

# ***e-DIVA*: Uma Infraestrutura Computacional para Gestão de Dados do Programa PELD**

**Andréa C. F. Albuquerque<sup>1</sup>, José L. Campos Dos Santos<sup>1</sup>, Sidnéia A. Amadio<sup>2</sup>,  
Cláudia P. de Deus<sup>2</sup>, Maria Angélica de A. Corrêa<sup>3</sup>, Jansen A. S. Zuanon<sup>2</sup>**

<sup>1</sup> Laboratório de Interoperabilidade Semântica (LIS), Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus-AM, Brasil

<sup>2</sup> Coordenação de Biodiversidade (COBIO), Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus-AM, Brasil.

<sup>3</sup> Departamento de Ciências Pesqueiras, Faculdade de Ciências Agrárias (FCA), Universidade Federal do Amazonas (UFAM), Manaus-AM, Brasil.

andreaalb.1993@gmail.com, laurindo.campos@inpa.gov.br,  
amadio@inpa.gov.br, claudias@inpa.gov.br, angelicacorrea2011@gmail.com, zuanon@inpa.gov.br

**Abstract.** *In order to identify the impacts of climatic variation on the aquatic life of the Central Amazon Basin, a new Long-Term Ecological Research Program: Varzea Diversity (PELD-DIVA) was proposed. This work presents e-DIVA computational infrastructure, which is a process of several ichthyological databases integration related to different research sites, besides offering a computational architecture for management and analysis of the data related to the Program.*

**Resumo.** *Com o propósito de identificar os impactos da variação climática na vida aquática da várzea da Amazônia Central, foi proposto um novo sítio do Programa de Pesquisa Ecológica de Longa Duração: Diversidade da Várzea (PELD-DIVA). Este trabalho apresenta a infraestrutura computacional e-DIVA, que constitui um processo de integração de diversas bases de dados ictiológicas referentes a diferentes sítios de pesquisa, além de oferecer uma arquitetura computacional para gestão e análise dos dados relacionados ao Programa.*

## **1. Introdução**

Desde a revolução industrial, o planeta vem apresentando um aumento significativo nos gases de efeito estufa (GEE) na atmosfera [Parmesan, 2006]. Como consequência, observa-se um aquecimento geral do planeta com impacto biológico ainda não devidamente dimensionado. As mudanças climáticas interferem proeminentemente no solo, água, nas espécies animais e vegetais, potencializando as preocupações nas áreas de segurança alimentar [Val, 2011]. Ações para reverter o quadro alarmante de mudanças climáticas são essenciais e, como mecanismo para essa reversão, é imprescindível o conhecimento dessas mudanças sobre a vida no planeta. A Amazônia, por conta da vasta extensão territorial, hidrografia surpreendente e fauna e flora peculiar, merece atenção especial dos governos e da comunidade científica.

A Amazônia é megadiversa, tanto em termos de espécies habitando a região como um todo, como coexistindo em um mesmo ponto. Entretanto, apesar de ser a região de

maior biodiversidade do planeta, apenas uma fração dessa biodiversidade é conhecida [Oliveira Júnior, 2015].

A bacia do rio Amazonas é a maior bacia hidrográfica do mundo despejando cerca de um sexto de toda a água doce dos continentes para os oceanos do mundo. A idade deste ecossistema, sua posição geográfica em proximidade com a linha do equador e a enorme diversidade de seus *habitats* aquáticos, produziram uma rica fauna ictiológica. Cerca de 2.500 espécies de peixes já foram descritas, quantidade superior a que é encontrada nas demais bacias do mundo e, estima-se que mais de 1.000 novas espécies permaneçam desconhecidas. A capacidade pesqueira atinge o volume de 450.000 toneladas de peixe por ano, favorecendo uma dieta proteica saudável para a região amazônica como um todo [Junk *et al.*, 2007]. Ao longo das últimas décadas, a pressão sobre os ecossistemas e os *habitats* aquáticos tem aumentado constantemente, principalmente devido às mudanças climáticas, à destruição em grande escala da cobertura vegetal natural pelas agroindústrias e madeireiras. Estas atividades colocaram a biodiversidade aquática, em grave risco. Muitas espécies de peixes de cabeceiras têm distribuições restritas e são, portanto, particularmente vulneráveis à degradação ambiental em larga escala.

O conhecimento da diversidade biológica de água doce e o entendimento da distribuição dessas espécies são estágios importantes para sua conservação. O estudo da biodiversidade de peixes, além de gerar conhecimento científico sobre a região, pode responder a questões ecológicas relevantes, podendo ser considerados indicadores de qualidade do ecossistema. As populações de peixes de um rio são reflexos do conjunto de fatores bióticos e abióticos que ocorrem em determinado momento e de aspectos como a disponibilidade de locais de alimentação, refúgio e reprodução, fundamentais para que ocorra o estabelecimento das espécies [Bennemann *et al.*, 2000]. Neste contexto, investigar os impactos das variações climáticas em ambientes aquáticos na Amazônia, contribui para entender os efeitos desses extremos sobre a vida na região, contribuindo para os processos de tomada de decisão.

