GeoBuddies: **Anotación semántica colaborativa con dispositivos móviles en el Camino de Santiago**

WP2

Index

- WP2 Overview
- Starting point: Knowledge bases and proposed solution
- Step 1: Authomatic ontology building
- Step 2: Mapping Definition
- Step 3: Mapping Discovery
- Step 4: Mapping Interpretation
- Step 5: Integration all Sources Information
- Step 6: Mapping Composition
- Step 7: Multilinguality
- Work to be done to December 2008

WP.2 Ontology-based multilingual information integration

- Objective 3. **semantic accessibility to IGN resources and other data available in the Web**
  - harmonised views of IGN databases → OntoBCN
  - Multilingual content (Spanish, Galician and Basque)
  - Information integration techniques
- Main Steps and results
  1. Domain Ontology
  2. Onto-BD mapping definition ($R_2$0 is not enough)
  3. Onto-BD mapping discovery algorithm
  4. Mapping execution (ODEMapster is not enough)
  5. Information integration with external data
  6. Composition procedural and declarative mappings
  7. Multilinguality
- Strong cooperation IGN-UPM (2007,2008)

Starting point

- Monolingual **Knowledge bases** of IGN (spanish): **NC** (Nomenclátor Conciso), **NGN** (Nomenclátor Geográfico Nacional), **BCN200** (Base Cartográfica Nacional escala 1:200.000), **BCN25** (Base Cartográfica Nacional escala 1:25.000)

- Monolingual **Knowledge bases** of CC.AA. (spanish, basque, galician): Castilla y León, Cataluña, Euskadi, Extremadura, Galicia, La Rioja, Madrid, Murcia, Navarra.

- Creation of an **ontology** mapped with knowledge bases

Knowledge Bases

Nomenclátor Conciso (NC)

Nomenclator conciso - Bloc de notas

Archivo Edición Formato Ayuda

```
Comunidad Autónoma
Provincia
Comarca administrativa
Capital
Población 1
Población 2
Sierra
Pico
Puerto
Área geográfica
Río
Canal
Lago
Embalse
Cabo
Golfo
Playa
Isla
Mar
Lugar protegido
Aeropuerto
Puerto marítimo
```

Base Cartográfica N. (BCN200)

Bcn200.tbl - Bloc de notas

Archivo Edición Formato Ayuda

```
CODIGO LV COL PS LC SIMB. NOMBRE
------ -- --- -- -- ----- ---
010101 01 000 00  6 00006 LIM
010102 01 015 00  6 03846 LIM
010201 01 030 03  4 07708 LIM
010301 01 045 06  2 11570 LIM
010401 01 060 06  3 15411 LIM
010501 01 075 00  1 19201 AGU
015101 01 090 00  6 23046 MUN
015131 01 105 00  6 26886 MUN
015191 01 120 00  6 30726 MUN
015201 01 135 03  4 34588 PRC
015231 01 150 03  4 38428 PRC
015291 01 165 03  4 42268 PRC
015301 01 180 06  2 46130 AUT
015331 01 195 06  2 49970 AUT
015391 01 210 06  2 53810 AUT
015401 02 000 06  3 00051 NAC
015431 02 015 06  3 03891 NAC
015491 02 030 06  3 07731 NAC
015501 02 045 00  6 11526 MUN
015531 02 060 00  6 15366 MUN
015591 02 075 00  6 19206 MUN
015601 02 090 00  6 23046 MUN
015631 02 105 00  6 26886 MUN
015691 02 120 00  6 30726 MUN
015701 02 135 00  6 34566 MUN
015731 02 150 00  6 38406 MUN
015791 02 165 00  6 42246 MUN
015801 02 180 00  6 46086 TER
015831 02 210 00  6 53766 TER
015891 02 225 00  6 57606 TER
015901 03 000 03  3 00027 TER
015931 03 015 03  3 03867 TER
015991 03 030 03  3 07707 TER
016001 03 045 03  3 11547 COM
016031 03 060 03  3 15387 COM
016091 03 075 03  3 19227 COM
016101 03 090 03  1 23065 PAR
016131 03 105 03  1 26905 PAR
016191 03 120 03  1 30745 PAR
016201 03 135 03  1 34585 PAR
016231 03 150 03  1 38425 PAR
016291 03 165 03  1 42265 PAR
016301 03 180 03  1 46105 RES
016331 03 195 03  1 49945 RES
016391 03 210 03  1 53785 RES
016401 04 000 06  0 00048 ZON
016500 04 015 06  3 03891 ZON
016501 04 030 06  3 07731 ZON
018101 04 045 01  0 11528 ADU
018201 04 060 04  0 15392 MO
018202 04 075 04  0 19232 MO
021101 08 006 00  1 01537 ESC
021102 08 021 00  1 05377 ACA
```

Nomenclator geográfico - Bloc de notas

Archivo Edición Formato Ayuda

```
NOMENCLÁTOR GEOGRÁFICO
ENTIDADES

Nación
Región geográfica
Capital de Nación
Elevación orográfica
Comunidad Autónoma
Llanura/Raso
Ciudad con Estatuto de Autonomía
Depresión orográfica
Capital de Comunidad Autónoma
Accidente costero
Provincia
Accidente marítimo
Capital de Provincia
Accidente hidrográfico
Coprincipado
Corriente fluvial
Capital de Coprincipado
Canal
Comarca
Embalse
Capital de Comarca
Lago/Laguna
Isla Humedal
Capital de Isla
Isla fluvial
Municipio
Isla marítima
Capital de Municipio
Garganta/Hoz
E.A.T.I.M.
Lugar/Paraje
Capital de E.A.T.I.M.
Paso/Collado
Población
Puerto de montaña
Comunidad de Municipios
Puerto comercial
Enclave
Helipuerto comercial
Territorio anejo
Aeródromo/Aeropuerto
Territorio autonómico
Estación de ferrocarril
Zona neutral
```

