



Arquitetura de Serviços Web para Cooperação em Modelagem de Biodiversidade

Proposta de Tese de Doutorado em
Computação Aplicada

Karla Donato Fook

Orientadores:

Dr. Antônio Miguel V. Monteiro

Dr. Gilberto Câmara



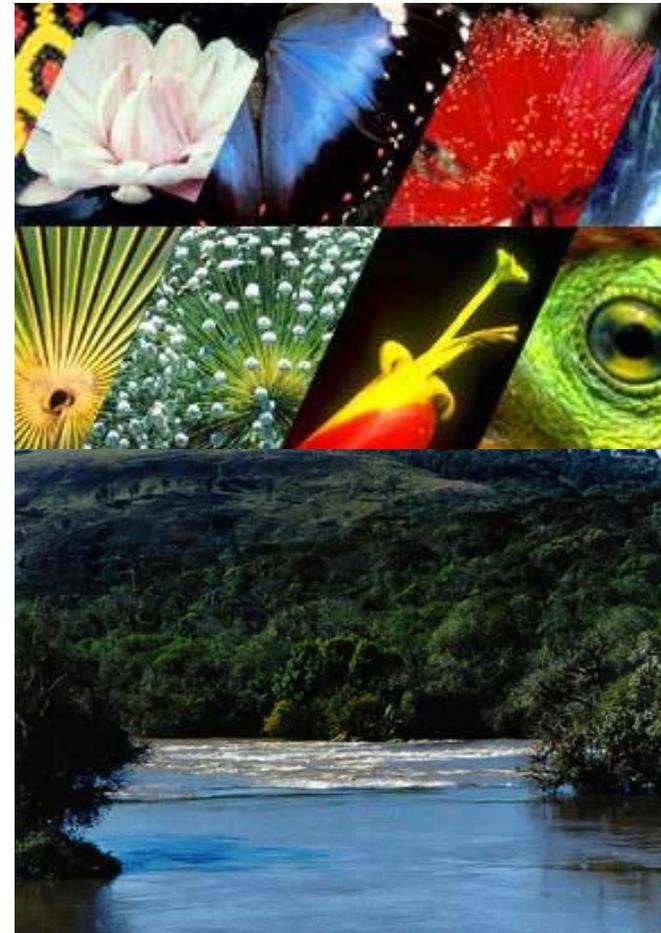
Sumário

- Introdução
- Revisão Bibliográfica
- WBCMS – *Web Biodiversity Collaborative Modelling Service*
- Considerações Finais



Introdução

- Os recursos biológicos da biodiversidade dão suporte a funções e setores essenciais como (Hobern, 2004)
 - Alimentação e Agricultura
 - Indústria Farmacêutica
 - Medicina
 - Tratamento do desperdício dos referidos recursos
- A proteção da biodiversidade é uma das questões mais urgentes e importantes da atualidade





Introdução – Motivação

- Pesquisadores que trabalham com a biodiversidade tentam prover subsídios para uso consciente dos recursos naturais
- Por que modelar a biodiversidade?
 - Recuperação ambiental
 - Conservação de espécies
 - Análise e solução de problemas como
 - Aplicações de previsão de distribuição de espécies
 - Impacto de mudanças climáticas
 - Problemas relacionados com a expansão de espécies invasoras



Introdução – Motivação

Evitar a degradação ambiental e a extinção de espécies



Dryocopus galeatus



Ilex aquifolium



Pinus sylvestris



Phoenicopterus ruber

www.biodiversitas.org.br/f_ameaca/index_lista.htm





Introdução – Motivação

- Cientistas que trabalham com a informação de biodiversidade usam recursos como
 - Fontes de dados
 - Análises estatísticas
 - Ferramentas para modelagem
 - Softwares para apresentação e visualização
- Estes recursos podem estar em várias plataformas computacionais locais e remotas (White, 2004)



Introdução – Motivação

- A colaboração entre pesquisadores não está somente na troca de dados, mas também na interação entre os modelos científicos e suas implementações, assim como na agregação de programas e de resultados de experimentos (Osthoff *et al.*, 2004)
- O conhecimento adquirido por pesquisadores pode ser um recurso a ser compartilhado
- A colaboração, a partir deste conhecimento, pode auxiliar pesquisadores
 - No avanço de seus estudos
 - Na aplicação de conhecimentos consolidados para solucionar novos problemas
 - Na aquisição de novos conhecimentos





Introdução – Questões

- Qual a melhor forma de suportar aplicações distribuídas sobre dados espaciais para modelagem de biodiversidade?
- Como prover colaboração numa rede para modelagem da biodiversidade?



Introdução – Hipóteses

- A cooperação entre pesquisadores de biodiversidade é necessária para uma boa modelagem da distribuição de espécies
- Ambientes computacionais adequados podem promover a cooperação através da integração da *semântica* de dados, modelos e resultados
- O arcabouço conceitual de serviços *web* fornece uma base para o desenvolvimento de ambientes cooperativos para modelagem de biodiversidade



Introdução – Objetivos

- Geral
 - Propor e desenvolver um ambiente cooperativo para a modelagem de biodiversidade. Este ambiente deve dar suporte aos pesquisadores que colaboram em ambientes distribuídos
- Específico
 - Oferecer um ambiente no qual dois ou mais usuários possam colaborar disponibilizando resultados de modelagem, independentes da aplicação e da plataforma



Introdução – Metodologia

- Buscar tecnologias e esforços existentes na literatura
- Identificar elementos básicos que permitam a construção do ambiente
- Avaliar ambiente proposto em um estudo de caso (prova de conceitos)



Revisão Bibliográfica



Informática da Biodiversidade

- Um desafio é a construção de uma infra-estrutura computacional para permitir acesso, visualização, interpretação e análise de informações da biodiversidade
(Canhos et al., 2004)
- Projetos ligados à biodiversidade e *e-Science* se destacam por apoiar as comunidades científicas através de cooperação de recursos
- Uma infra-estrutura computacional que promova a integração de recursos e a cooperação entre seus usuários requer a utilização conjunta de tecnologias
 - *GIS*
 - *Web Services*
 - Arquiteturas distribuídas





GIS

- A tecnologia GIS permite o armazenamento e a apresentação de dados espaço-temporais, sendo aplicada

- Administração de recursos
- Telecomunicações
- Planejamento urbano e regional
- Em todas as ciências que tratam da superfície da Terra

(Câmara *et al.*, 2003; Goodchild, 2000)

- GIS deixam de ser sistemas isolados e *standalone*, trabalhando em arquiteturas cliente-servidor para ser pequenas aplicações baseadas em *Web*

(Newcomer, 2002; Curbera *et. Al*, 2002; Panatkool and Laoveerakul, 2002)





Sistemas *Web*

- A concepção de sistemas de informação tem sido aplicada a uma arquitetura distribuída na internet, promovendo os sistemas *web*
- Características principais
 - Heterogeneidade dos dados
 - Modo como os dados são publicados e disponibilizados ao público
 - Várias interfaces e capacidades de consulta oferecidas pelas diferentes fontes

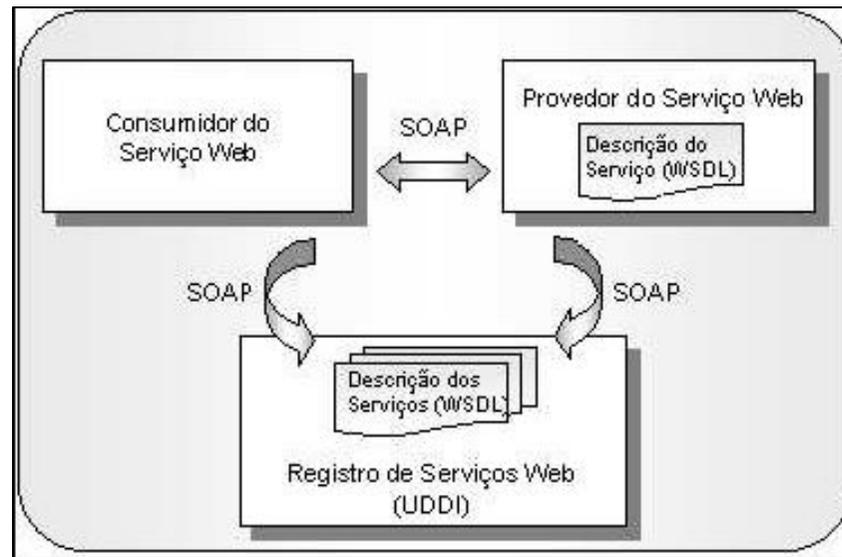
(Hernandez, 2004)





Web Services

- Executam uma tarefa específica ou um conjunto de tarefas
- São descritos em termos de arquitetura orientada a serviço
(Newcomer, 2002; Curbera et. Al, 2002; Panatkool and Laoveerakul, 2002)
- Habilitam computação distribuída baseada em componente e suporta esforços de integração de sistemas de informação
(Haines, 2004; OGC, 2004)



Arquitetura Web Services (Marquezan, 2004)





GIS Web

- A tecnologia *Web Services* permite que desenvolvedores construam componentes GIS independentes, que podem ser executados em qualquer plataforma usando qualquer linguagem (Panatkool *et al.*, 2002)
- Uma infra-estrutura computacional será proposta sob uma arquitetura distribuída
- Os tipos de processos distribuídos abordados são clientes, servidores e *peers* (Tarouco, 2005)
 - Os processos cliente e servidor são distintos e interagem entre si
 - Os processos *peer* são idênticos um ao outro, interagindo de forma cooperativa para realizar uma tarefa



