How INSPIRE has influenced the redesign of the French topographic database
Plan

- Context
- INSPIRE influence
- Conclusions
GENERAL CONTEXT
Context

- **BD UNI v1:**
  - Current (internal) production database
  - Large scale topographic data base (around 10K)
  - IGN main data

- **BD UNI v2 project:**
  - Redesign of data specification and of data production process
    - For various reasons
  - Specification work took place in 2016
Objectives

- Be more reactive – quicker updates
- Be more collaborative
- External products closer to production database
- Be simpler
- Be richer
- Be closer to INSPIRE
- Be closer to MGCP (Defense)
- Fulfil better user requirements
Methodology

Specifications of current product (BD UNI v1)
User requirements (marketing survey)
MGCP requirements
Matching tables exercise from BD UNI v1 to INSPIRE

Working Group (IGN staff)

Specifications of new product (BD UNI v2)

The purpose was to make transformations to INSPIRE easier and of better quality.
Concerned INSPIRE themes

- Considered for BD UNI v2: AU, GN, TN, BU
- Themes AD, LC and HY have been considered
  - in other dedicated products
  - with external stakeholders
  - with different methodologies
- Theme US poorly considered for INSPIRE
  - IGN not referent data producer for electric lines
  - No big issues regarding governmental services
INSPIRE INFLUENCE
Avoid wrong transformations

- Railway Station example
  - In BD UNI v1, a point of interest (outside the network)
  - In INSPIRE, it may be a RailwayStationArea or a RailwayStationNode

- Matching table:
  - Correspondence between our POI and INSPIRE nodes
    - Key feature type in railway network
  - But does not fit with the INSPIRE definition

![Diagram showing correspondence between POI and RailwayStationNode]
Avoid wrong transformations

- Railway Station example
  - In BD UNI v2, decision to capture railway stations as areas
  - => correct matching with INSPIRE RailwayStationArea
Avoid loss of information

**Case 1: VerticalPosition**

- BD UNI richer than INSPIRE
- But we have what INSPIRE expects
- Not an issue
- No change

<table>
<thead>
<tr>
<th>Classe : Troncon de route</th>
<th>Classe : VerticalPosition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attr : Position par rapport au sol</td>
<td>Attr : verticalPosition</td>
</tr>
<tr>
<td>4</td>
<td>suspendedOrElevated</td>
</tr>
<tr>
<td>3</td>
<td>onGround</td>
</tr>
<tr>
<td>2</td>
<td>underground</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Gué ou radier</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td></td>
</tr>
</tbody>
</table>
Avoid loss of information

- Case 2: DesignSpeed

- What INSPIRE expects
- Distinction between ordinary trains and high speed trains
- DesignSpeed

- What we have in source data (BD UNI v1)
- No information at all about train speed
- What we can provide currently for INSPIRE

Rich
Poor
Avoid loss of information

- **Case 2: DesignSpeed**
  - BD UNI : we make distinction between
    - Train
    - High speed train
  - INSPIRE : DesignSpeed

- **IGN decision:**
  - No matching => lost of valuable information
  - We have included the DesignSpeed information in specification of new product BD UNI v2
  - We expect potential partnership to get this information
### Example: administrative hierarchy

<table>
<thead>
<tr>
<th>Commune</th>
<th>Attributes of « commune »</th>
<th>Attributes of the upper AU</th>
</tr>
</thead>
<tbody>
<tr>
<td>géometrie</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code canton</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code arrondissement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom arrondissement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code arrondissement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom arrondissement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code région</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom région</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- In existing data, IGN provides level 5 (commune) and attributes of upper levels are carried by “commune” => it is up to user to build upper levels

In source data, a key feature « Commune » - municipality
Make transformations easier

- **Example: administrative hierarchy**
  - INSPIRE requires a feature type for each level of AU
  - Current matching rules:
    - **Create new features** for upper level AU
    - Get their geometry by merging the geometries of lower level
    - Provide unique and persistent identifiers
Make transformations easier

- Example: administrative hierarchy

IGN has external identifiers for “Commune” ... but not for the upper levels
Decision was to use **thematic identifier** based on SHN (from EuroBoundaryMap) for all levels of AU =>
complex transformation because of some specificities (e.g. over-sea territories)