Dgn.cod - Bloc de notas

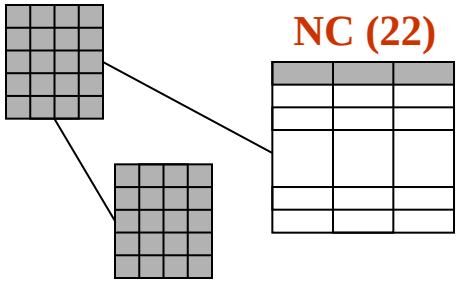
Archivo Edición Formato Ayuda

```
!Modificada 3/04/2007
!
! fichero creado por genlib
! Diccionario de conversión DGN -> EDM.
! FORMATO:
!   Tipo_dgn Entidad Tipo_istram Grupo Código_bcn Cerrado Trato [!comentario]
!
! Tipo_ dgn...NNSCCCGG                    Codigo_bcn...TTGGSS
!       NN  : Nivel elemento                      TT : Tema
!       S   : Estilo linea dgn                    GG : Grupo
!       CCC : Color  linea dgn                    SS : Subgrupo
!       GG  : Grosor linea dgn
!
! Entidad                                 Tipo_istram....???
!       104 : polilínea
!       203 : célula se convierte a símbolo
!        -1 : célula se explota en sus componentes
!       304 : rótulo
!
! Grupo
!         0 : sin determinar
!         1 : carreteras
!         2 : hidrografía
!         3 : conducciones
!         4 : administrativo
!
!         En textos el grupo corresponde a la fuente Microstation + Mayúsculas (M) / Ma
!
! Cerrado
!        en lineas                                en textos
!                1 : perimetral                           n : altura
!                0 : entidad lineal abierta
!               -1 : cultivo perimetral
!               -2 : cultivo linea abierta
!
! Trato
!   I: Intocable  A: Altimetría  N: No tratar  T: Textos Asociados
!   S: Textos Sueltos  C: Cultivo  F: Solo salida !: Tratar normalmente
!                                          TTGGSS
02000900   104     1   0   090101   1   !I   Marco de hoja
02300902   104     2   0   100200   0   !    Base Geodésica de Madridejos
06003900   104     3   0   025102   0   !    Acantilado
06006900   104     4   0   025302   0   !    Costa rocosa no acantilada
06009900   104     5   2   037402   1   !    Playa fluvial de guijarros. Contorno
06012900   104     6   0   025501   1   !    Lavas. Contorno
06015900   104     7   0   058303   0   !I   Dique de hormigón >15 metros
06018900   104     8   0   058304   0   !I   Dique de hormigón < 15 metros
07013400   104     9   0   058302   0   !I   Dique de tierra
07016400   104    10   0   055401   1   !    Vertedero. Contorno
11003003   104    11   1   062202   0   !    Autopista. Enlace
11012000   104    12   0   056091   1   !I   Patio. Contorno
13003300   104    13   1   060101   0   !    Autopista. Eje
13303300   104    14   1   060131   0   !    Autopista en Contrucción. Eje
14002401   104    15   1   066901   1   !I   Puesto de S.O.S.
14003301   104    16   1   067901   1   !I   Peaje
15003003   104    17   1   062204   0   !    Autovía. Enlace
15003004   104    18   1   060701   0   !    Autovía
```

N. Geográfico Nacional (NGN)

Base Cartográfica N. (BCN25)

Initial Knowledge Bases (IGN)

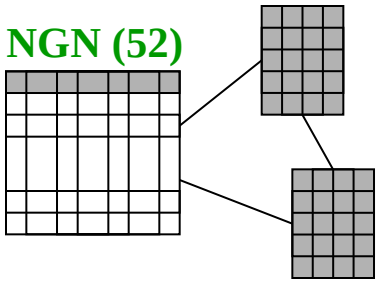


NC (22)

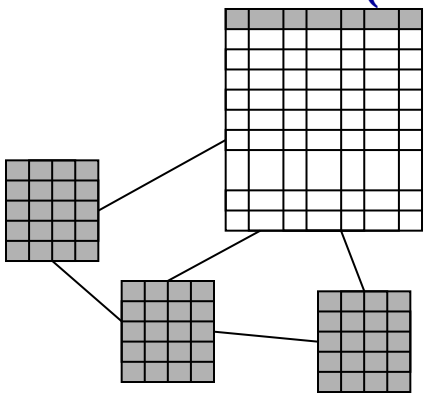
- Río

NGN (52)

- Corriente fluvial

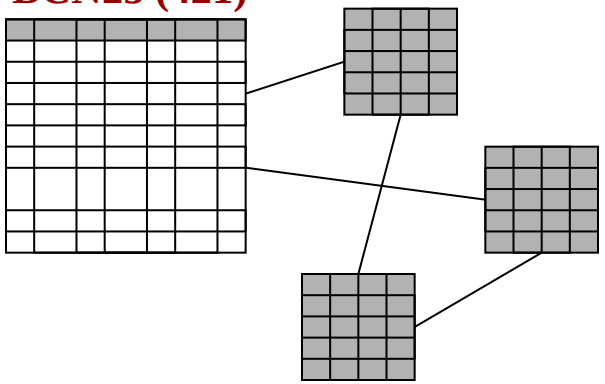


BCN200 (224)