Sistemas Distribuídos

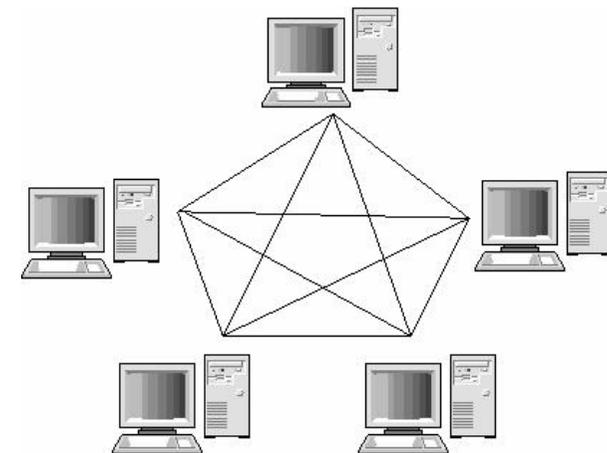
- Sistemas Cliente-Servidor
 - A computação depende de uma organização central ou hierárquica,
 - Os integrantes da rede não possuem o mesmo poder computacional em servir e utilizar recursos

- Sistemas *Peer-To-Peer*
 - Formados por nós que possuem o mesmo poder computacional, podendo assumir papel de cliente ou de servidor

 - O aumento da complexidade e sofisticação das aplicações, e a necessidade crescente para interoperabilidade propiciou uma forte convergência de interesse entre *peer-to-peer* e grade computacional (Foster *et al.*, 2001)



Sistema cliente-servidor



Sistema peer-to-peer



Sistemas Distribuídos

- Grade Computacional
 - Está interessada no “compartilhamento coordenado de recursos e resolução de problemas em organizações virtuais dinâmicas e multi-institucionais”
- Organização Virtual (VO)
 - Conjunto de indivíduos e/ou instituições definido por regras que estipulam o compartilhamento controlado
- Aplicações
 - Bioinformática
 - Geoinformática
 - Projetos colaborativos





Trabalhos Relacionados

- Arquiteturas e abordagens para encadeamento de GI *Services* na *Web*
 - (Alameh, 2001), (Tsou and Battenfield, 2002), (Bernard et. al, 2003)
- Computação em Grade no âmbito GIS *Web*
 - (Panatkool and Laoveerakul, 2002), (Pinto *et al.*, 2003), (Di et al. 2003), (Zhao et al., 2004), (Jones et al., 2003), (Cirne at al., 2003)
- Open Source Software e Cooperação
 - (Anderson et al., 2003), (Pinto *et al.*, 2003)



GIS Web e Informática da Biodiversidade

GIS Web / Informática da Biodiversidade		
Características	Desafios	Desafios abordados na Proposta
<ul style="list-style-type: none">• Grande volume de dados espaciais• Necessidade de Protocolos e Padrões• Heterogeneidade dos dados• Modo como os dados são publicados e disponibilizados ao público• Esforços colaborativos para disponibilização de dados na	<ul style="list-style-type: none">• Geoprocessamento• Geosimulação• Semântica da geoinformação• Encadeamento de serviços / <i>Workflow</i>• Computação em Grade• Integração nas infra-estruturas de TI	<ul style="list-style-type: none">• Encadeamento de serviços / <i>Workflow</i>• Integração nas infra-estruturas de TI• Cooperação

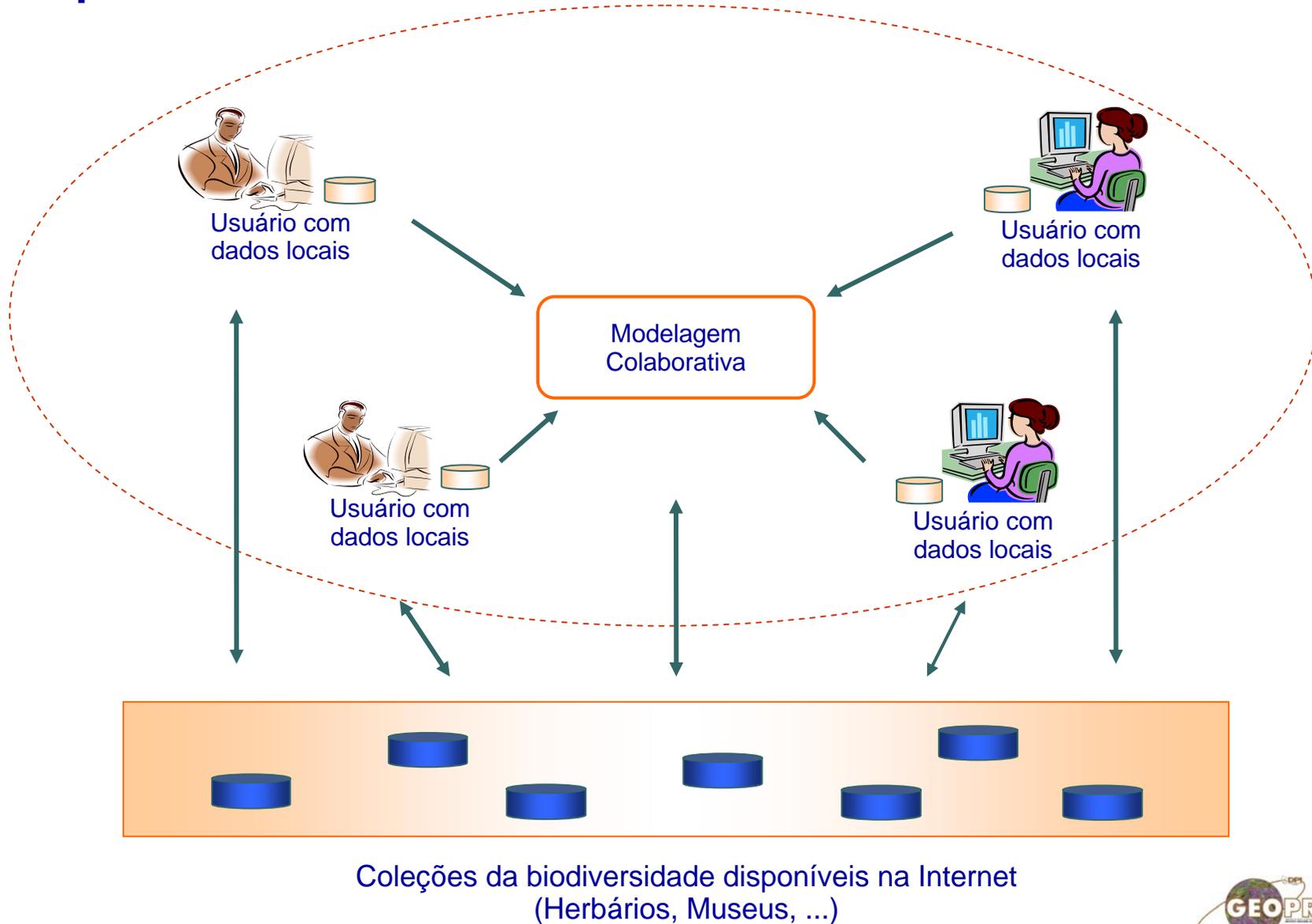
Web



Arquitetura de Serviços *Web* para Cooperação em Modelagem de Biodiversidade



Arquitetura de Serviços Web para Cooperação em Modelagem de Biodiversidade





Arquitetura de Serviços *Web* para Cooperação em Modelagem de Biodiversidade

- Além de dados e serviços, o conhecimento é um recurso presente que suporta descobertas e desenvolvimento contínuo de novos estudos e tecnologias
- O conhecimento obtido ao término de uma modelagem pode ser compartilhado através dos resultados da modelagem, aqui chamados de instâncias de modelos
- As instâncias de modelos conterão informações conceituais do modelo, além de suas condições de geração



Arquitetura de Serviços *Web* para Cooperação em Modelagem de Biodiversidade

- Considerar 3 tipos de cooperação em um ambiente de Modelagem de Biodiversidade:
 1. Instituições e/ou usuários cooperam com dados
 1. Instituições e/ou usuários cooperam com serviços
 1. Usuários cooperam com instâncias de modelos



Arquitetura de Serviços *Web* para Cooperação em Modelagem de Biodiversidade

- Um serviço de modelagem cooperativa de biodiversidade, baseado em um catálogo de instâncias de modelos, pode ajudar a pesquisadores a encontrar respostas para questões como:
 - Qual a espécie modelada?
 - De onde vieram os dados?
 - Quais as variáveis ambientais usadas?
 - Qual o algoritmo usado?
 - Se eu tiver um problema, como posso buscar resultados semelhantes?