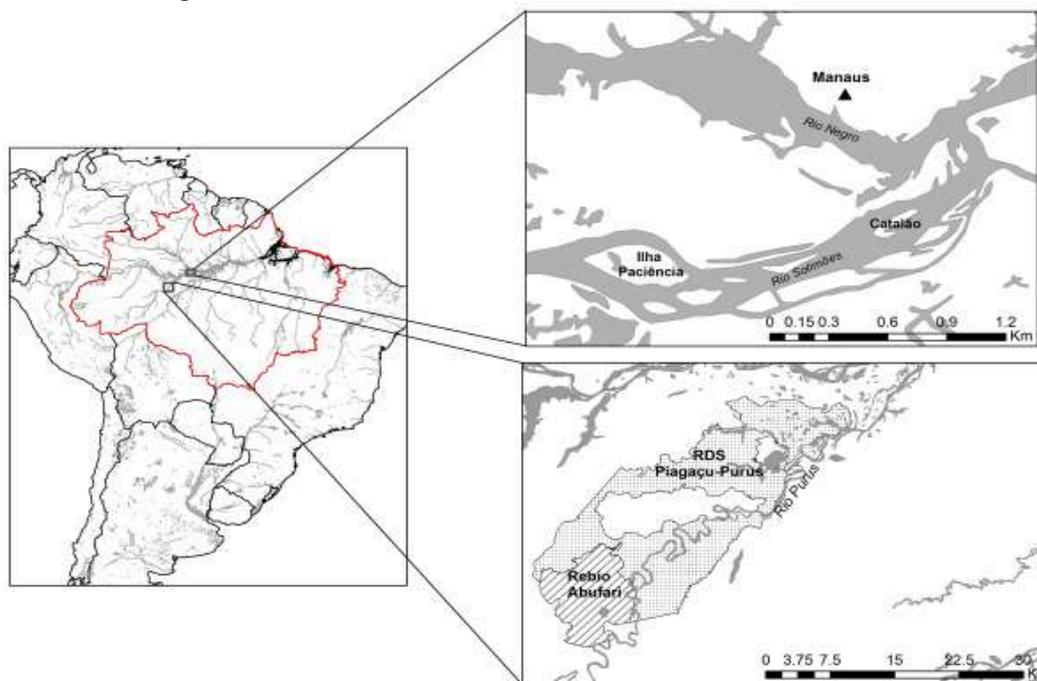
O grande volume de dados gerados por pesquisas e experimentos na área de biodiversidade é da ordem de exabyte (EB) [Li e Chen, 2014]. Para que estes dados sejam compreendidos sobre a perspectiva de observação de fenômenos, é necessário lidar com a complexidade das informações de biodiversidade, integrar os dados de fontes heterogêneas, definir os tipos de padrões e tecnologias que podem ser utilizadas como ferramentas de análise e síntese, entre outros.

Com o objetivo de identificar os impactos da variação climática na vida aquática da várzea da Amazônia Central, foi proposto um novo sítio PELD (Programa de Pesquisa Ecológica de Longa Duração): Diversidade da Várzea (DIVA). Para atender as demandas do Programa, foi desenvolvido o *e-DIVA* que propõe uma estratégia de integração das diversas bases de dados de biodiversidade geradas e relacionadas no âmbito do PELD, além de oferecer uma arquitetura computacional para gestão e análise dos dados relacionados ao PELD-DIVA.

### **1.1. PELD – DIVA: Cenário**

O PELD-DIVA (<http://peld-diva.bio.br>) busca responder questões pertinentes aos efeitos das mudanças climáticas nas entidades bióticas e abióticas das planícies alagáveis da Amazônia Central, a fim de preservar a integridade ecológica, principalmente da ictiofauna. O objetivo geral do Programa é avaliar a dinâmica temporal

de longo prazo da diversidade de assembléias de peixes em resposta a diferentes situações de conservação ambiental, esferas de governança e tipos de manejo dos ambientes aquáticos e dos recursos pesqueiros. O sítio PELD compreende o Lago Catalão e Ilha da Paciência, no baixo rio Solimões, e Reserva de Desenvolvimento Sustentável (RDS) Piagaçu-Purus e Reserva Biológica (REBIO) Abufari, no baixo rio Purus, conforme apresentado na Figura 1.