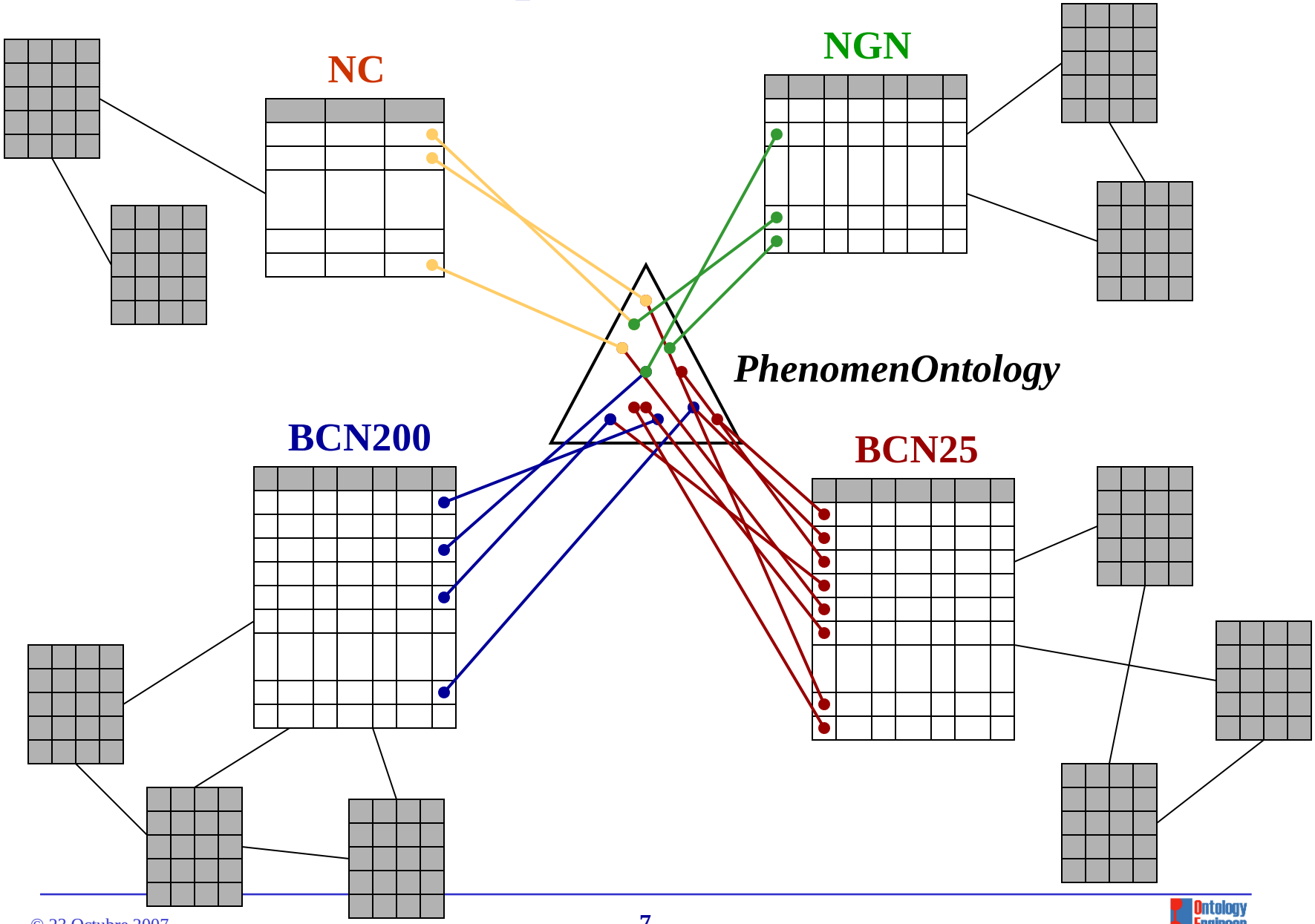
- Ríos, arroyos 3º y 4º, ramblas. Lagunas medianas
- Ríos y rías medianos de 2º orden. Lagunas > 100 Ha
- Ríos y rías principales 1º orden
- Ríos, arroyos, 5º y 6º, lagunas y charcas pequenas, fuentes, manantiales, pozos, etc.

BCN25 (421)



- **4 Knowledge Bases: NC, NGN, BCN200, BCN25**
- **Heterogeneity and granularity example**
- **Vocabulary differences: "*Autovia*", "*Autovías*", "*AUTOVIA.*"**