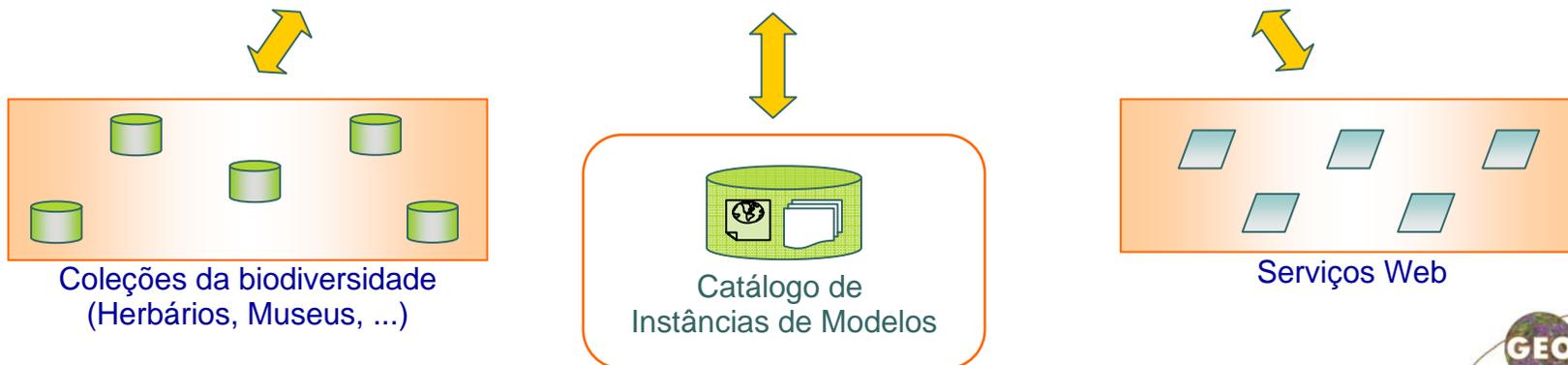
Arquitetura de Serviços Web para Cooperação em Modelagem de Biodiversidade



WBCMS

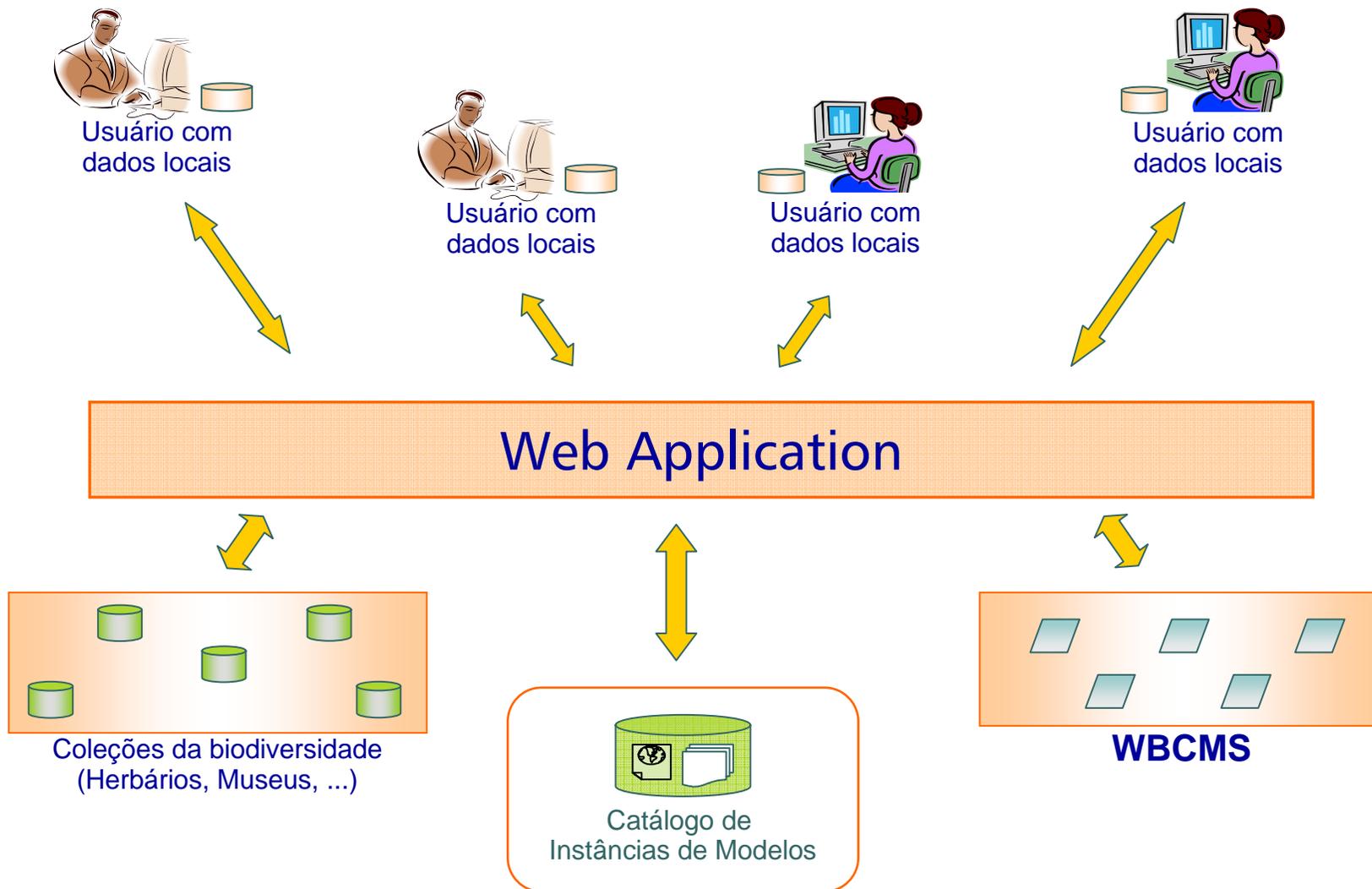
Componente de Catalogação

Componente de Acesso





Arquitetura de Serviços Web para Cooperação em Modelagem de Biodiversidade





WBCMS – *WEB Biodiversity Collaborative Modelling Service*

- Serviço *Web* que será projetado e desenvolvido para dar suporte à modelagem cooperativa de biodiversidade
- Constrói catálogos de instâncias de modelos e permite a sua recuperação

