-10, ISBN 978-1-906638-71-9.
- [19]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L.; MAGALHÃES NETTO, J. F. "A Strategy for Biodiversity Knowledge Acquisition Based on Domain Ontology". In Proceedings of the 9th. International Conference on Intelligent Systems Design and Application (ISDA 2009). November 30 th –December 2th, 2009. Pisa, Italy, pp. 1143-1148, ISBN 978-0-7695-3872-3.
- [20]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L.; MAGALHÃES NETTO, J. F. "Biodiversity Ontology and Semantic Web for Improving Biological Data Integration". Em: Anais da Semana de Informática, Geotecnologias e Encontro de Software Livre em Santarém – SIGES 2009, 6ª. edição. 21 a 25 de Outubro de 2009, Santarém, Pará, Brasil. ISSN/ISBN 978-85-247-0420-8.
- [21]ALBUQUERQUE, A. C. F.; CAMPOS DOS SANTOS, J. L.; MAGALHÃES NETTO, J. F. "Ontologia para Integração de Dados de Biodiversidade". Em: Anais da I Escola Regional de Informática, Regional Norte 1. Interação X Computação – ERIN 2009. 04 a 06 de Março de 2009, Manaus, Amazonas, Brasil. ISBN 978-85-7669-222-5.
- [22] AMORA, A. de C.; CAMPOS DOS SANTOS, J. L.; LIMA, L. P.; CARDOSO, A. V. "Tecnologia Web 2.0 para disseminação de informações sobre a Biodiversidade da Amazônia". Em Conferência Científica Internacional - Amazônia em Perspectiva Ciência Integrada para um Futuro Sustentável, 17 a 20 de novembro de 2008. Manaus, Amazonas, Brasil.