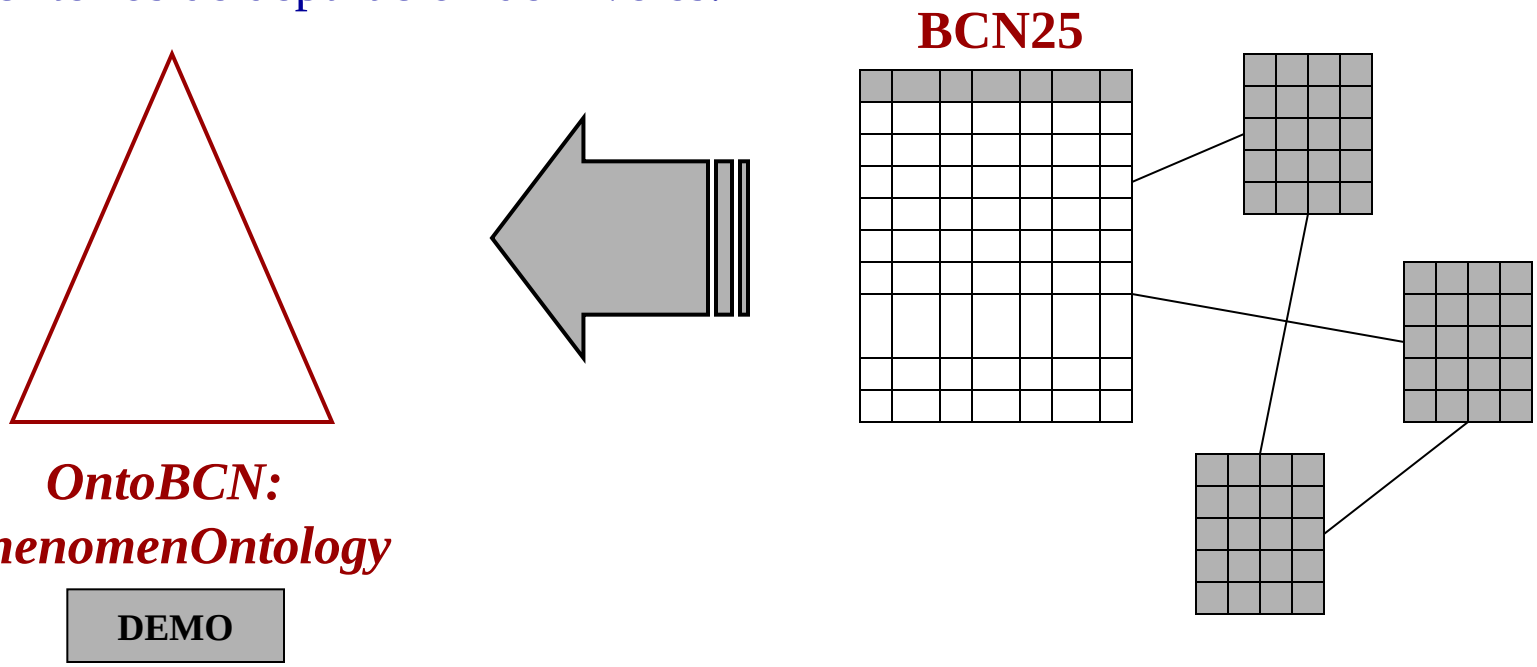
Proposed Solution



Step 1: Authomatic Ontology Building: PhenomenOntology

- Generación automática de la ontología a partir de BCN25 (*ontobcn*)

  → criterios de depuración de niveles.



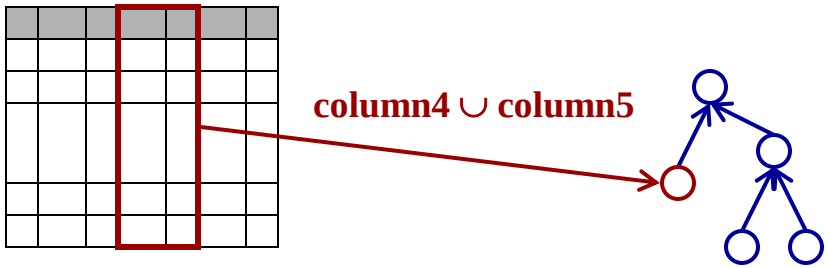
- Detección automática de diferencias lingüísticas (*linsearch*): plurales, tildes, diéresis, signos de puntuación, mayúsculas.
- Evaluándose en el IGN

Step 2: Onto-BD Mapping (R_2O)

R_2O is a language for expliciting procedural mappings between relational models and ontologies. (**model – model**)

ODEMapster is a compiler that cans execute mappings for:

- carrying out ontology population (DB → ontology)
- responsing queries to ontology with DB information (ontology → DB → ontology)

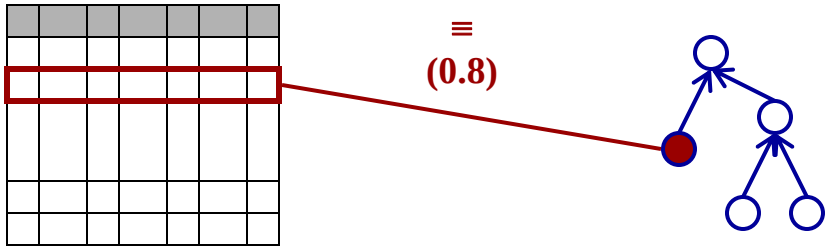



```
conceptmap-def
  name Cliente
  identified-by Usuarios.ID
  uri-as
    <transformation>
  applies-if
    <cond-expr>
  joins-via
    <cond-expr>
  documentation ”Correspondence description”
```

Step 2: Onto-BD Mapping (IGN case)

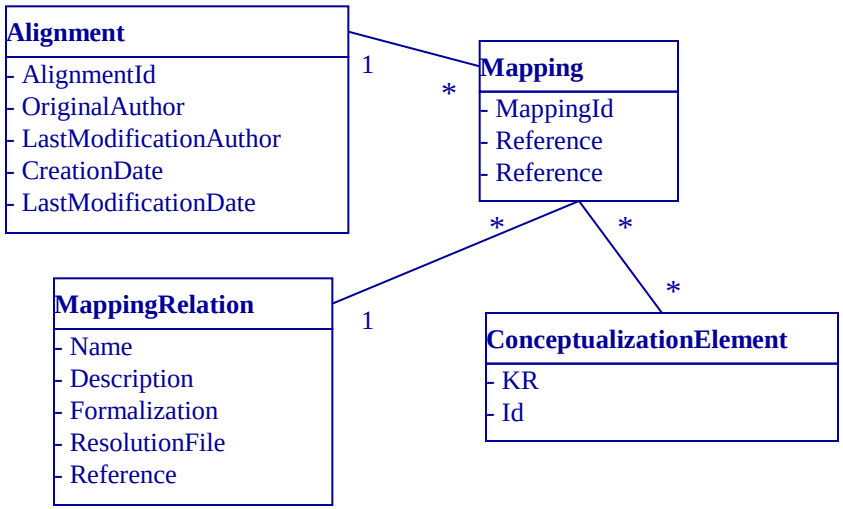
Ontology concepts are expressed as **instances of relational model**. (**model – individuals**)

Mappings does not express procedural relation (**actions condicionated to events**), express declarative relation (**facts with certainty**).