Modelagem Colaborativa de Biodiversidade

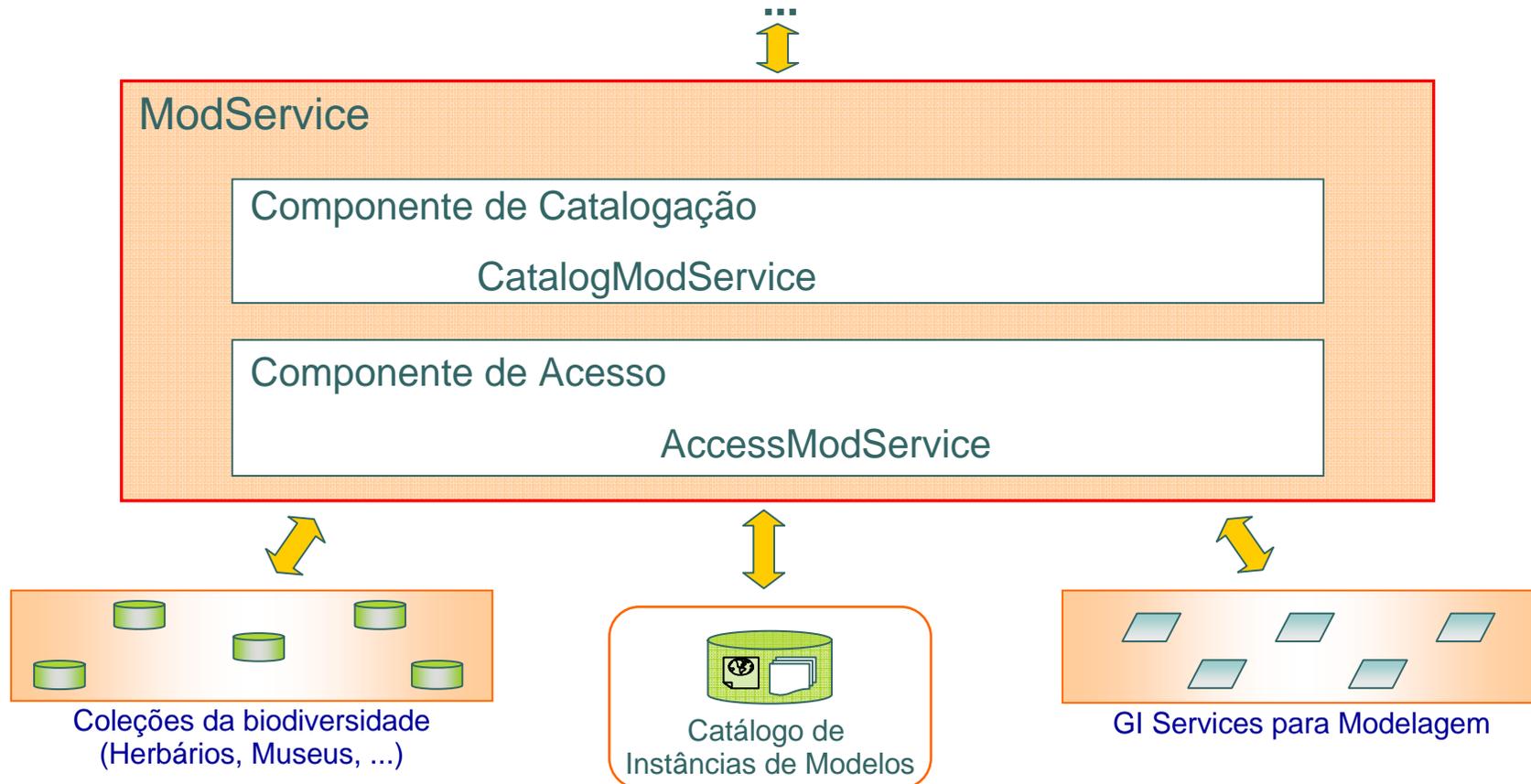
Catálogo de
Instâncias de Modelos

Acesso ao Catálogo de
Instâncias de Modelos





WBCMS – *WEB Biodiversity Collaborative Modelling Service*



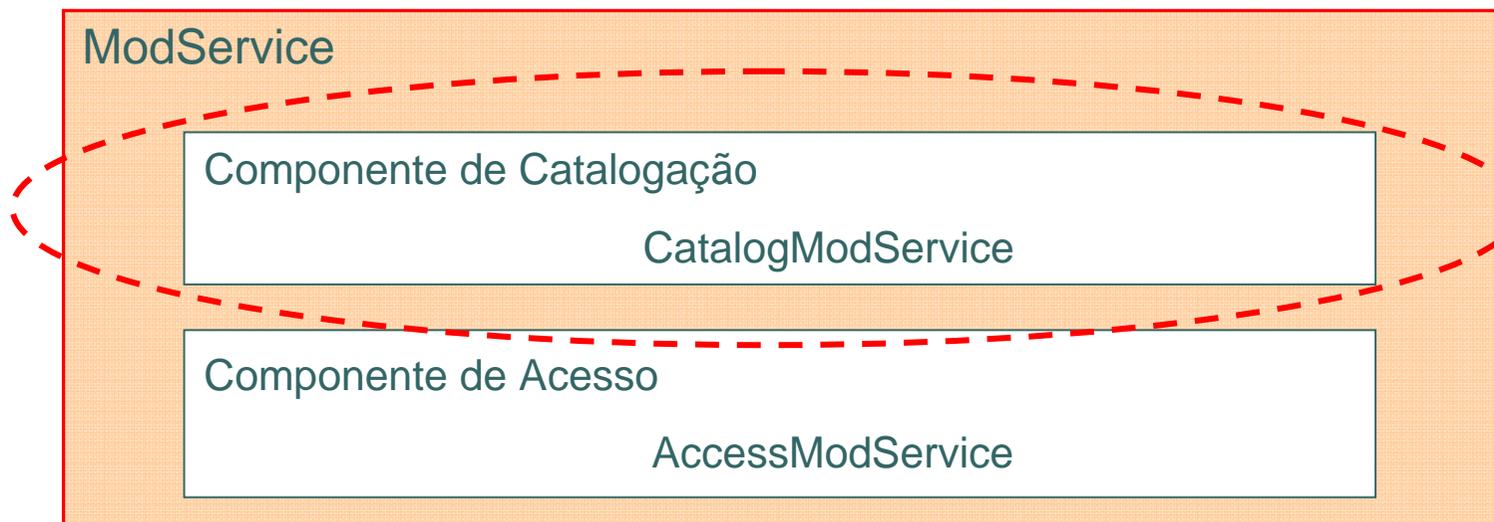
- **ModService**

- Promove Integração de Dados Espaciais e Serviços de Modelagem
- Contempla Serviços de Dados e Serviços de Processos
- Compatível com W3C e OGC





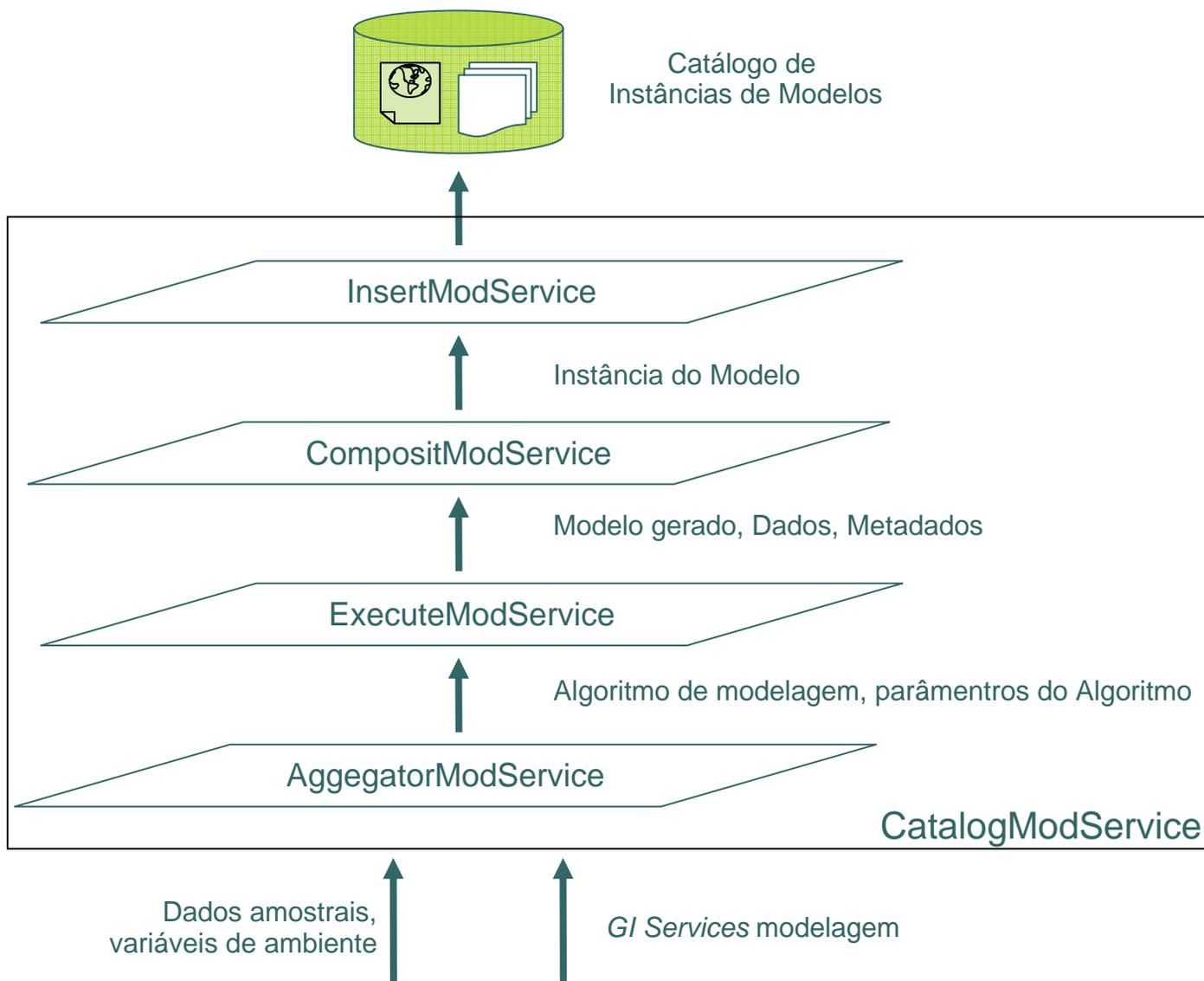
WBCMS – *WEB Biodiversity Collaborative Modelling Service / Catalogação*



- **CatalogModService**
 - Consiste de um conjunto de *GI Services* que
 - recebe os dados de entrada
 - executa algoritmo de modelagem
 - cataloga a instância do modelo

