



# Analyse & évaluation des infrastructures de données géographiques par l'analyse des catalogues via l'application DAOBS

Etienne Taffoureau, François-Xavier Prunayre

## ► To cite this version:

Etienne Taffoureau, François-Xavier Prunayre. Analyse & évaluation des infrastructures de données géographiques par l'analyse des catalogues via l'application DAOBS. FOSS4G 2016 : Rencontre francophone des développeurs et utilisateurs de logiciels libres dédiés à la géomatique, OSGeo-fr, May 2016, Marne-La-Vallée, France. hal-01317487

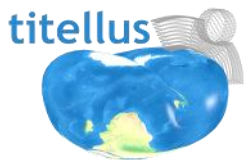
**HAL Id: hal-01317487**

**<https://brgm.hal.science/hal-01317487>**

Submitted on 18 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Analyse & évaluation des IDG par l'analyse des catalogues

Comment évaluer le déploiement des IDG en France et l'évolution de la mise en œuvre de la directive INSPIRE ?  
Calcul d'indicateurs et analyse de données avec Daobs

Etienne Taffoureau – BRGM  
François-Xavier Prunayre - titellus



# Contexte du projet DAOBS

- > Outil permettant de réaliser des tableaux de bord
- > Techniquement, basé sur les composants Solr (moteur de recherche) et Banana (tableaux de bord)
- > Dans le cadre d'un groupe de travail INSPIRE en 2014-2015 (Maintenance & Implementation group)
- > Composé d'experts de plusieurs pays : FR, NL, DE, UK, IT, PL, SE...
- > Piloté par l'Agence Européenne de l'Environnement (EEA)
- > Dashboard européen :
  - <https://inspire-dashboard.eea.europa.eu/official>
  - <https://inspire-dashboard.eea.europa.eu/dashboard>

# Objectifs visés

## Simplifier le rapportage INSPIRE annuel

- Calcul d'indicateurs automatique à partir des catalogues
- Génération d'un fichier XML standard (au lieu de fichier Excel)