**Figura 1 - Localização das áreas que compõe o Sítio PELD-DIVA. (Fonte: cortesia PELD-CNPq 2016)**

As áreas que compõem este PELD representam unidades regionais que variam desde a região do Catalão (área fora de Unidade de Conservação - UC) e as UC cuja governança acontece em diferentes escalas administrativas. O Catalão é uma região onde o INPA mantém uma série histórica de pesquisas sobre a ictiofauna local iniciadas em 1999. É uma área sem governança direta, localizada próxima à cidade de Manaus. REBIO Abufari é uma UC federal cuja responsabilidade é do Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio). A Reserva de Desenvolvimento Sustentável (RDS PP) é uma UC estadual administrada pela Secretaria de Meio Ambiente (SEMA) tendo como cogestor o Instituto Piagaçu. A Área de Proteção Ambiental (APA) Municipal Ilha da Paciência é de responsabilidade da Secretaria de Meio Ambiente e Turismo de Iranduba, tendo os Acordos de Pesca com manejo de pirarucu reconhecido pelo Departamento Estadual de Unidades de Conservação (DEUC) na SEMA - estadual. Os estudos ictiofaunísticos tiveram início em 2006 e são desenvolvidos pela Universidade Federal do Amazonas (UFAM).

Na Ilha da Paciência e RDS PP o manejo dos recursos pesqueiros são, portanto, desenvolvidos por meio de instrumentos normativos legais reconhecidos pelos órgãos responsáveis, sendo que as categorias de manejo estabelecidas são semelhantes nessas duas regiões.

Dentre os objetivos específicos, destacam-se: determinar os fatores ambientais, em escala local e de paisagem, que influenciam a diversidade de peixes; investigar o efeito da atividade pesqueira em lagos com diferentes níveis de proteção e de manejo sobre a diversidade de peixes; identificar os efeitos parciais da pesca e das flutuações hidrológicas interanuais (e suas interações) sobre a diversidade de peixes; estudar os efeitos socioeconômicos e culturais das diferentes formas de manejo de lagos nas áreas de várzea selecionadas; investigar o nível de governança ambiental, em diferentes esferas de gestão, no tocante à diversidade de peixes nas áreas de várzea e seus efeitos nos arranjos produtivos locais; analisar os efeitos das ações de manejo, incluindo aquelas voltadas para a proteção de predadores de topo da cadeia, como o pirarucu (*Arapaima gigas*), sobre a diversidade de peixes; analisar os efeitos socioeconômicos e culturais do manejo do pirarucu em áreas de várzea na Amazônia Central; e, curadoria dos dados científicos para provimento de bases de dados em ambientes de *workflow* científico.

## **1.2. Ciência da Computação e da Informação: Contexto Tecnológico**

Dados de biodiversidade são complexos, considerando seus formatos e modelos [Mayer-Schönberger e Cukler, 2014]. Os mesmos são encontrados de forma abundante e dispersa, com natureza semi-estruturada, organizados de forma distinta, dependente do elicitador ou do especialista que o gerou, caracterizando a problemática da integração de dados de biodiversidade. 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. Os dados e conhecimentos científicos sobre biodiversidade exercem um importante papel no atendimento a importantes demandas, 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 a adoção/customização ou o desenvolvimento de sistemas de informações capazes de extrair, armazenar, gerenciar, analisar, integrar e disseminar dados das diversas fontes, além de linguagens, protocolos e mecanismos de representação que possibilitem a automação ou semi-automatização das atividades [Campos dos Santos *et al.*, 2000; Umminger e Young, 1997].

Um dos problemas que a biodiversidade tem enfrentado é o volume crescente de dados gerados independentemente por grupos individuais, que trabalham em uma mesma área do domínio, mas com semântica própria. A integração destas várias semânticas é necessária para gerar novos conhecimentos e auxiliar os processos de tomada de decisão. Alguns dos trabalhos correntes em integração de dados estão focados no tema da integração semântica, que objetiva atenuar os conflitos semânticos entre fontes de dados heterogêneas (problemas relacionados a conceitos semanticamente equivalentes ou conceitos semanticamente relacionados/não-relacionados), ao invés de projetar a estrutura da arquitetura de integração. Uma das estratégias para lidar com tais problemas é o uso de elementos integradores como as ontologias para ajudar a eliminar conflitos semânticos.

Borst (1997) define ontologia como uma especificação formal e explícita de uma conceitualização compartilhada. Nessa definição, formal significa legível por computadores; especificação explícita diz respeito a conceitos, atributos, relações, restrições e axiomas que são explicitamente definidos; compartilhado quer dizer conhecimento consensual; e conceitualização diz respeito a um modelo abstrato de algum fenômeno do mundo real.

As bio-ontologias, ontologias ligadas à biologia e ciências da vida, têm auxiliado e contribuído para grandes avanços na área, pois permitem a associação de significado aos dados gerados em experimentos, além de possibilitar a integração das várias fontes de conhecimento, uma nova abordagem para preparação de experimentos ou ainda a busca de respostas usando resultados de fontes distintas através da aquisição de novos conhecimentos. Esta tecnologia é adequada ao suporte de atividades de análise e integração de dados de qualquer natureza.