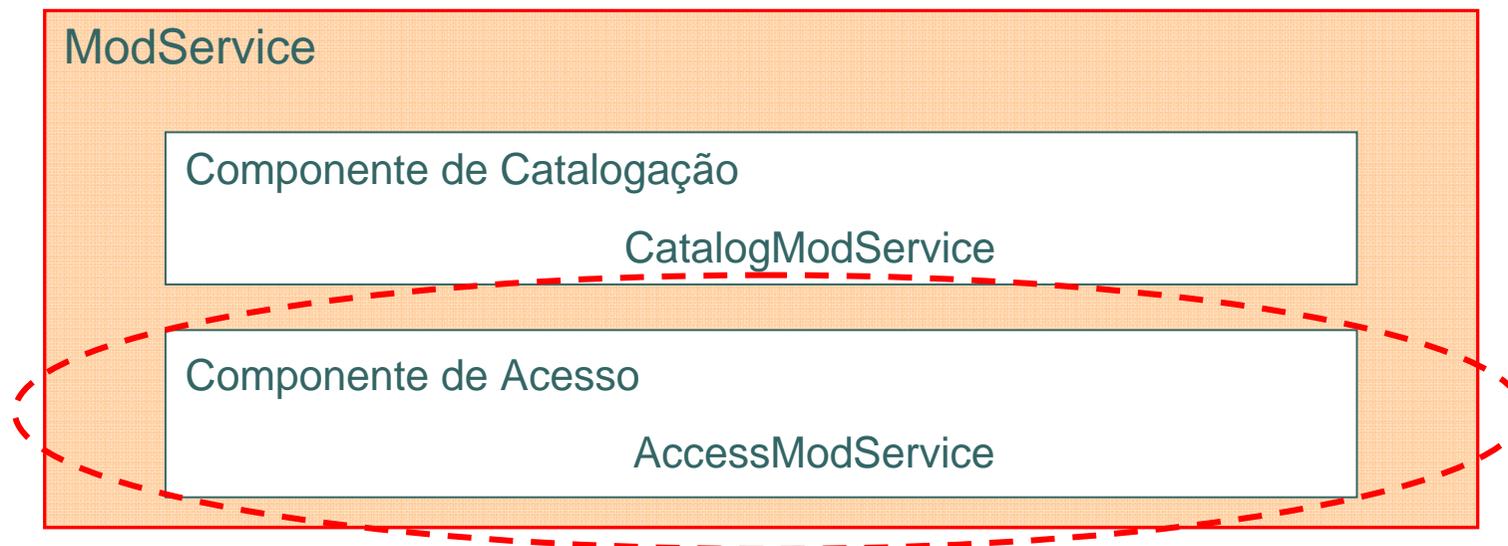


WBCMS – *WEB Biodiversity Collaborative Modelling Service / Catalogação*





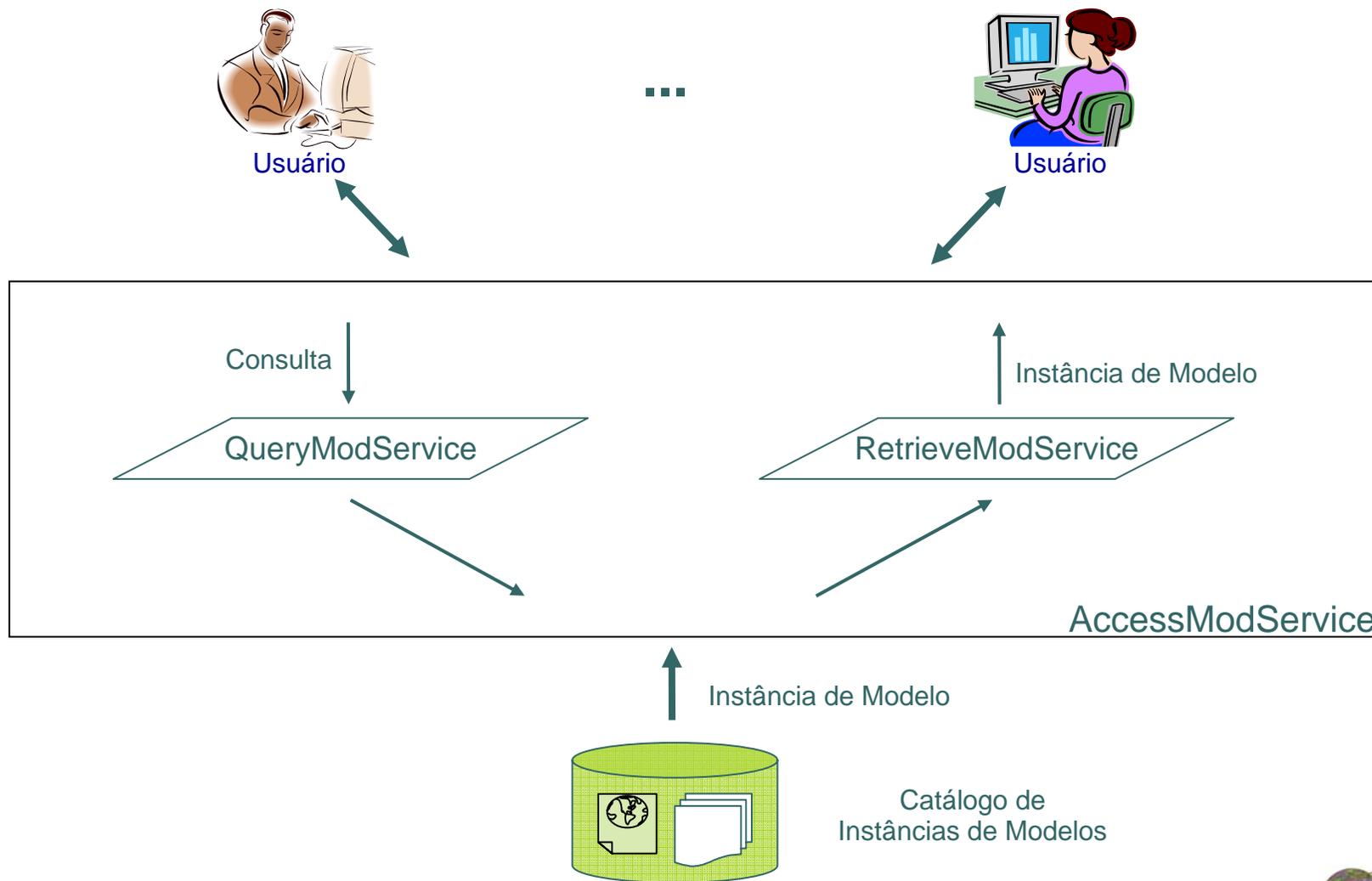
WBCMS – *WEB Biodiversity Collaborative Modelling Service* / Acesso ao Catálogo



- AccessModService
 - Consiste de um conjunto de *GI Services* que permite o acesso ao catálogo de instâncias de modelos



WBCMS – *WEB Biodiversity Collaborative Modelling Service* / Acesso ao Catálogo





Estudo de Caso



Projeto openModeller

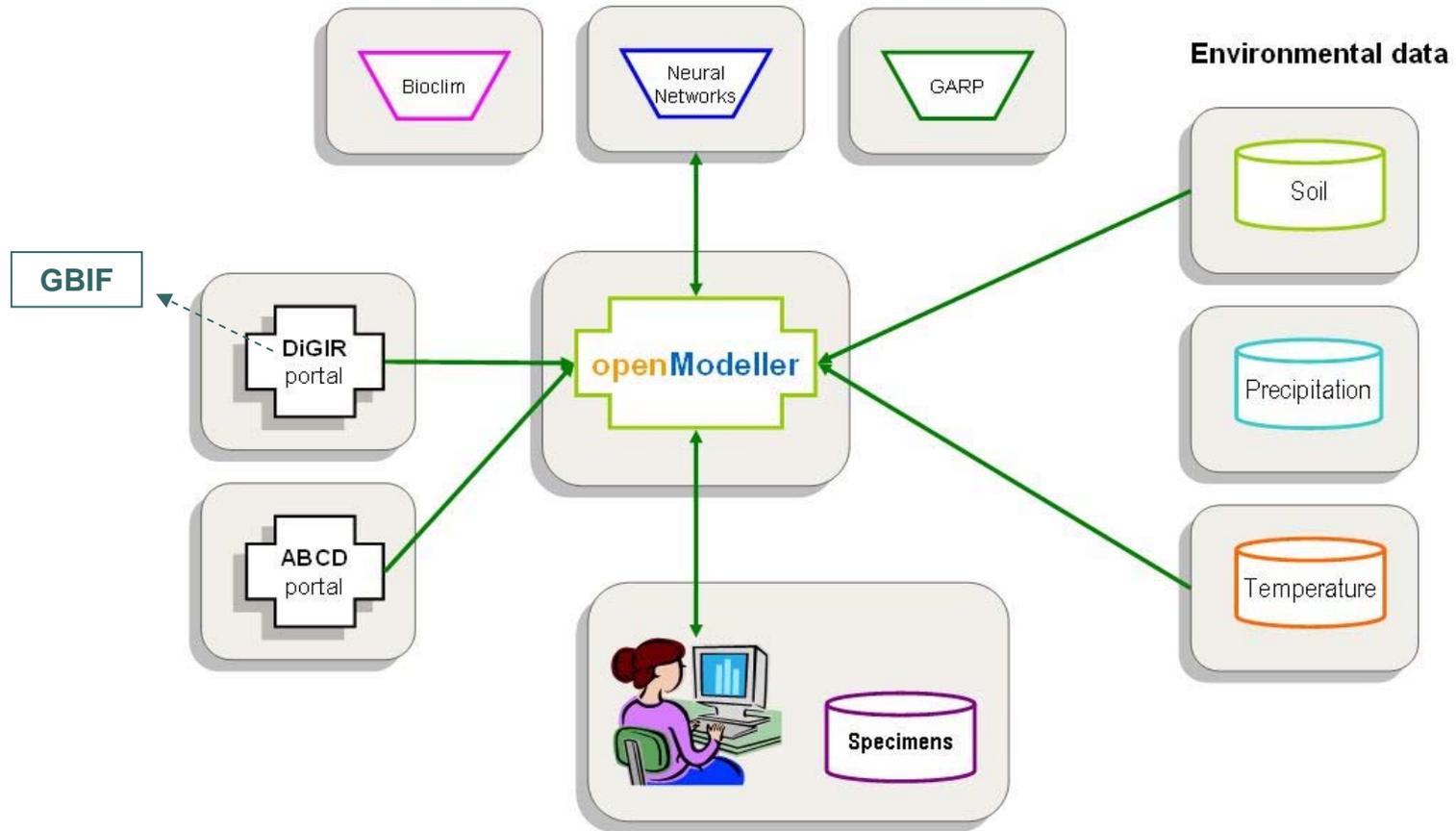
- Prover um ambiente computacional para modelagem de distribuição de espécies
- Utiliza o portal GBIF (*Global Biodiversity Information Facility*) para a busca dos dados das instituições catalogadas na rede
- Parte do projeto *speciesLink*
 - Sistema Distribuído de Informação que integra em tempo real
 - dados primários de coleções científicas do Estado de São Paulo
 - dados de observação do programa Biota/FAPESP
 - dados de acervos de algumas coleções do exterior
- O uso do openModeller como um componente de uma Grade foi proposto em 2005, no *BiodiversityWorld Grid Workshop* (Giovanni, 2005)





openModeller

Modelling algorithms



Adaptado de (Muñoz, 2004) 