Step 2: Mapping Model Definition

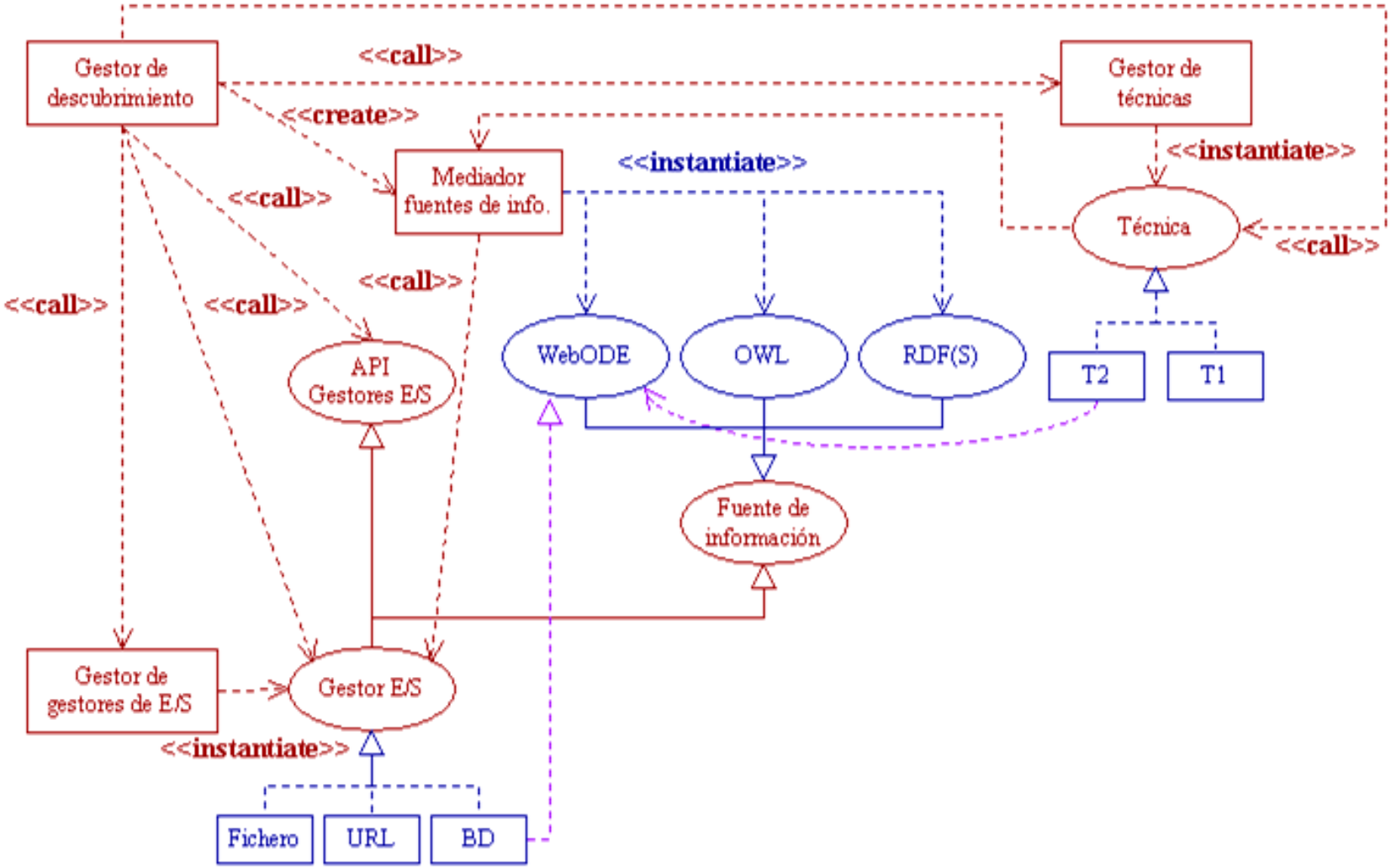
- Model of declarative mappings and XML example




```
<Mapping>
    <MappingId>BCN1_Fenomenos.Phenomens.030701-Phenomtology.Castillo</MappingId>
    <Certainty>1.0</Certainty>
    <Reference>Manual Mapping by JARG</Reference>
    <MappingRelation>
      <Name>Igualdad</Name>
      <Description>All elements are semantically equivalents.</Description>
      <Formalization>A=B=C=...</Formalization>
      <ResolutionFile>--None--</ResolutionFile>
      <Reference>OEG Mapping Relations</Reference>
    </MappingRelation>
      <ConceptualizationElement>
        <KR>DB</KR>
        <Id>BCN1_Fenomenos.Phenomens.030701</Id>
      </ConceptualizationElement>
      <ConceptualizationElement>
        <KR>Ontology</KR>
        <Id>Phenomtology.Castillo</Id>
      </ConceptualizationElement>
  </Mapping>
```

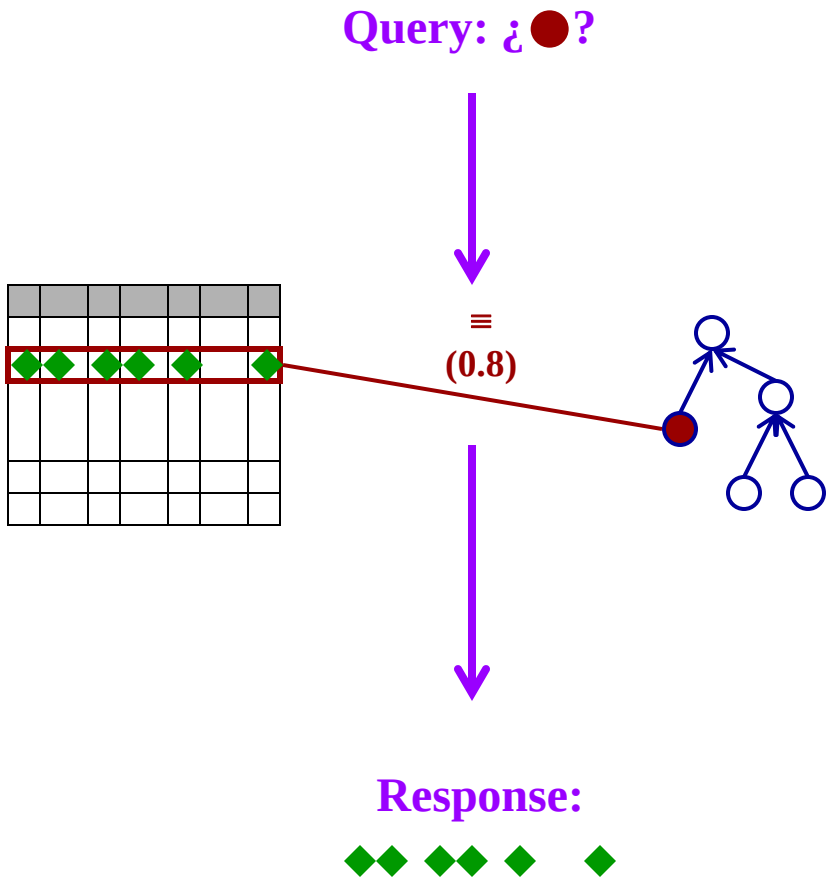
- Storage Systems
  - XML file (XSD defined yet)
  - WebODE now
  - To be decided new storage system

