

## Modelo de Dados para Geração de Ontologias Geográficas

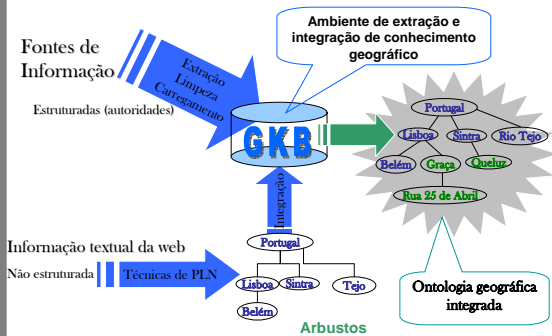
Marcirio Silveira Chaves  
 Catarina Rodrigues  
 Mário J. Silva  
 {mchaves,crodrigues,mjs}@xldb.fc.ul.pt

Departamento de Informática  
 Faculdade de Ciências  
 Universidade de Lisboa

## Motivação

- ❑ Informação geográfica de Portugal dispersa
- ❑ Repositório único -> Ontologia Geográfica
- ❑ Informação geográfica relacionada
- ❑ Domínio geográfico administrativo e físico
- ❑ Fácil reuso -> OWL

## Contexto

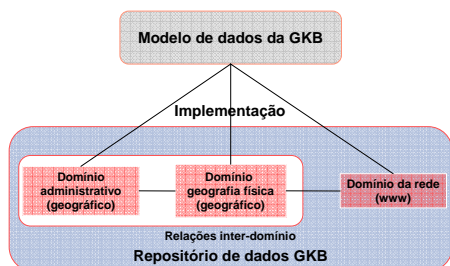


## Contexto

- ❑ GREASE - <http://xldb.fc.ul.pt/grease>
- ❑ *Geographic Knowledge Base* (GKB)
  - ambiente para integrar dados geográficos e gerar ontologias
  - baseado em um modelo capaz de suportar o armazenamento de informação geográfica

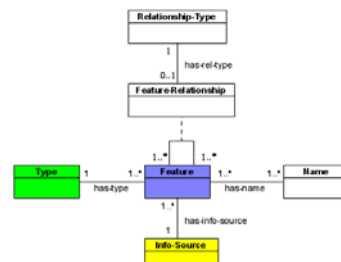
## Contexto

- ❑ Domínios da GKB



## Contexto

- ❑ GKB 1.0 -> Geo-Net-PT01 - <http://xldb.fc.ul.pt/geonetpt>
- ❑ GKB 1.0 Meta-modelo



## Estrutura da Apresentação

- ❑ Requisitos da GKB 2.0
- ❑ Modelo para o Domínio Físico
- ❑ Publicando o conteúdo da GKB para Aplicações
- ❑ Trabalhos Relacionados
- ❑ Considerações Finais

19-Feb-07

XATA 2007

12

## Requisitos do Meta-modelo da GKB 2.0

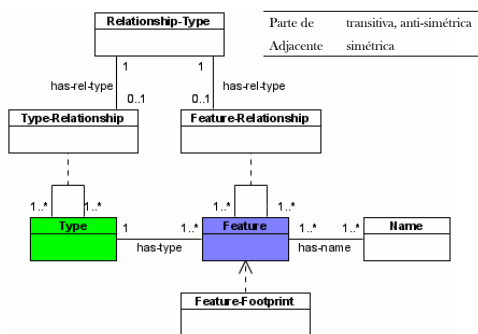
- ❑ Suporte para relacionamentos entre tipos
- ❑ Suporte para conjunto de atributos genéricos
- ❑ Melhor controle das fontes de informação
- ❑ Suporte multi-língua