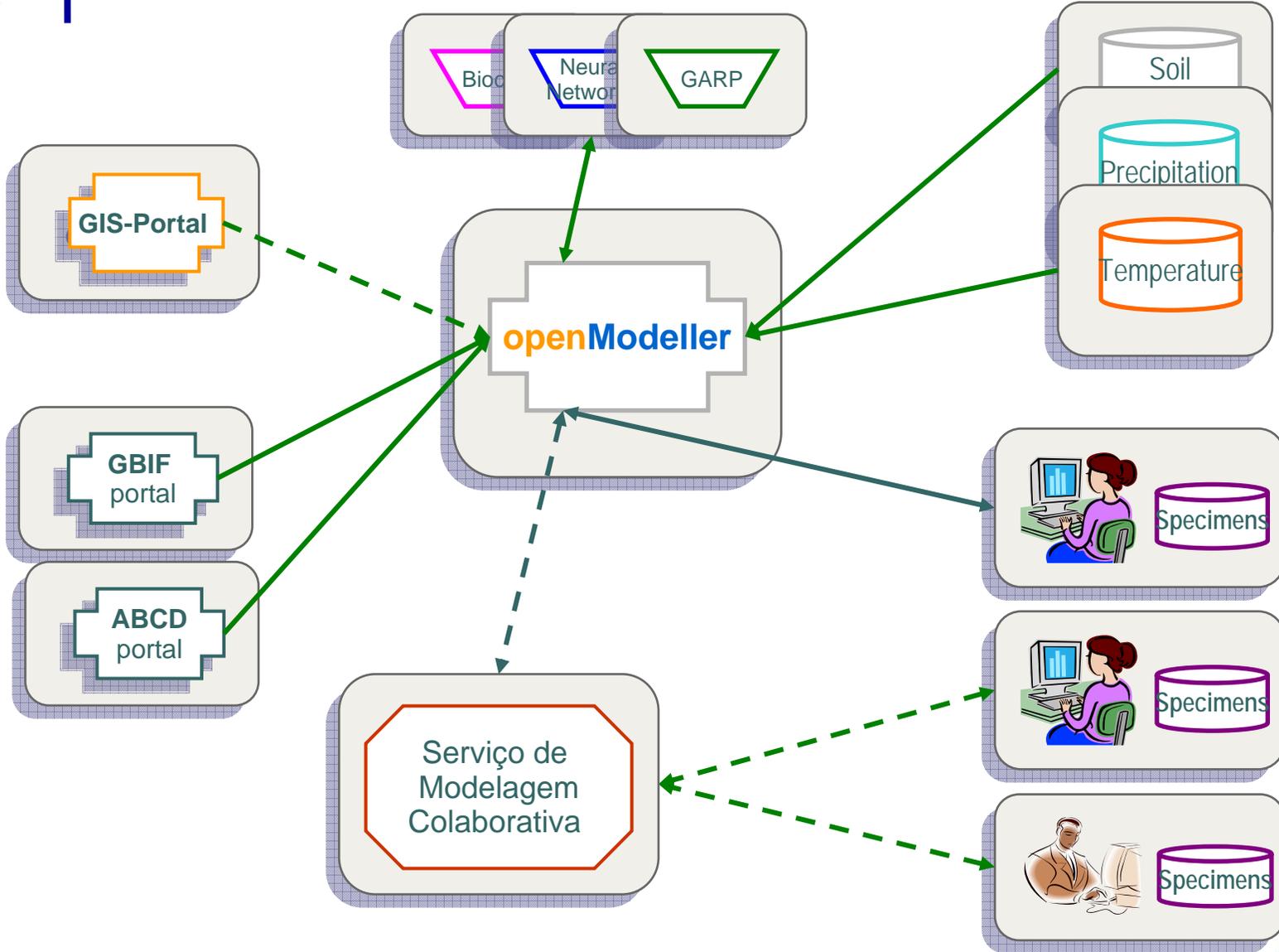




openModeller

Modelling algorithms

Environmental data

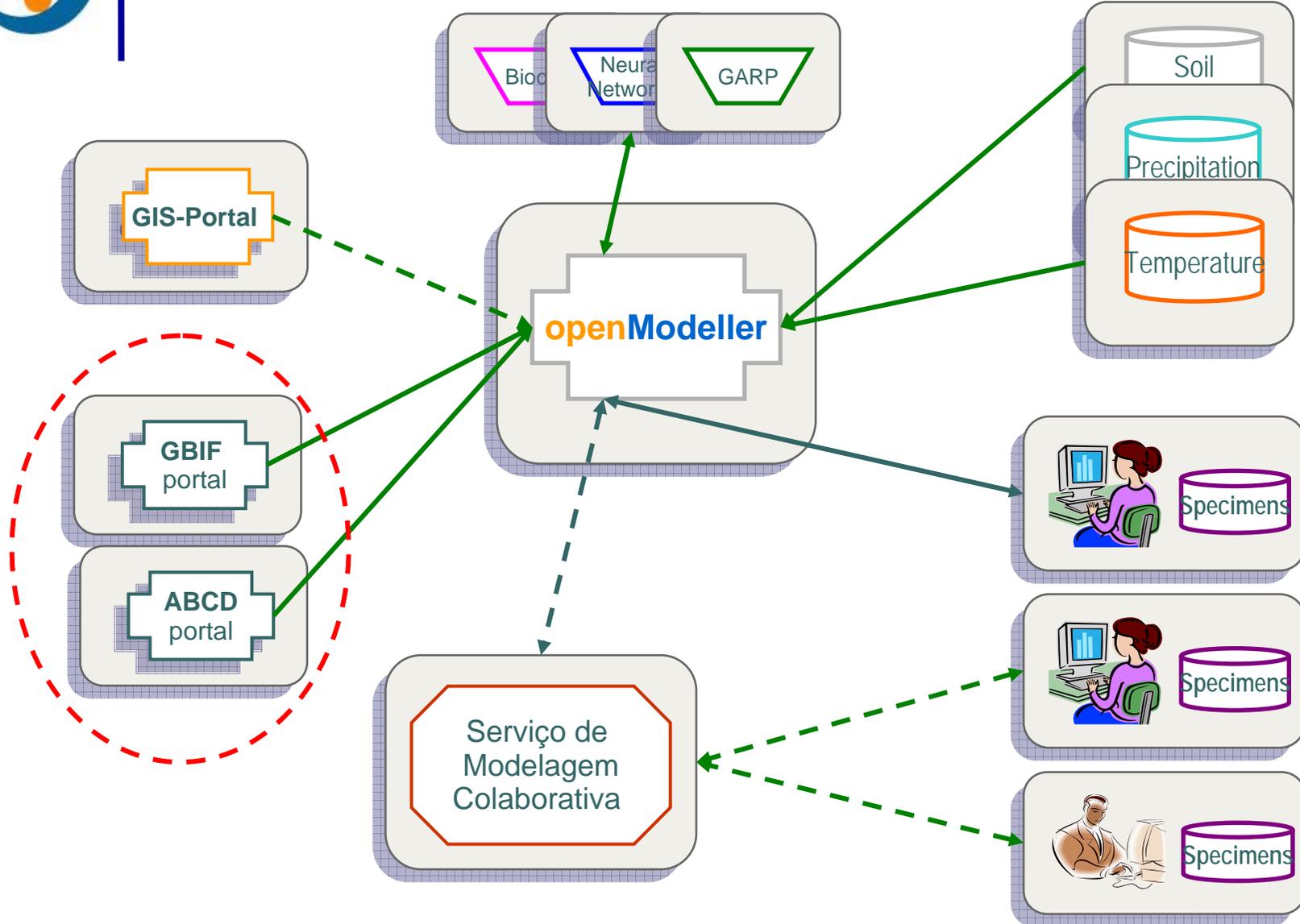




openModeller

Modelling algorithms

Environmental data



1. Instituições e/ou usuários cooperam com dados

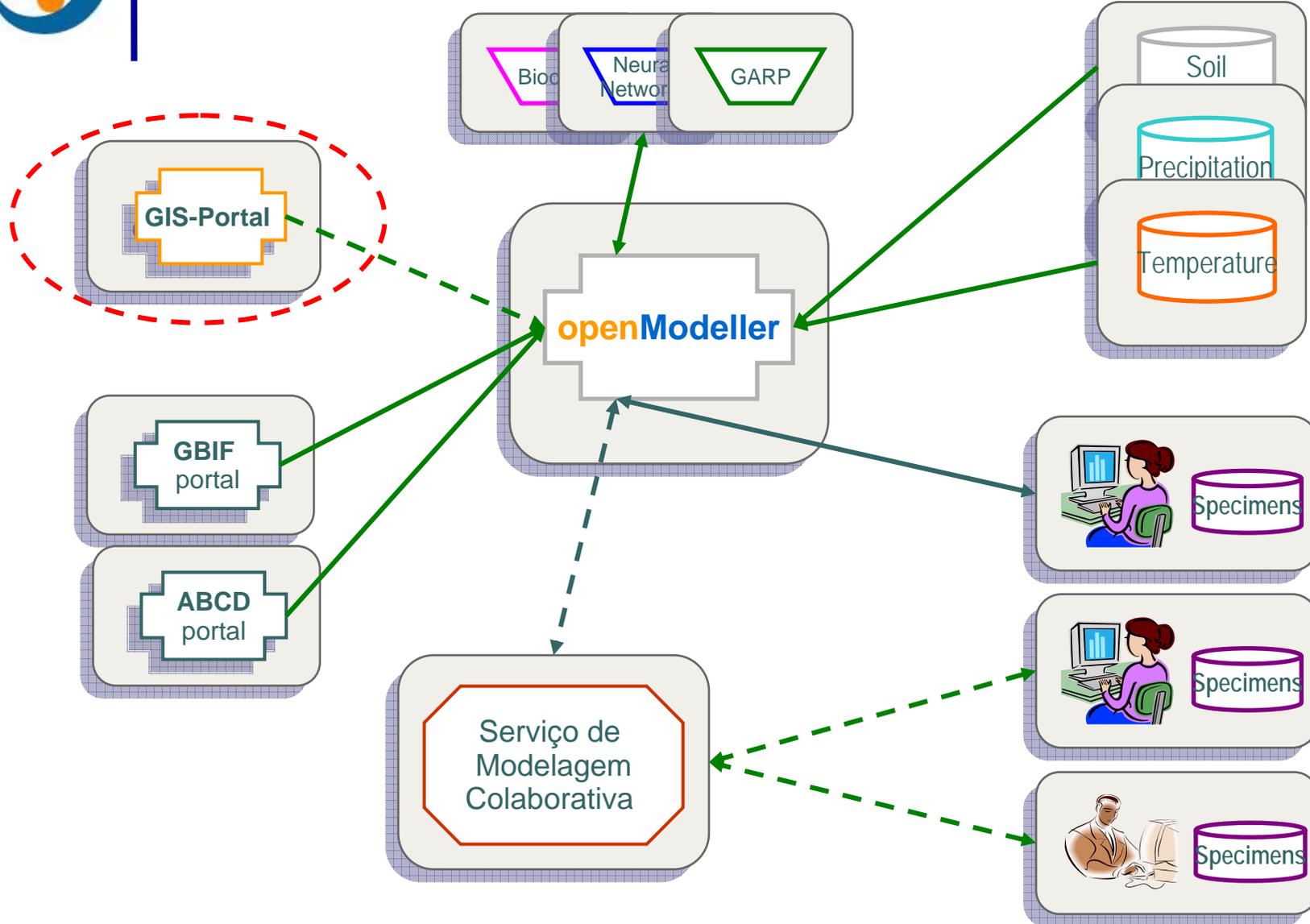




openModeller

Modelling algorithms