Step 3: Mapping discovery (in progress)

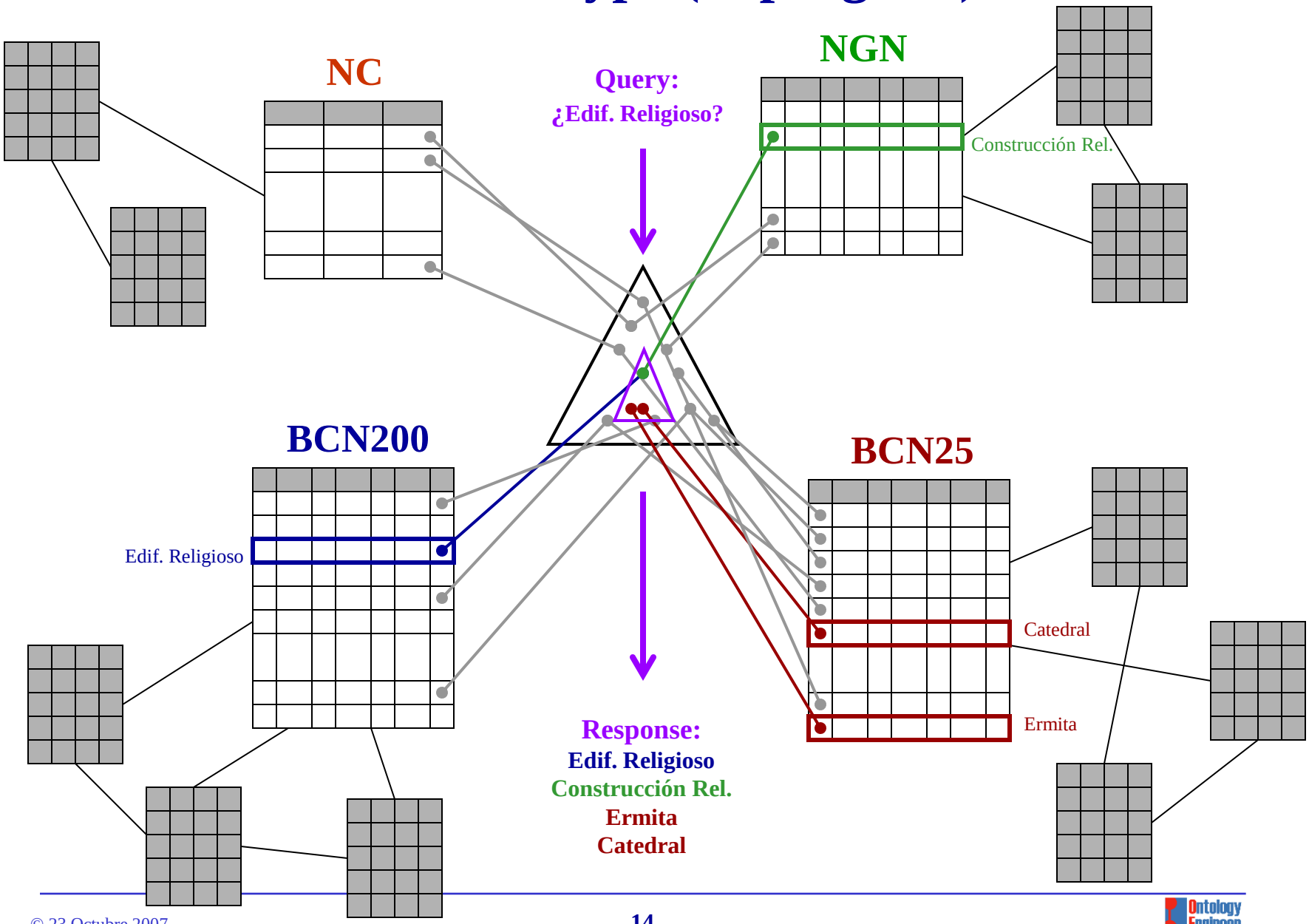


Step 4: Onto –BD Mapping interpretation (in progress)