Assim, foi desenvolvida a OntoBio (Albuquerque *et al.*, 2015), uma ontologia formal de biodiversidade desenvolvida no cenário de coleções biológicas e eventos de coleta de campo do INPA, com o propósito de mitigar os problemas oriundos da interoperabilidade dos dados de biodiversidade (heterogeneidade semântica). Esta tecnologia constitui-se em uma ferramenta com funcionalidades de integração e estruturação de biodados, tendo sido modelada com base no conhecimento explícito e consensual sobre biodiversidade, e constitui um importante ponto de partida para a validação da aplicação de ontologias formais à aquisição e gestão de conhecimento de biodiversidade, bem como recurso de estruturação de conhecimento do PELD-DIVA (projeto lógico de bancos de dados de biodiversidade).

## **2. e-DIVA como Ambiente Computacional**

O e-DIVA promove mecanismos de manipulação da complexidade dos dados e das informações e conhecimento gerados no âmbito do PELD-DIVA e integra estes dados (recursos heterogêneos) permitindo inclusive a geração de novo conhecimento (através da OntoBio) e gerando parâmetros e recursos a aplicação em *workflow* científico.

Ainda, ressalta-se que este ambiente evidencia questões recorrentes de Inteligência Artificial (IA) e Banco de Dados (BD) como aquisição e elicitación do conhecimento, automatização do processo de integração do conhecimento, representação e modelagem do conhecimento e a complexidade dos dados do domínio em questão (parâmetros espaço-temporais, estrutura indefinida, multidimensionalidade e multimídia, vocabulário incógnito expresso por uma linguagem particular, grande volume de dados, existência de numerosos modelos e formatos de dados) (Albuquerque *et al.*, 2009; Elmasri e Navathe, 2016).

Desta forma, o *e-DIVA* implementa uma arquitetura computacional para análise dos dados de biodiversidade da várzea da Amazônia Central, gerados no âmbito do PELD-DIVA, utilizando a OntoBio como elemento estruturante para integração e organização das diferentes fontes de dados, possibilitando o processamento automático do conhecimento.

### **2.1 Integração de Bases de Dados Ictiológicas**

A base de dados do e-DIVA é composta da integração das bases de dados de biodiversidade do Catalão, Ilha da Paciência, Purus, Abufari e do próprio PELD-DIVA (dados coletados a partir de 2018) e contempla aspectos da biodiversidade (dinâmica temporal da diversidade de assembléias de peixes em resposta a diferentes situações de conservação ambiental) e sócio-econômicos (esferas de governança e tipos de manejo dos ambientes aquáticos e dos recursos pesqueiros). Após uma avaliação semântica dos dados contidos nas bases de biodiversidade já existentes (através da OntoBio), foi possível

verificar que as bases disponíveis eram similares, sendo os dados do PELD-DIVA mais abrangentes, o que em muito facilitou o processo de integração de dados.

## 2.2 Módulo de Biodiversidade

A descrição do domínio da aplicação do Módulo de Biodiversidade é apresentada a seguir e ilustrada na Figura 2. A fase de levantamento e análise de requisitos do Módulo de Biodiversidade foi baseada em reuniões com os pesquisadores e nos conceitos descritos na OntoBio.

Em toda excursão, são aplicados os protocolos de coleta e processamento adotados no Programa. A metodologia de campo é, além da coleta de peixes, verificar dados como: o local de coleta, data, hora, posição geográfica do local e instrumentos utilizados na pesca, como malhadeira, redinha, tarrafa, entre outros. A tarrafa, por exemplo, é um material apropriado para águas rasas, águas turvas e corredeiras. Após a coleta, os peixes são levados para estudo no INPA, onde são verificadas suas características.