Environmental data



1. Instituições e/ou usuários cooperam com serviços

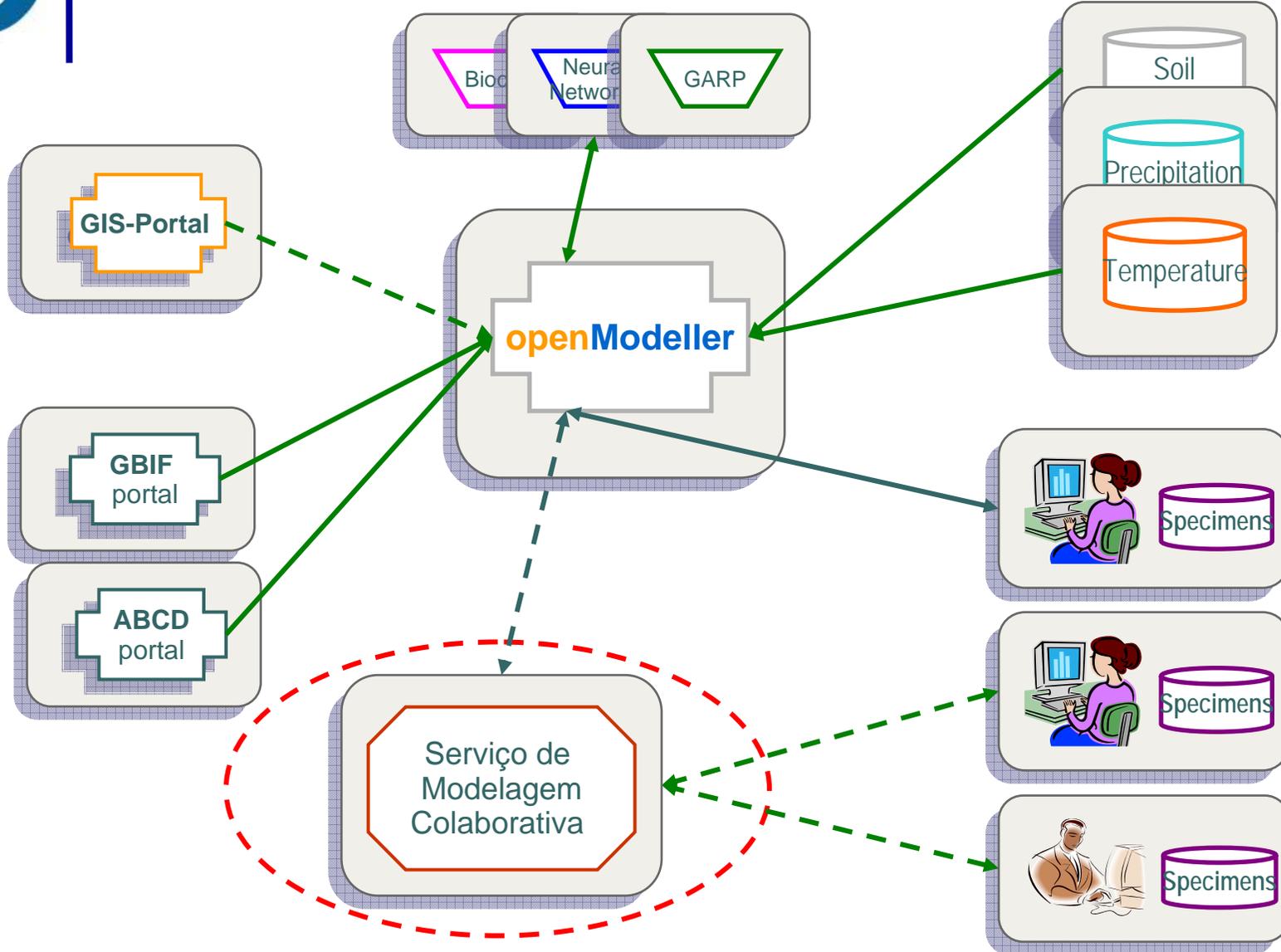




openModeller

Modelling algorithms

Environmental data



1. Usuários cooperam com Instâncias de Modelos





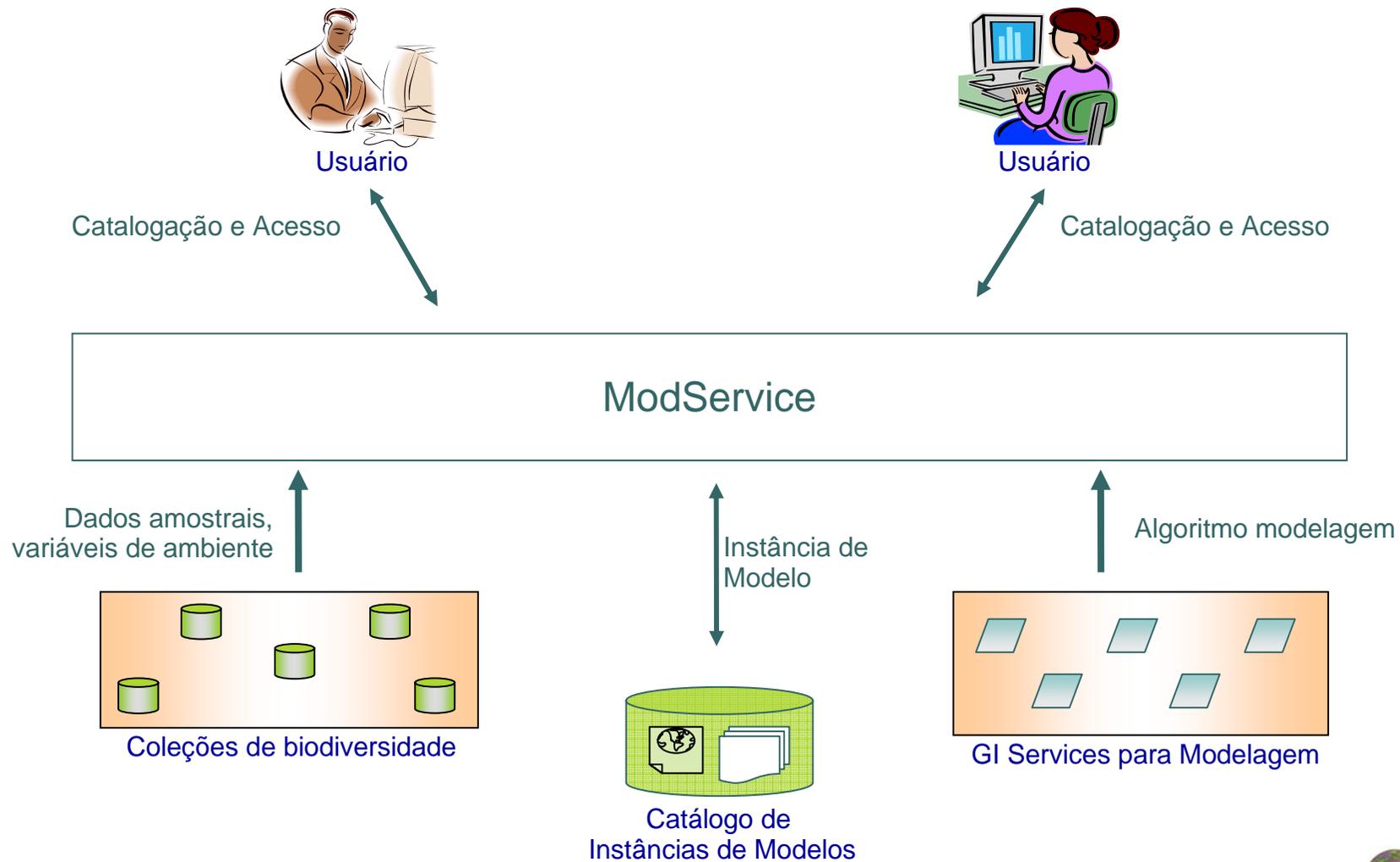
WBCMS – *WEB Biodiversity Collaborative Modelling Service* – openModeller