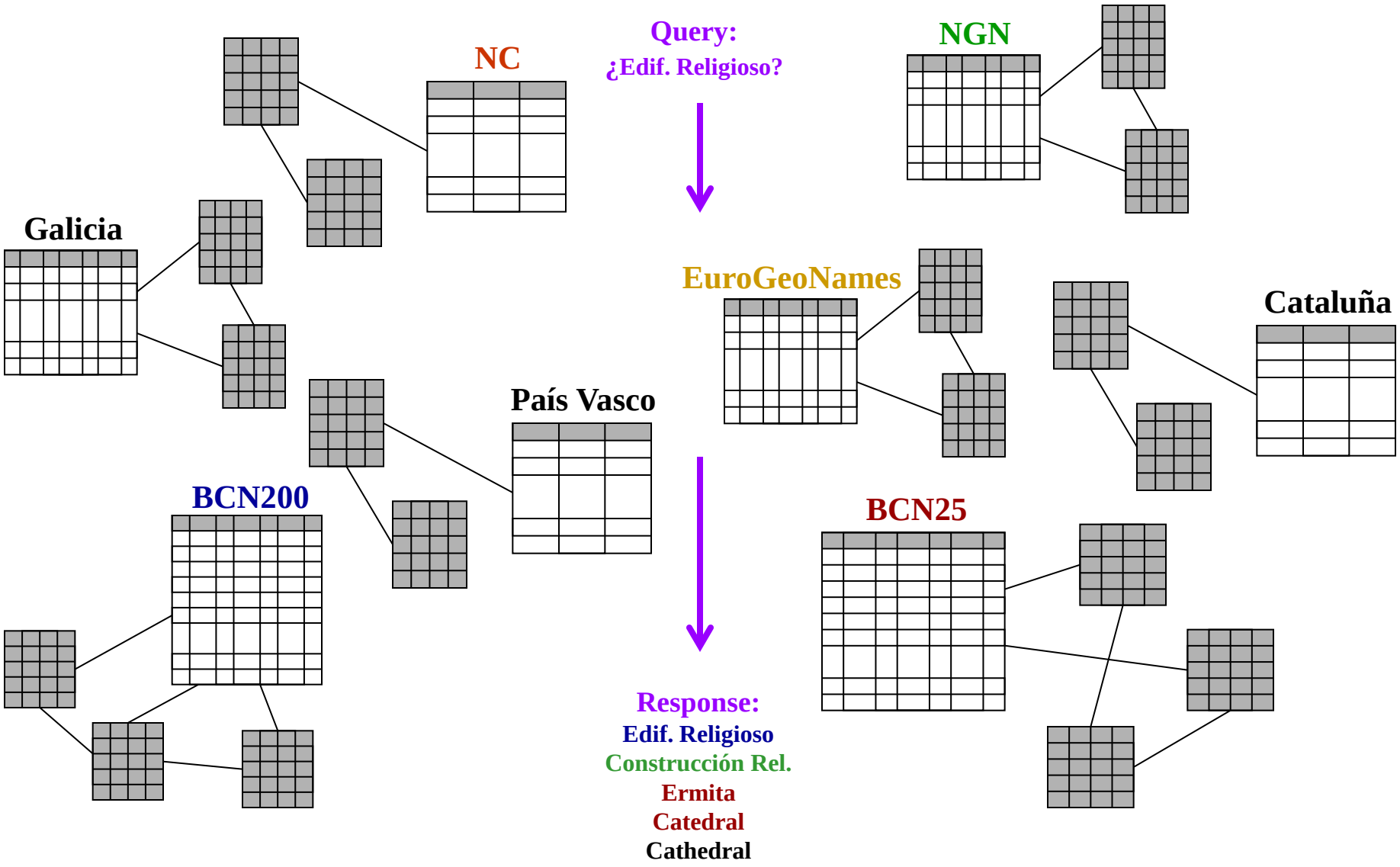
Mapping interpreter permits response queries using DB information transparently through declarative mappings



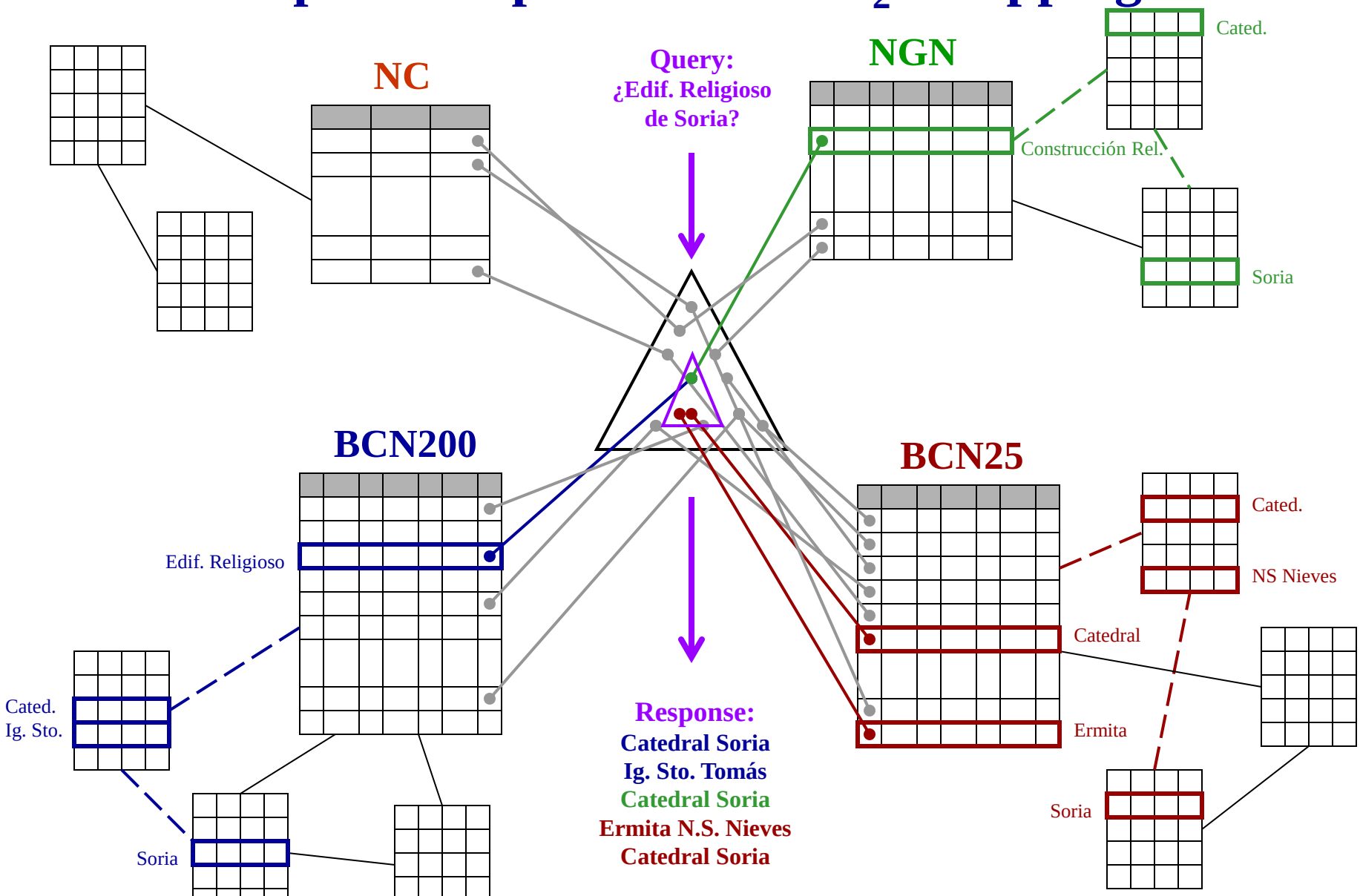
First Prototype (in progress)