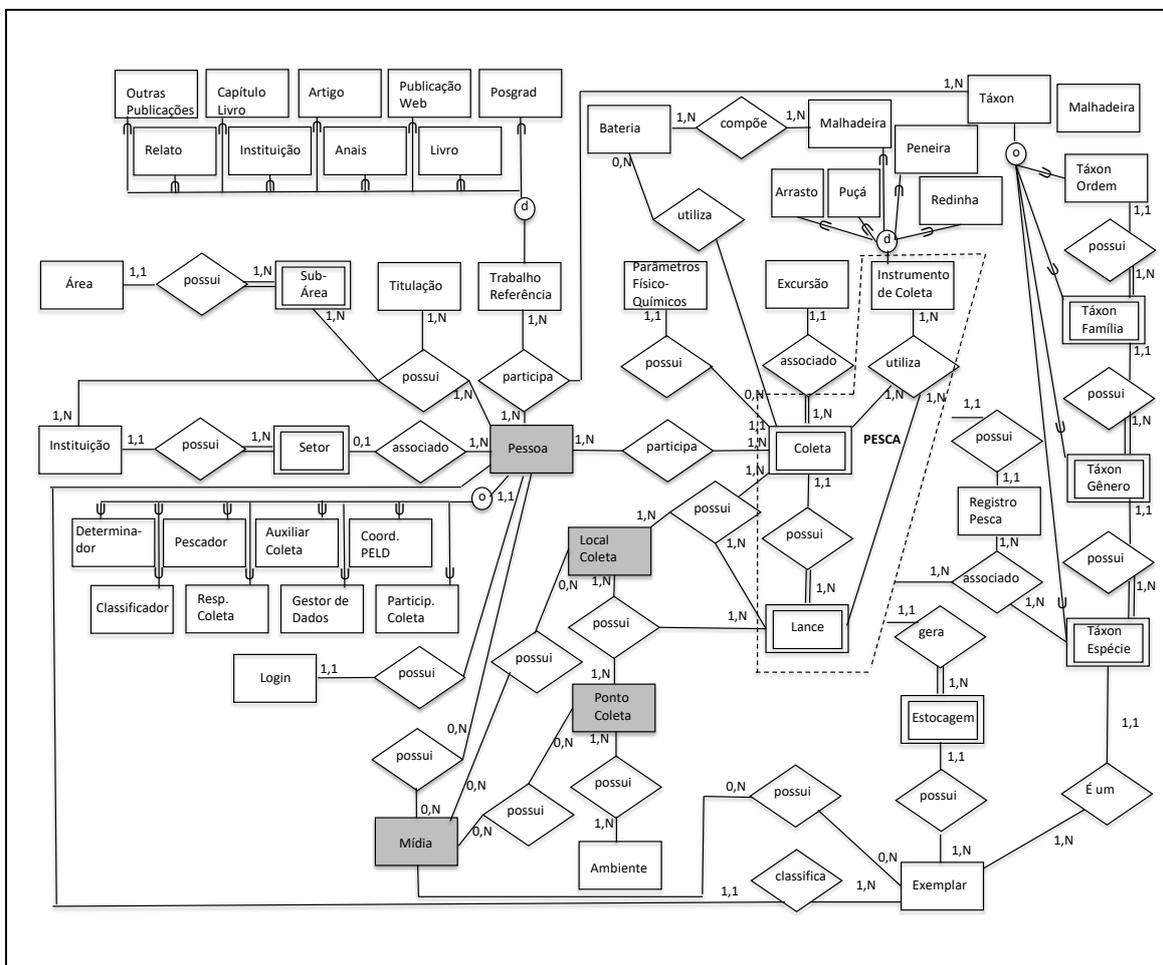


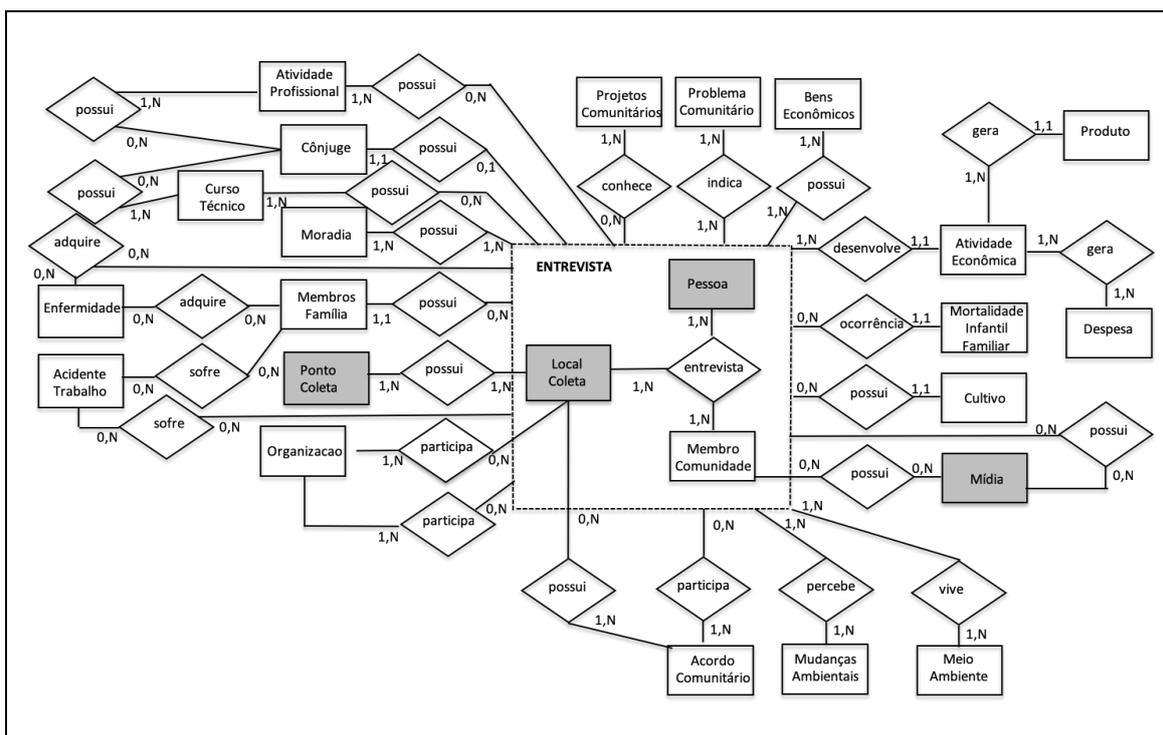
Figura 2 – Esquema Conceitual do Módulo de Biodiversidade.