<table>
<thead>
<tr>
<th>AdministrativeUnits Element</th>
<th>Attribut</th>
<th>Type</th>
<th>Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier</td>
<td>localId</td>
<td>string</td>
<td>case 1 : ARRONDIS :</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;FR&quot;+ &quot;93&quot; + &quot;13&quot; + &quot;3&quot; + ExtractString(NUMINSEE,3,3) where NUMINSEE like &quot;13%&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;FR&quot;+ &quot;11&quot; + &quot;75&quot; + &quot;1&quot; + ExtractString(NUMINSEE,3,3) where NUMINSEE like &quot;75%&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;FR&quot;+ &quot;82&quot; + &quot;69&quot; + &quot;1&quot; + ExtractString(NUMINSEE,3,3) where NUMINSEE like &quot;69%&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 2 : COMMUNE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>if (dataset name contains “FR” or dataset name contains &quot;20&quot;) , &quot;FR&quot;+ INSEEEREG + INSEEDEP + INSEEARD + ExtractString(NUMINSEE,3,3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>if not (dataset name contains &quot;FR&quot; or dataset name contains &quot;20&quot;) , &quot;FR&quot;+ INSEEEREG + INSEEDEP + INSEEARD + ExtractString(NUMINSEE,4,2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 3 : Arrondissement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>if (dataset name contains “FR” or dataset name contains &quot;20&quot;) , &quot;FR&quot;+ INSEEEREG + INSEEDEP + INSEEARD + &quot;000&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>if not (dataset name contains &quot;FR&quot; or dataset name contains &quot;20&quot;) , &quot;FR&quot;+ INSEEEREG + INSEEDEP + INSEEARD + &quot;00&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 4 : Département</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>if (dataset name contains “FR” or dataset name contains &quot;20&quot;) , &quot;FR&quot;+ INSEEEREG + INSEEDEP + &quot;00000&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>if not (dataset name contains &quot;FR&quot; or dataset name contains &quot;20&quot;) , &quot;FR&quot;+ INSEEEREG + INSEEDEP + &quot;000&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 5 : Région</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;FR&quot;+ INSEEEREG + &quot;000000&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Case 6 : État</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>&quot;FR&quot;+ &quot;000000000&quot;</td>
</tr>
</tbody>
</table>
Example: administrative hierarchy

Current situation:

- Complex transformation
- Confusion between external identifier (inspireId) and thematic identifier

Decision for new product:

- Create a feature type for each level of AU
- Manage in production database a unique and persistent identifier for each feature
Pumping up our data model

- Case of geographical names
  - Lot of information related to geographical name(s) in source data

A feature may have several names in source data
Pumping up our data model

- **Case of geographical names**

  - **Current situation:**
    - Lot of information related to geographical name(s) in source data
    - Our old product is in traditional database
      - Fixed multiplicity for attribute values
      - Example:
        - name-1, name-1.status, …
        - name-2, name-2.status, ….
  
  - Named places are grouped in a theme “Points of Interest”
  
  - **Advantage:**
    - the complex set of attributes applies only to places having a name
    - Easy to manage (on production side)
  
  - **Drawback:**
    - The name is carried by a POI and not by the “true” feature
    - **Not user-friendly**, not in line with INSPIRE
Pumping up our data model

Case of geographical names

- Decision for new product:
  - Model close to INSPIRE
    - Named place
    - Carrying unlimited number of names
    - Names described by their spelling and by “metadata” attributes: language, source, status, ...

```xml
<featureType>
  NamedPlace
+ geometry: GM_Object
+ inspireId: Identifier
+ name: GeographicalName [1..*]
  «voidable, lifeCycleInfo»
+ beginLifespanVersion: DateTime
+ endLifespanVersion: DateTime [0..1]
  «voidable»
+ leastDetailedViewingResolution: MD_Resolution [0..1]
+ mostDetailedViewingResolution: MD_Resolution [0..1]
+ localType: LocalisedCharacterString [1..*]
+ type: NamedPlaceTypeValue [1..*]
+ relatedSpatialObject: Identifier [0..*]
</featureType>
<dataType>
  GeographicalName
+ spelling: SpellingOfName [1..*]
  «voidable»
+ language: CharacterString
+ nativeness: NativenessValue
+ nameStatus: NameStatusValue
+ sourceOfName: CharacterString
+ pronunciation: PronunciationOfName
+ grammaticalGender: GrammaticalGenderValue [0..1]
+ grammaticalNumber: GrammaticalNumberValue [0..1]
</dataType>
```
Pumping up our data model