Step 5: Integration of IGN+CCAA+European



Step 6: Composition with $R_2 0$ mappings



Step 7: Multilinguality

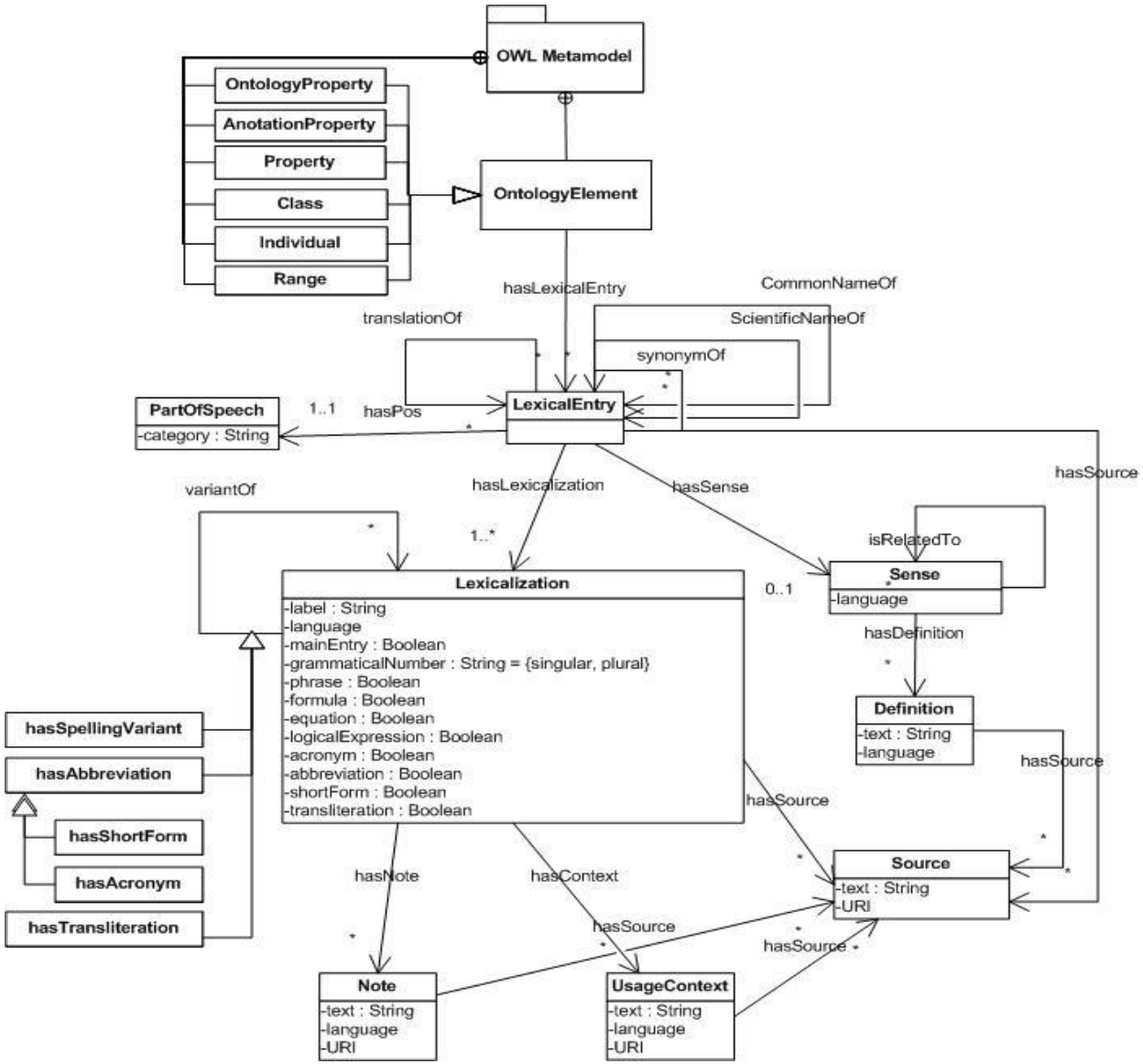
- ¿Quién va a interactuar con la ontología?
  - Un programa -> la multilingualidad no es necesaria
  - Un usuario -> interesa que la ontología sea multilingüe

- Si interesa que sea multilingüe aunque no hay necesidad,
  - ¿Cómo se va a proporcionar?
  - ¿En qué nivel ha de estar la multilingualidad?

Posible solutions

- Solution 1
  - NeonMultilingual Model
  - Label Translator
- Solution 2
  - Neon Multilingual Meta-model
    - Add linguistic information → elements are part of an ontology
      - classes, properties or relations
  - LabelTranslator plugin
    - Used sources
      - EWN databases
      - Web Translation resources
    - Linguistic information
      - Translated labels
      - Gloss or definition
    - Supported languages
      - English, German and Spanish

NeonMultilingual Model



Work to be done until Dec. 08

- OntoBCN
- Mapping Model def.