19-Feb-07

XATA 2007

13

## GKB 2.0 - Meta-Modelo Base

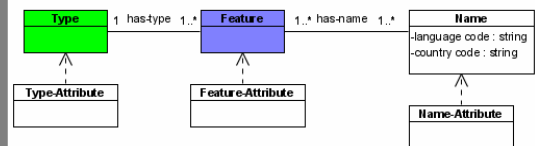


19-Feb-07

XATA 2007

14

## Representação dos Nomes e Atributos

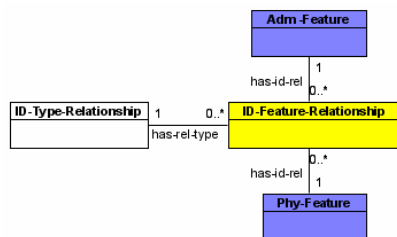


19-Feb-07

XATA 2007

15

## Relacionamentos Inter-domínio

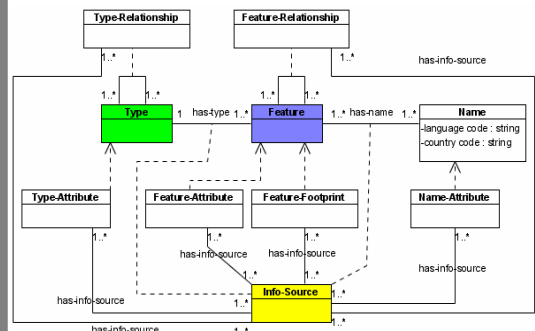


19-Feb-07

XATA 2007

16

## Fontes de Informação



19-Feb-07

XATA 2007

17

## Domínio Físico

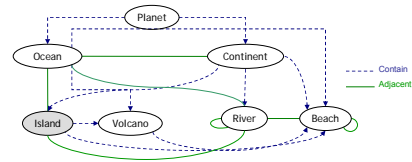
- Maior complexidade de modelação e integração
- Conjunto mais rico de relacionamentos
  - Explícitas
    - Ex.: contém (total ou parcialmente), intersecta, adjacente, etc
  - Implícitas
    - A oeste de..., a X Km de..., à menor distância, com o menor custo, etc.

19-Feb-07

XATA 2007

18

## Modelo para o Domínio Físico - Conceitos

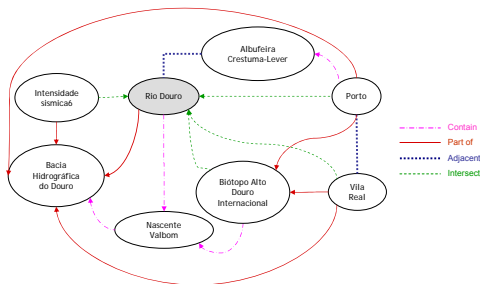


19-Feb-07

XATA 2007

19

## Modelo para o Domínio Físico - Ocorrências



19-Feb-07

XATA 2007

20

## Tipos e Fontes de Informação - Domínio Físico

Tipos	# de features	Fontes de Informação
<b>Rede Hidrográfica</b>	<b>7,589</b>	
Edifício Notável (e.g. igreja)	3,877	□ Ministério do Ambiente
Estabelecimento Dormida	2,847	□ Instituto Geográfico do Exército
Espécie Notável (e.g. sobreiro)	2,432	□ Instituto da Água
Litologia	2,245	□ Instituto Nacional de Estatística
Albufeira (ponto)	1,800	□ Instituto de Pesquisa da Marinha
<b>Praia</b>	<b>591</b>	□ etc...
<b>Museu</b>	<b>524</b>	
Estação Arqueológica	456	
Paul (e.g. Boquilobo)	441	
Σ 31 tipos	3,112	
<b>Total features</b>	<b>25,914</b>	

19-Feb-07

XATA 2007

21

## Estrutura da Apresentação

- Requisitos da GKB 2.0
- Modelo para o Domínio Físico
- Publicando o conteúdo da GKB para Aplicações
- Trabalhos Relacionados
- Considerações Finais

19-Feb-07

XATA 2007

22

## Publicando o conteúdo da GKB para Aplicações

- Diferentes visões sobre o mesmo conteúdo
- Visão centrada em nomes
  - Aplicações
    - Reconhecedor de Entidades Mencionadas
- Visão centrada em *features*
  - Aplicações
    - Recuperação de Informação Geográfica
- MySQL -> PostgreSQL (PostGIS)
- Perl (importação e exportação de dados)
- OWL (*Lite e Full*)

19-Feb-07

XATA 2007

23

## Publicando o conteúdo da GKB para Aplicações

- Visão centrada em nomes

```
<gn:Geo-Name rdf:ID="GEO_ADM_80">
  <gn:names>
    <rdf:Bag>
      <rdf:li gname="Lisboa" xml:lang="PT-PT" gnatt="P" gnis="INE"/>
      <rdf:li gname="Lissabon" xml:lang="DE-DE" gnatt="A" gnis="WIKI"/>
      <rdf:li gname="Ατσαβόνα" xml:lang="EL-GR" gnatt="H" gnis="WIKI"/>
    </rdf:Bag>
  </gn:names>
  <gn:geo_type_id rdf:resource="#PRO"/>
  <gn:geo_type_id rdf:resource="#CON"/>
  <gn:geo_type_id rdf:resource="#RUA"/>
  <gn:geo_type_id rdf:resource="#AVE"/>
</gn:Geo-Name>
```