## > Suivre l'évolution des indicateurs dans des tableaux de bord INSPIRE



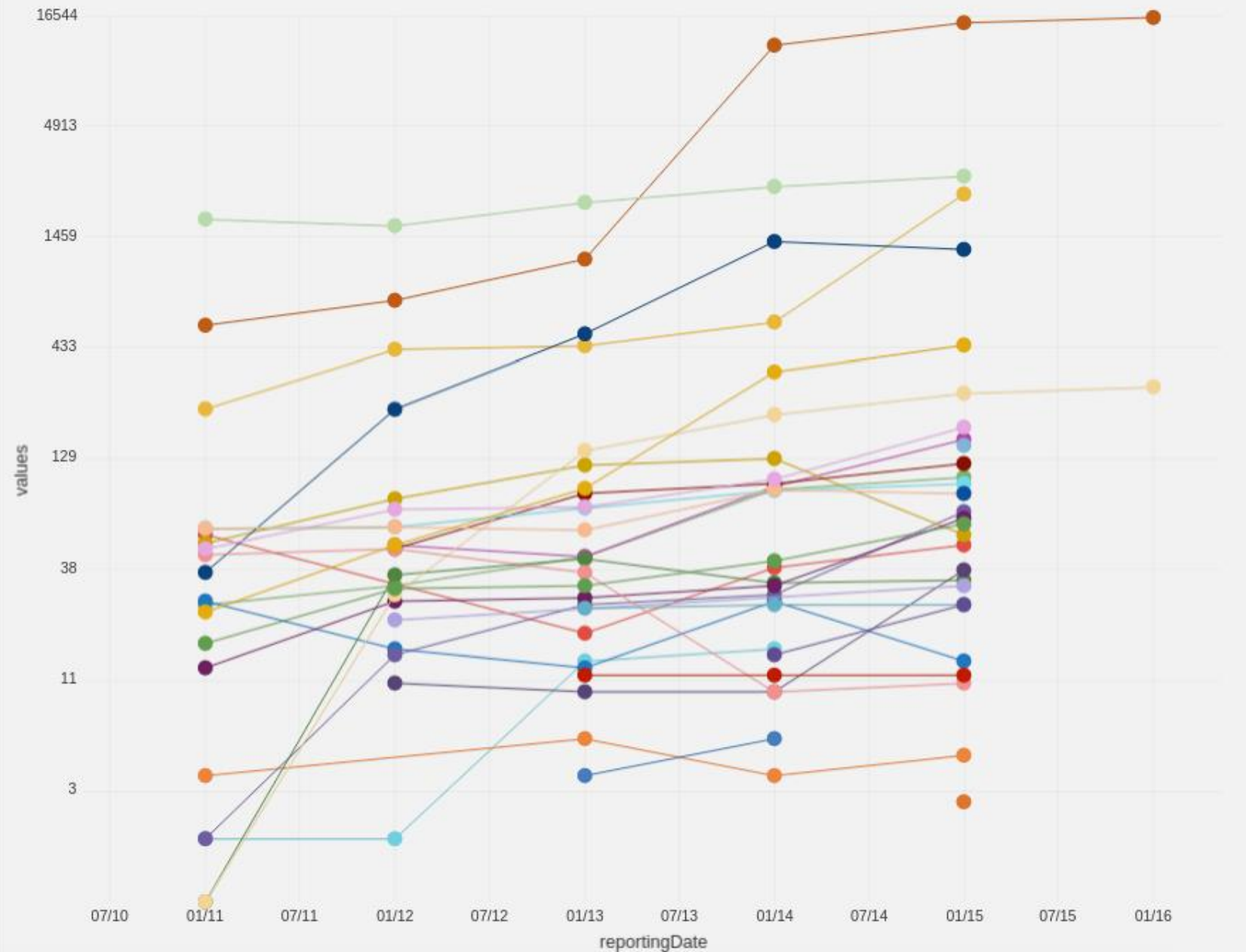
# NUMBER OF NETWORK SERVICES REPORTED




 HISTOGRAM

AT (5) BE (5) BG (4) CH (1) CY (3) CZ (5) DE (6) DK (5) EE (5) EL (5) ES (5) FI (5) FR (5) HR (2) HU (4) IE (4) IS (3) IT (5)  
 LI (3) LT (5) LU (5) LV (5) MT (3) NL (6) NO (5) PL (5) PT (5) RO (5) SE (5) SI (5) SK (5) UK (5) indicatorValue values per 1y | (144 hits) |

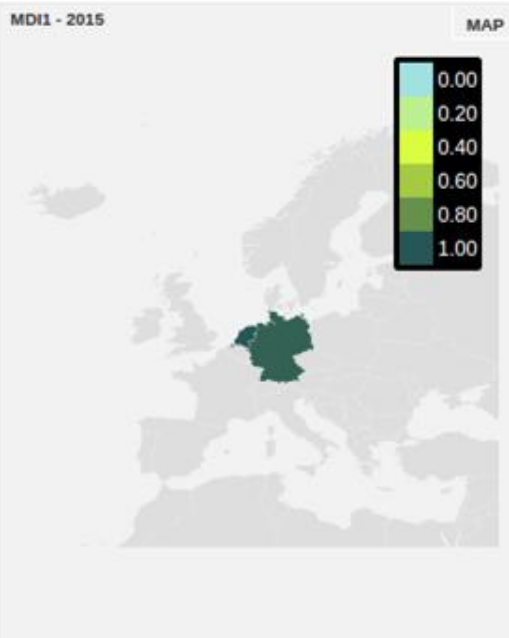
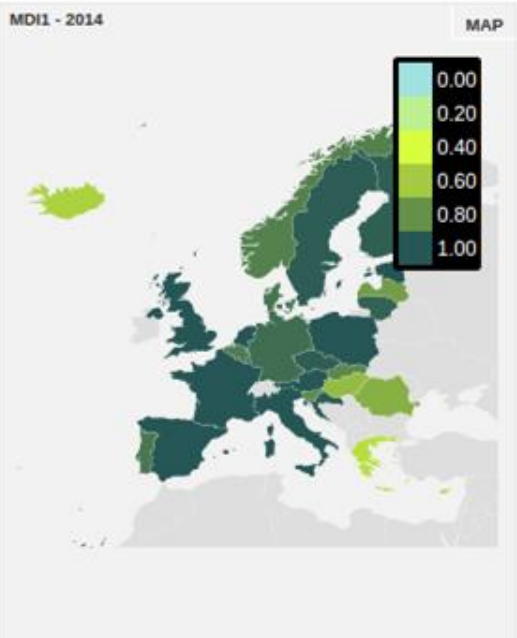
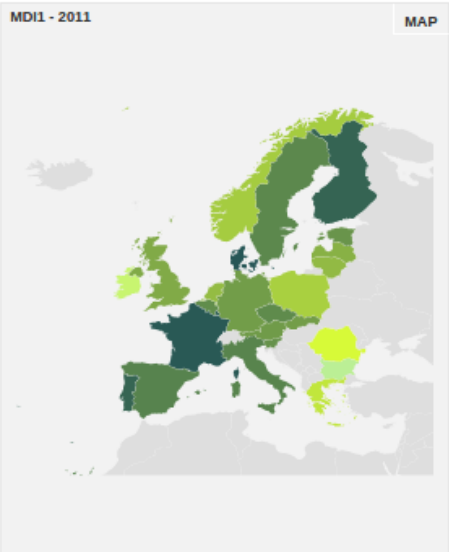
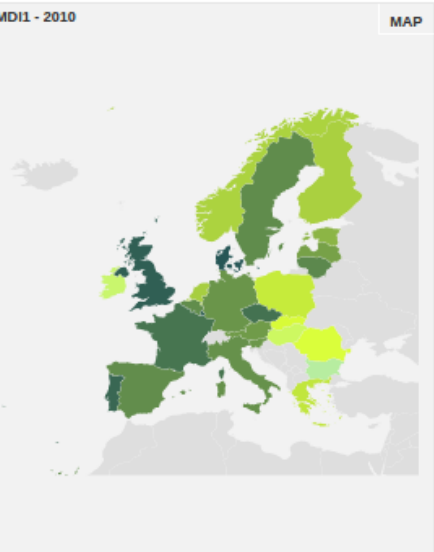
Time correction : browser