Plusieurs solutions de modélisation dans la BDUni v2 :

2) Créer un champ unique ‘JSON’

Champ JSON : champ de type clé-valeur avec saisie d’un nombre de toponymes illimité

<table>
<thead>
<tr>
<th>Graphie</th>
<th>Origine</th>
<th>Date de validation</th>
<th>Statut</th>
</tr>
</thead>
<tbody>
<tr>
<td>le vieil armand</td>
<td>BDTopo</td>
<td>31/07/1994</td>
<td>Classique</td>
</tr>
<tr>
<td>hartmannswillerkopf</td>
<td>BDTopo</td>
<td>31/07/1994</td>
<td>Régional</td>
</tr>
<tr>
<td>Hartmannswillerkopf (le vieil armand)</td>
<td>BDTopo</td>
<td>31/07/1994</td>
<td>Cartographique</td>
</tr>
<tr>
<td>au vieil armand</td>
<td>EJN</td>
<td></td>
<td>Cadastral</td>
</tr>
<tr>
<td>le vieil armand ou hartmannswillerkopf</td>
<td>SDIS</td>
<td></td>
<td>Partenariale</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Use of JSON attributes

New tools to be developed to capture and manage this kind of attributes

POI

```json
{  Importance : 8,
  CDB_Sélection : GE,
  Commentaire : Blabla,
  Toponymie : [
    { Statut : Validé,
      Graphie : mon lieu-dit,
      Source : BDTopo,
      Date : 26/05/2010,
      Validation demandée : -,
      Id_partenaire : ... }
    { Statut : Collecté,
      Graphie : cet endroit,
      Source : Mairie,
      Date : 26/05/2010,
      Validation demandée : oui,
      Id_partenaire : ... }
  ], ...
}
```

Hierarchic structure in our new product!

STATUT

Validé
Collecté
Partenariale
Régional
Autre
(BAN ?)
Enrich our data model

- **Example: Buildings**
  - **Current situation:**
    - INSPIRE requires
      - current use - number of dwellings -....
      - date of construction - material of roof
      - number of floors - material of structure
    - This information is also required by our users
    - But is not or poorly available in our current product
Enrich our data model

- Example: Buildings
  - Decision for new product
    - These attributes are considered as core information
    - Include these attributes in data model
  - Struggle to get source information
    - Data available in land registry (Cadastre)
    - Integration test was performed
      - technical difficulties to match IGN buildings with land registry ones
      - privacy issues
LEARNINGS AND CONCLUSIONS
# Modelling approach

- **Data model prepared by Excel tables**

<table>
<thead>
<tr>
<th>Bâti</th>
<th>Besoin(s)</th>
<th>Valeurs de la 'Désignation'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inspire / MGCP</td>
<td>&lt;sans valeur&gt; réellement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>idem BDUni</td>
</tr>
<tr>
<td></td>
<td></td>
<td>idem BDUni</td>
</tr>
<tr>
<td></td>
<td></td>
<td>remplace &lt;sans valeur&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inspire + MGCP (abandonné, détruit, démantelé, endommagé) ... MAIS pas si facile...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Métadonnées d'appariement concaténées : TA, TX, TY, Id Parcelle, Type BDP, Anc. SG2D, Anc. CLEABS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Année de la source de la donnée : pertinent pour les nouvelles données.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bâti</th>
<th>Valeurs de 'Nature' Inspire supplémentaires :</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[Auvent]</td>
</tr>
<tr>
<td></td>
<td>[Hangar]</td>
</tr>
<tr>
<td></td>
<td>[Mosquée] &gt;&gt; PAI BDUni</td>
</tr>
<tr>
<td></td>
<td>[Synagogue] &gt;&gt; PAI BDUni</td>
</tr>
<tr>
<td></td>
<td>[Temple] &gt;&gt; PAI BDUni</td>
</tr>
<tr>
<td></td>
<td>[Habitation troglodytique] &gt;&gt; PAI BDUni</td>
</tr>
</tbody>
</table>