Para as pesquisas realizadas, é importante armazenar dados como o ano hidrológico (diferente do ano gregoriano, pois não começa em janeiro e termina em dezembro), dependente do período hidrológico, ou seja, para completar um ano hidrológico é preciso passar por quatro períodos: enchente, seca, vazante e cheia, os quais não têm um tempo de duração pré-definido. Quanto à localização do peixe são utilizadas geo-coordenadas. A

morfometria do peixe considerada refere-se ao comprimento padrão e total. A taxonomia também é verificada (ordem, gênero, espécie, família). Podem ser registradas multimídias dos indivíduos, por exemplo, fotos e vídeos. Em laboratório são arquivadas caracterizações como gordura, peso total do peixe, peso eviscerado, peso do fígado, estômago, peso da gônada, peso da amostra da gônada. São ainda verificadas características dos peixes como um todo. Para uma melhor análise, são identificados aspectos abióticos (parâmetros físico-químicos) como: temperatura da água, o tipo e pH da água, a temperatura do ar, oxigênio dissolvido, condutividade, pressão da água e o tempo. Os dados dos peixes são correlacionados com o nível da água que ocorre todos os anos, sendo este dado obtido na Portobrás, (porto de Manaus).

### 2.3 Módulo Socioeconômico

A descrição do domínio da aplicação do Módulo Socioeconômico do PELD-DIVA é apresentada a seguir e ilustrado através da Figura 3. A fase de levantamento e análise de requisitos do PELD-DIVA, Módulo Socioeconômico foi baseada em reuniões e nos questionários elaborados, tabulados, e já então utilizados (validados) nas visitas às várias comunidades para avaliação das atividades de manejo. Uma das dificuldades constituiu-se em mapear os questionários para uma base de dados íntegra e coerente para os pesquisadores. Os questionários estão disponíveis no link <http://peld-diva.bio.br>.



**Figura 3 – Esquema Conceitual do Módulo Socioeconômico.**

Este módulo registra dados socioeconômicos dos membros de uma comunidade. As comunidades encontram-se localizadas nas áreas de atuação do Programa. Os registros são feitos em entrevistas, permitindo que um membro de comunidade tenha várias entrevistas referentes a diferentes espaço-temporais. Esta peculiaridade é necessária para que se possa avaliar os impactos das variações climáticas e do manejo da pesca e de atividades econômicas ao longo do tempo em um indivíduo, grupo familiar ou comunidade. Nas entrevistas, informações sobre a organização familiar, formação

acadêmica e profissional dos membros da família, moradia, atividades profissionais, bens econômicos, atividades econômicas e produtos gerados, acordos comunitários dos quais a comunidade e o entrevistado participam e percepção sobre variações climáticas e ambientais são registradas.

A integração do esquema de dados de biodiversidade com o esquema de dados socioeconômico se dá através das relações *Local\_Coleta*, *Ponto\_Coleta*, *Pessoa* e *Midia*. Estas entidades e relações associadas, encontram-se nos esquemas (sombreadas de cinza) apresentados nas Figuras 2 e 3.

#### **4. Aplicação da Infraestrutura do *e-DIVA* para Uso em Workflow Científico**

Biocientistas são solicitados a fornecerem evidências de mudanças na biodiversidade, identificar fatores que causam seu declínio, além de prever o impacto e sugerir formas de combater a perda da mesma. A alteração na distribuição das espécies, a natureza dinâmica dos ecossistemas e o aumento dos riscos de extinção, muitos dos quais surgem das atividades antropogênicas, têm impacto em importantes áreas de interesse social (saúde e bem-estar humanos, segurança alimentar, serviços ecossistêmicos, bioeconomia, etc.). Assim, cada vez mais, pede-se aos cientistas que forneçam apoio decisório baseado no conhecimento para o manejo da biodiversidade e do uso da terra em múltiplas escalas e granularidades, desde genomas até espécies e ecossistemas, para prevenir ou pelo menos mitigar tais perdas. A geração de evidências suficientes e o fornecimento de apoio à decisão dependem cada vez mais de grandes conjuntos de dados mantidos em formatos digitais e da aplicação de recursos computacionais substanciais e capacidade de analisar, modelar, simular e prever o uso desses dados. A computação tem se tornado parte integrante e imprescindível para o sucesso na realização de pesquisas científicas das mais variadas áreas, conhecido hoje como *e-science*.