ble

# Existence of metadata (spatial data sets)

- Existence of metadata (data sets and services) - MDi1



# Contexte du projet DAOBS

## > Dashboard européen :

- <https://inspire-dashboard.eea.europa.eu/official>
- <https://inspire-dashboard.eea.europa.eu/dashboard>

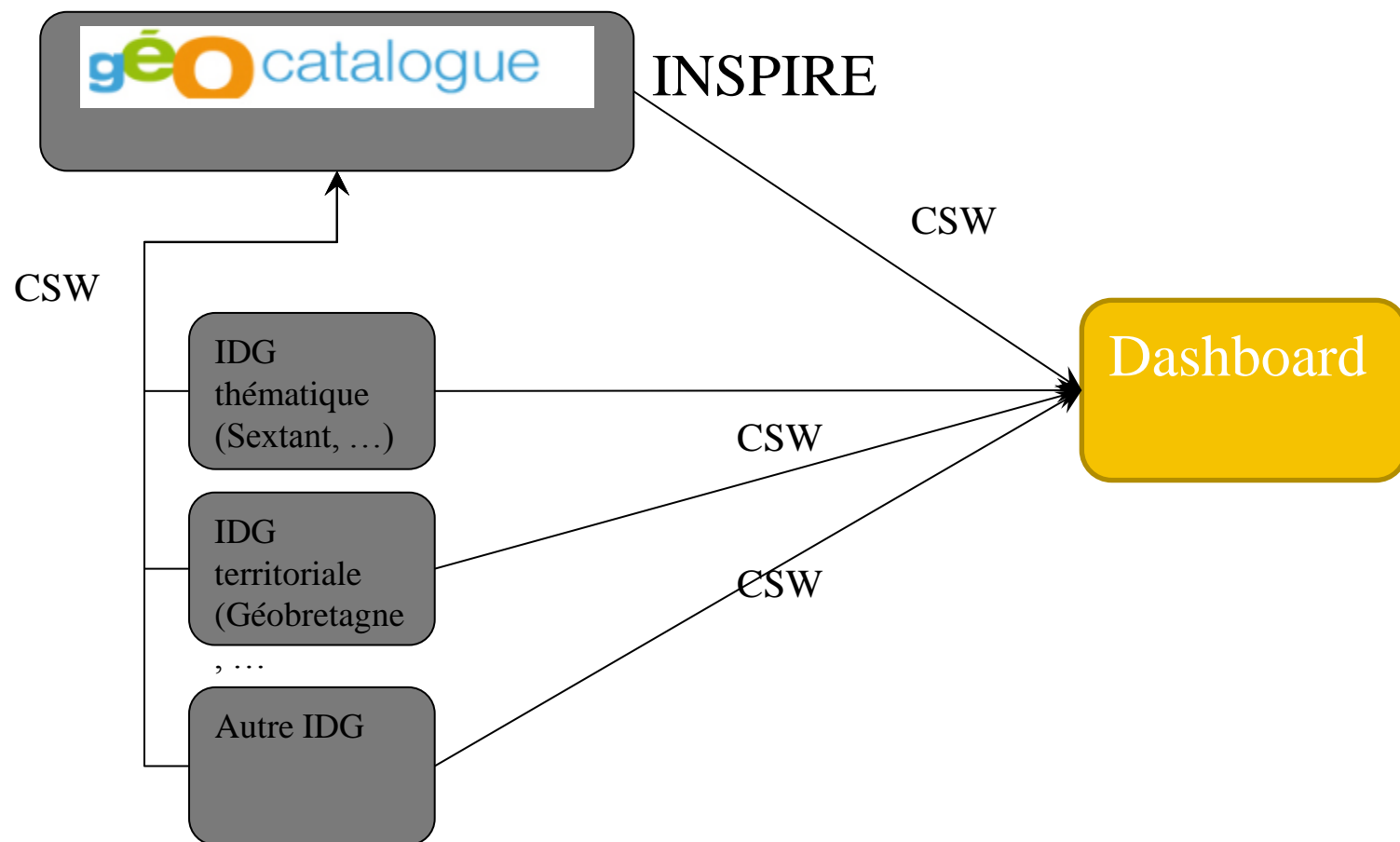
## > Projet opensource publié sur Github :