- **Data model prepared by Excel tables**

<table>
<thead>
<tr>
<th>Bâti</th>
<th>Besoin(s)</th>
<th>Valeurs de la 'Désignation'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature Liste</td>
<td>Inspire / MGCP</td>
<td>&lt;sans valeur&gt; réellement</td>
</tr>
<tr>
<td></td>
<td>idem BDUni</td>
<td></td>
</tr>
<tr>
<td></td>
<td>idem BDUni</td>
<td></td>
</tr>
<tr>
<td></td>
<td>remplace &lt;sans valeur&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inspire + MGCP (abandonné, détruit, démantelé, endommagé) ... MAIS pas si facile...</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Métadonnées d'appariement concaténées : TA, TX, TY, Id Parcelle, Type BDP, Anc. SG2D, Anc. CLEABS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Année de la source de la donnée : pertinent pour les nouvelles données.</td>
<td></td>
</tr>
</tbody>
</table>

- **Data model prepared by Excel tables**
Modelling approach

- Why no UML model?
  - Not in the missions of the Working Group
    - Mission was to decide on content
  - Not (yet) in the IGN culture
    - UML is not a “reflex” among IGN staff
  - UML model not seen as useful
    - New product in simple structure
      - No inheritance
      - Few associations
    - => graphical representation not so useful
  - May come in future
Flexibility regarding INSPIRE

- INSPIRE has significantly influenced the design of our new product.
- But there will remain many differences or even discrepancies between BD UNI v2 and INSPIRE.
  - Repartition in themes is not the same.
    - Example: Ferry crossings are
      - in Water Transport Network in INSPIRE.
      - In Road Transport Network in BD UNI v2.
  - Missing attributes, additional ones.
  - ....
#### Why adopting INSPIRE (sometimes)

<table>
<thead>
<tr>
<th>Reason</th>
<th>Examples</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoid « wrong » transformations. Ensure minimum quality of INSPIRE data. INSPIRE as reasonable constraint</td>
<td>Railway station captured as area (instead of POI)</td>
<td>😞</td>
</tr>
<tr>
<td>INSPIRE helps us to « push » user requirements. INSPIRE as an opportunity.</td>
<td>Enrichment of theme Buildings Processing of Geographical Names</td>
<td>😊</td>
</tr>
</tbody>
</table>
Why not (always) adopting INSPIRE

- INSPIRE not the main driver;
  - Main driver: user requirements
  - Starting point was specification of old product and not the INSPIRE data models

- No need to adopt INSPIRE “natively” in production if transformations don’t raise issues

- INSPIRE not always seen as good practice

  - Example: Transport Network
    - in INSPIRE, transport properties are feature types attached by linear referencing to the transport objects
    - In our source data, transport properties are attributes directly carried by the transport objects
    - Easier to manage in production and to use by GIS
    - => INSPIRE modelling approach was not adopted
Why not (always) adopting INSPIRE

- Take into account production constraints:
  - INSPIRE does not mandate capture of new data
  - But INSPIRE pushed us to enrich our new product (e.g. BU)
  - Enrichments limited to:
    - What is considered as useful
    - What is considered as (more or less) feasible, e.g. more collaborative capture or search for new partnership
  - More flexible specifications
    - Core content: with some quality measure and guarantee
    - Extended content: included in the model but no guarantee
Étapes suivantes

- Production d’une version test des données:
  - Migration des données dans le nouveau modèle BD UNI v2
    - France entière
  - Enrichissements en cours:
    - Collecter: partenariat avec le cadastre pour le thème BU
    - Calculer : ex: attribut vitesse moyenne sur tronçons de route
    - Exposer des attributs internes (ex toponymes variés) => mise à niveau
    - Exposer des attributs vides (ex: DesignSpeed) dans l’attente d’une source de données
Étapes suivantes

- **Validation des nouvelles spécifications**
  - Enquête qualitative
    - Entretiens avec un petit nombre d’utilisateurs
    - Juin 2017
    - => les décisions vont dans le bon sens
    - => besoin de documenter le passage BD UNI v1 vers BD UNI v2
  - Enquête quantitative :
    - En cours
    - Questionnaire en ligne
Étapes suivantes

- Préparation des produits externes
  - Simplification par rapport à la méthode précédente
    - Produit externe = vue, sous-ensemble de la base de production
  - Organisation en thèmes proches d’INSPIRE
    - BD UNI v1 : thème « fourre-tout » sur les points d’intérêt
    - BD UNI v2 : les points d’intérêt sont répartis dans leurs thèmes
      - Services, activités: US, PF, AF
      - Toponymes: GN