As atividades de *e-science* estão acompanhadas de desafios que vão desde a capacidade de consumo e análise de grande volume de dados até o desenvolvimento de novas ferramentas para atender novas demandas. Dentre os desafios, enfatiza-se: como entender e organizar os recursos computacionais; como compartilhar e reutilizar experimentos bem-sucedidos (ferramentas e dados), e como prover interoperabilidade entre dados e ferramentas de diferentes locais e sendo utilizados por usuários com perfis distintos.

Novas práticas de pesquisa que exploram dados e análises baseadas em dados precisam de infraestrutura que permita confiabilidade, robustez, repetibilidade, proveniência e reprodutibilidade para investigações científicas grandes e complexas. A infraestrutura *e-DIVA* oferece uma arquitetura computacional adequada para estes propósitos e constitui recurso adequado para aplicação em ambiente de *workflow* científico, mitigando questões de metadados, interoperabilidade semântica (OntoBio), integração e proveniência de dados, permitindo delineamento de experimentos como parâmetro para o projeto a ser executado no *workflow*.

A Figura 4 apresenta a infraestrutura projetada para o *e-DIVA*, na qual as bases de dados de biodiversidade referentes aos sítios de pesquisa que compõem o PELD-DIVA podem ser gerenciadas individualmente pelos usuários dos projetos do INPA e UFAM. Os esquemas lógicos de dados são integrados com o suporte da OntoBio e do conhecimento elicitado via formulários sócio-econômicos. O Esquema Lógico Integrado do BD é mapeado para o esquema físico do BD que é povoado com os dados das bases de dados

que compõem o PELD-DIVA. Adicionalmente, ilustra-se o ambiente de workflow científico que durante seu processo de execução tem como fonte dado, o *e-DIVA* BD. Uma vez que a fonte de dado que instancia o workflow (*e-DIVA*) atende o processo de validação (proveniência dos dados), não é necessário informações sobre a proveniência dos mesmos.

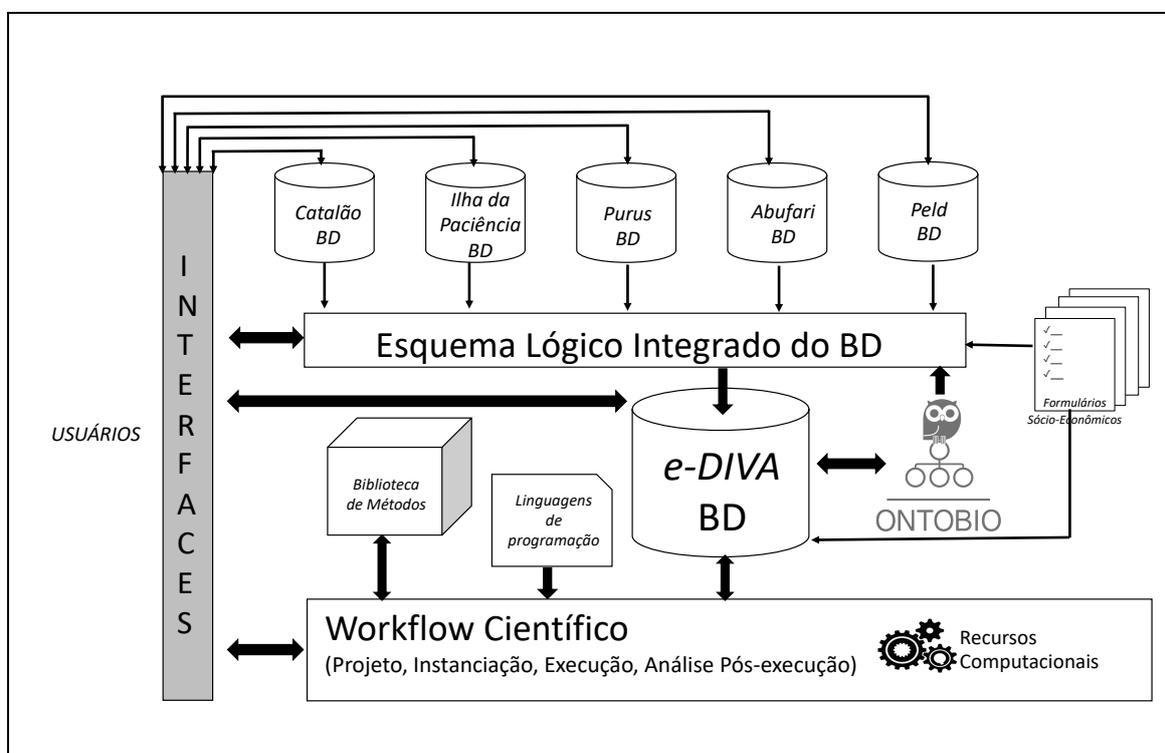


Figura 4 – Infraestrutura do e-DIVA.