- A Modelagem Colaborativa da Biodiversidade é alcançada através do WBCMS, que provê
 - Catalogação de resultados de modelagens de biodiversidade
 - Acesso aos catálogos de resultados de modelagens de biodiversidade
- Ao gerar um modelo no openModeller, o usuário poderá disponibilizar sua instância para os demais usuários
- O resultado do modelo contemplará metadados do modelo
 - autor, dados de entrada/pontos de ocorrência, variáveis ambientais, figuras (JPEG ou GIF) do mapa de distribuição de espécies, etc.





WBCMS – *WEB Biodiversity Collaborative Modelling Service* – openModeller





Cronograma

	Jul-Set 2006	Out-Dez 2006	Jan-Mar 2007	Abr-Jun 2007	Jul-Set 2007	Out-Dez 2007	Jan-Mar 2008
Construção dos elementos conceituais e teóricos	X						
Especificação do ambiente computacional proposto	X						
Desenvolvimento do ambiente computacional		X	X	X	X		
Experimentação (Estudo de Caso)			X	X	X	X	
Documentação do Trabalho							
Redação de artigos	X			X		X	
Redação da Tese	X	X	X	X	X	X	X
Defesa da Tese							X



Considerações Finais

- Contribuições para o avanço do conhecimento científico
 - Ambiente computacional voltado a serviços, que suporta cooperação entre pesquisadores através do compartilhamento de conhecimentos obtidos após modelagens
- Resultados Esperados
 - Produção de um documento com conclusões a cerca do ganho do uso de um ambiente computacional, que permita a cooperação dos usuários através de resultados de modelagens



Obrigada



Leontopithecus rosalia



Alouatta ululata



Sagulus midas



Saimiri sciureus



Referências

- Alameh, N. **Scalable and Extensible Infrastructures for Distributing Interoperable Geographic Information Services on the Internet**. Massachusetts Institute of Technology, Massachusetts, 2001.
- Alameh, N. Chaining geographic information web services. **IEEE Internet Computing**, 2003.
- Anderson, G.; Moreno-Sanchez, R. Building Web-Based Spatial Information Solutions around Open Specifications and Open Source Software. **Transactions in GIS**, v. 7, n. 4, p. 447-466, 2003.
- Bernard, L.; Einspanier, U.; Lutz, M.; Portele, C. Interoperability in GI Service Chains-The Way Forward. In: 6th AGILE Conference on Geographic Information Science, 2003, Muenster. Disponível em: musil.uni-muenster.de.
- Câmara, G.; Monteiro, A. M.; Davis, C. **Introdução à Ciência da Geoinformação**. São José dos Campos: INPE, 2003. Disponível em: www.dpi.inpe.br/gilberto/livro.
- Canhos, V. P.; Souza, S.; Giovanni, R. D.; Canhos, D. A. L. Global biodiversity informatics: setting the scene for a "new world" of ecological modeling. **Biodiversity Informatics**, v. 1, p. 1-13
- Cirne, W.; Paranhos, D.; Costa, L.; Santos-Neto, E.; Brasileiro, F.; Sauv e, J.; Silva, F. A. B. d.; Barros, C. O.; Silveira, C. Running Bag-of-Tasks Applications on Computational Grids: The MyGrid Approach. In: ICCP - International Conference on Parallel Processing, 2003, Kaohsiung. 32: IEEE Computer Society, p. 407.
- Curbera, F.; Duftler, M.; Khalaf, R.; Nagy, W.; Mukhi, N.; Weerawarana, S. Unraveling the Web services web: an introduction to SOAP, WSDL, and UDDI. **IEEE Internet Computing**, 2002.
- Di, L.; Chen, A.; Yang, W.; Zhao, P. The Integration of Grid Technology with OGC Web Services (OWS) in NWGISS for NASA EOS Data. In: HPDC12 (Twelfth IEEE International Symposium on High-Performance Distributed Computing) & GGF8 (The Eighth Global Grid Forum), 2003, Seattle, Washington, USA.





Referências

- Foster, I.; Kesselman, C.; Tuecke, S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. **International J. Supercomputer Applications**, v. 15, n. 3
- Giovanni, R. D. The openModeller project. In: BiodiversityWorld GRID workshop, 2005, Edinburgh. e-Science Institute, Disponível em: <http://www.nesc.ac.uk/esi/events/594/>.
- Goodchild, M. F. The current status of GIS and spatial analysis. **Journal of Geographical Systems**, v. 2, p. 5-10
- Haines, M. N. Web services as information systems innovation: a theoretical framework for Web service technology adoption. In: IEEE International Conference on Web Services, 2004, San Diego, California, USA. p. 11-16.
- Hernandez, T.; Kambhampati, S. Integration of Biological Sources: Current Systems and Challenges Ahead. **SIGMOD Record**, v. 33, n. 3
- Hobern, D., 2004, **Architecture and Tools for Biodiversity Data Exchange**, Converging Sciences Conference, Trento, Italy.
- Jones, A. C.; White, R. J.; Pittas, N.; Gray, W. A.; Sutton, T.; Xu, X.; Bromley, O.; Caithness, N.; Bisby, F. A.; Fiddian, N. J.; Scoble, M.; Culham, A.; Williams, P. BiodiversityWorld: An architecture for an extensible virtual laboratory for analysing biodiversity patterns. In: UK e-Science All Hands Meeting, 2003, Nottingham, UK. Cox, S.J., p. 759-765. Disponível em: <http://www.bdworld.org/index.php?option=content&task=view&id=11&Itemid=26>.
- Marquezan, C. C. **Estudo das características de serviçosWeb (Web Services) e de seus cenários de aplicação**. Universidade Federal do Rio Grande do Sul - Instituto de Informática. Programa de Pós-Graduação em Computação, 2004. 61 p.
- Newcomer, E., 2002, **Understanding Web Services- XML, WSDL, SOAP and UDDI**, In: Addison-Wesley, ed.
- OGC. **The Spatial Web. An Open GIS Consortium (OGC) White Paper**. 2004. Disponível em: <http://www.opengis.org>.





Referências

- Osthoff, C.; Almeida, R. A. d.; C.V.Monteiro, A.; Strauch, J.; Souza, J. M. d.; Brito, H. M. d. **MODGRID – Um ambiente na WEB para desenvolvimento e execução de modelos espaciais em um ambiente de Grades Computacionais**. Petrópolis: LNCC - Laboratório Nacional de Computação Científica, 2004.
- Panatkool, A.; Laoveeraku, S. Decentralized GIS Web Services on Grid. In: Open source GIS - GRASS users conference, 2002, Trento, Italy.
- Pinto, G. d. R. B.; Medeiros, S. P. J.; Souza, J. M. d.; Strauch, J. C. M.; Marques, C. R. F. Spatial data integration in a collaborative design framework. **Communications of the ACM**, v. 46, n. 3, p. 86-90
- Tarouco, L. M. R., 2005, **Tutorial Cliente-Servidor**, In: http://penta2.ufrgs.br/redes296/cliente_ser/tutorial.htm, ed., Rio Grande do Sul, Universidade Federal do Rio Grande do Sul.
- Tsou, M. H.; Battenfield, B. P. A Dynamic Architecture for Distributing Geographic Information Services. **Transactions in GIS**, v. 6, n. 4, p. 355-381, 2002.
- Tu, S.; Xu, L.; Abdelguerfi, M.; Ratcliff, J. J. Applications: Achieving interoperability for integration of heterogeneous COTS geographic information systems In: 10th ACM international symposium on Advances in geographic information systems, 2002, McLean, VA (near Washington, D.C.). ACM Press,
- W3C. **W3C - Web Services Description Language (WSDL) 2.0**. 2004. Disponível em: <http://www.w3.org>.
- White, R., 2004, **Helping biodiversity researchers to do their work: collaborative e-Science and virtual organisations**, Converging Sciences Conference, Trento, Italy.
- Zhao, Y.; Wilde, M.; Foster, I.; Voekler, J.; Jordan, T.; Quigg, E.; Dobson, J. Grid middleware services for virtual data discovery, composition, and integration In: 2nd workshop on Middleware for Grid Computing 2004, Toronto, Ontario, Canada. ACM Press,