19-Feb-07

XATA 2007

24

## Publicando o conteúdo da GKB para Aplicações

- Visão centrada em *features* (1)

```
<gn:Geo_Feature rdf:ID="GEO_PHY_145">
  <gn:names>
    <rdf:Bag>
      <rdf:li gname="Douro" xml:lang="PT-PT" gnatt="P" gnis="IGOE"/>
      <rdf:li gname="Douro" xml:lang="ES-ES" gnatt="A" gnis="IGP"/>
    </rdf:Bag>
  </gn:names>
  <gn:geo_type_id gnis="IGEO" rdf:resource="#Rio"/>
  <rdf:comment>Serra de Urbiño - Spain</rdf:comment>
  <gn:source_river gnis="IGEO" rdf:resource="#GEO_PHY_120"/>
  <rdf:comment>Porto - Portugal</rdf:comment>
  <gn:outlet_river gnis="IGEO" rdf:resource="#GEO_ADM_288"/>
  <gn:tributary gnis="IGEO">
    <rdf:Bag>
      <rdf:li rdf:resource="#GEO_PHY_400"/>
      <rdf:li rdf:resource="#GEO_PHY_401"/>
    </rdf:Bag>
  </gn:tributary>
  <gn:length unit="km" gnis="IGEO">850</gn:length>
</gn:Geo_Feature>
```

19-Feb-07

XATA 2007

25

## Publicando o conteúdo da GKB para Aplicações

- Visão centrada em *features* (2)

```
<rdf:Bag><rdf:li>
  <gn:Geo_Relationship>
    <gn:rel_type_id rdf:resource="#INTERSECTS"/>
    <gn:geo_id><rdf:Bag>
      <rdf:li rdf:resource="#GEO_ADM_288"/>
      <rdf:li rdf:resource="#GEO_PHY_300"/>
    </rdf:Bag></gn:geo_id>
  </gn:Geo_Relationship>
</rdf:li><rdf:li>
  <gn:Geo_Relationship>
    <gn:rel_type_id rdf:resource="#ADJ"/>
    <gn:geo_id rdf:resource="#GEO_PHY_198"/>
  </gn:Geo_Relationship>
</rdf:li><rdf:li>
  <gn:Geo_Relationship>
    <gn:rel_type_id rdf:resource="#FKT"/>
    <gn:geo_id rdf:resource="#GEO_PHY_100"/>
  </gn:Geo_Relationship>
</rdf:li></rdf:Bag> </gn:Geo_Feature>
```

19-Feb-07

XATA 2007

26

## Estrutura da Apresentação

- Requisitos da GKB 2.0
- Modelo para o Domínio Físico
- Publicando o conteúdo da GKB para Aplicações
- Trabalhos Relacionados
- Considerações Finais

19-Feb-07

XATA 2007

27

## Trabalhos Relacionados

- *Getty Thesaurus of Geographic Names (TGN)*
  - detalhes de construção e procedimento de integração dos dados do recurso são desconhecidos
  - cerca de 1 milhão de locais em todo o mundo.
  - falta informação detalhada sobre Portugal
  - XML
- *Geonames World Ontology ([geonames.org](http://geonames.org))*
  - OWL *Lite* e *Full*
  - poucos detalhes sobre a integração dos dados
  - fontes de informação não são exportadas na ontologia
  - 6.2 milhões de nomes geográficos em todo o mundo
  - cobertura dos nomes para Portugal é muito limitada

19-Feb-07

XATA 2007

28

## Considerações Finais

- Nova versão da GKB, detalhando o modelo de dados estendido
- Modelo capaz de gerar ontologias representando o conhecimento geográfico integrado.
- Visões diferentes do conteúdo armazenado na GKB
- Ontologias para aplicações na Web Semântica
- Fácil reuso do conhecimento

19-Feb-07

XATA 2007

29

## Trabalho em Andamento

- ❑ Extração, Transformação e Carregamento dos dados geográficos cobrindo o domínio físico
- ❑ Geração e lançamento da Geo-Net-PT02

19-Feb-07

XATA 2007

30

Obrigado!



<http://xldb.fc.ul.pt/geonetpt>

Marcirio Chaves - [mchaves@xldb.fc.ul.pt](mailto:mchaves@xldb.fc.ul.pt)

Perguntas?

19-Feb-07

XATA 2007

31

## Agradecimentos

- ❑ Marcirio Chaves é bolseiro da Fundação para a Computação Científica Nacional (FCCN) e o trabalho descrito aqui foi (parcialmente) financiado pela Fundação para a Ciência e Tecnologia (FCT) através do projecto POSI/PLP/43931/2001, co-financiada pelo POSI, e pelo projecto POSC/339/1.3/C/NAC, financiado pelo POSC.
- ❑ Este trabalho é financiado pelo POSI/SRI/47071/2002 (GREASE) from FCT, co-financed by POSI

19-Feb-07

XATA 2007

32