## 5. Considerações Finais

Pesquisar a biodiversidade e dar suporte a políticas públicas depende cada vez mais de grandes conjuntos de dados mantidos eletronicamente e de capacidade computacional de analisar, modelar, simular e prever o uso desses dados. No entanto, os recursos de dados que estão fisicamente distribuídos e a falta de experiência com ferramentas analíticas avançadas constitui desafios para os cientistas contemporâneos. O *e-DIVA* foi projetado para atender estas demandas e permite avaliar a dinâmica temporal de longo prazo da diversidade de populações de peixes da várzea da Amazônia Central em resposta a diferentes situações de conservação ambiental, esferas de governança e tipos de manejo dos ambientes aquáticos e dos recursos pesqueiros.

Uma das questões a ser considerada é o uso exclusivo de formulários socioeconômicos utilizados nas entrevistas estruturadas, pois tendem a limitar a abrangência do domínio e podem comprometer o modelo da realidade (Esquema Lógico Integrado). A adoção de entrevistas semiestruturadas no processo de elicitação dos requisitos constitui um ferramental mais adequado e requer um número maior de especialistas para análise dos dados coletados.

## 6. Agradecimentos

Ao LIS/ INPA, a FAPEAM (Fundação de Amparo à Pesquisa do Estado do Amazonas), Edital N°005/2018 – FIXAM e ao CNPq (Conselho Nacional de Desenvolvimento

Científico e Tecnológico) programa 5731 - CNPq/CAPES/FAPs/BC-Fundo Newton - Pesquisa Ecológica de Longa Duração, processo 88887.144296/2017-00 por financiarem parcialmente esta pesquisa.

## Referências

- Albuquerque, A. C. F.; Campos dos Santos, J.; De Magalhães Netto, J. (2009) "Modeling Complex Domain Ontology Based on the Unified Foundational Ontology". Extended Proceedings of the 4th. Latin American Conference on Computer Human Interaction (CLIHIC 2009) in conjunction with the 7th. Latin American Web Congress (LAWEB 2009) / Alberto L. Morán, comp. Ensenada, Baja Calif. : Universidad Autónoma de Baja California. November 9-11th, Mérida, Yucatán, México. ISBN: 978-607-7753-32-2.
- Albuquerque, A. C. F.; Santos, J.L.C.; Castro Jr, A.N. (2015) "OntoBio: A Biodiversity Domain Ontology for Amazonian". Proceedings of 48th Hawaii International Conference on System Sciences. Kauai, Hawaii, January 5th – 8th. ISBN: 978-1-4799-7367-5
- Bennemann, S. T.; Shibatta, O. A.; Garavello, J. C. (2000) "Peixes da Bacia do Rio Tibagi: Uma Abordagem Ecológica". Londrina: EDUEL.
- Borst, W. N. (1997) "Construction of Engineering Ontologies for Knowledge Sharing and Reuse". PhD thesis, Univ. of Twente. CTIT Ph.D.-thesis series No. 97-14. ISBN 90-365-0988-2.
- Campos dos Santos, J.; de By, R.A. ; Magalhães, C. (2000) "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.
- Elmasri, R.; Navathe, S. (2016) "Fundamentals of Database Systems". Pearson, 7th. Edition. ISBN-10: 0133970779. ISBN-13: 978-0133970777
- Junk, W. J.; Soares, M. G. M.; Bayley, P. B. (2007) "Fresh Water Fishes of the Amazon River Basin: Their Biodiversity, Fisheries and Habitats". Journal Aquatic Ecosystem Health and Management, Vol. 10. Issue 2.
- Li, Y.; Chen, L. (2014) "Big biological Data Challenges and Opportunities". In Genomics Proteomics Bioinformatics 12(2014): 187-189. DOI: 10.1016/j.gpb.2014.10.001
- Mayer-Schönberger, V.; Cukler, K. (2014) "Big Data: A Revolution That Will Transform How We Live, Work, and Think". Eamon Dolan/Mariner Books; Reprint Edition (March 4, 2014). ISBN-10: 0544227751.
- Oliveira Júnior, A. (2015) "Amazônia: Paisagem e Região na Obra de Eidorfe Moreira". Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas, v. 10, n. 3, p. 569-581. DOI: 10.1590/1981-81222015000300003
- Parmesan, C. (2006) "Ecological and Evolutionary Responses to Recent Climate Change". Annual Review of Ecology, Evolution and Systematics 37: 637-669.
- Umminger, B.; Young, S. (1997) "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.
- Val, A. L. (2011) "As Mudanças Climáticas e os Peixes de Água Doce". Anais da 63a. Reunião Annual da SBPC, Goiânia, GO.