- <https://github.com/INSPIRE-MIF/daobs>
- Licence EUPL

## > Développé par titellus, Geocat.BV, CapGemini



# Principes du dashboard / #1 Moissonnage



> Tableaux de bord construits à partir des métadonnées ISO 19139 accessibles via des points d'accès CSW (=harvest)



# Principes du dashboard / #2 Analyses

- > Chaque fiche de métadonnées fait l'objet d'un document Solr**
- > Les informations exploitées par les tableaux de bord sont stockées dans des champs dédiés dans l'index Solr :**
  - Contenu des métadonnées ISO 19139 (e.g. gmd:title)
  - Calculés à partir des métadonnées (e.g. mot-clé « Opendata »)
  - Métadonnées liées (recordOperatedBy)
  - Synonymes issus de vocabulaires contrôlés (par exemple pour la gestion des producteurs)

# Principes du dashboard / #2 Analyses

## eg. utilisation des vocabulaires contrôlés

.thesaurus.external.parameter.NVS.P01	Q ⓘ III	Northward current velocity in the water body
.thesaurus.external.parameter.NVS.P02	Q ⓘ III	Horizontal velocity of the water column (currents)
.thesaurus.external.parameter.NVS.P03	Q ⓘ III	Currents
.thesaurus.external.theme.Inspire-theme	Q ⓘ III	Oceanographic geographical features
.thesaurus.local.theme.medsea.challenges	Q ⓘ III	Oil platform leaks
.thesaurus.local.theme.medsea.data.delivery.mechanism	Q ⓘ III	On-line downloading services
.thesaurus.local.theme.medsea.environmental.matrix	Q ⓘ III	Marine water
.thesaurus.local.theme.medsea.level.of.characteristics	Q ⓘ III	Forecast/Hindcast
.thesaurus.local.theme.medsea.production.mode	Q ⓘ III	Real-time

# Principes du dashboard / #2 Analyses

**eg. restriction aux listes de valeurs INSPIRE**



# Principes du dashboard / #2 Analyses

## eg. analyse synonymique

Ortofotosnímky ➡

Orthoimagery > Annex II



inspireTheme=Ortofotosnímky  
inspireTheme\_syn=Orthoimagery  
inspireAnnex=II

```
"Ορτοφωτογραφία": ["Orthoimagery"],  
"Orthofotografie": ["Orthoimagery"],  
"Ortofotosnímky": ["Orthoimagery"],  
"Ортоизображение": ["Orthoimagery"],  
"http://rdfdata.eionet.europa.eu/inspirethemes/themes/12": ["Orthoimagery"],  
"Geologia": ["Geology"],  
"Geologie": ["Geology"],  
"Geológia": ["Geology"],  
"Geology": ["Geology"]
```

```
"Coordinate reference systems": ["I"],  
"Elevation": ["II"],  
"Land cover": ["II"],  
"Orthoimagery": ["II"],
```

Index

ManagedSynonymFilterFactory

Synonyms



Géosciences pour une Terre durable

**brgm**

# Principes du dashboard / #2 Analyses

## eg. analyse synonymique

DDT 42

DDT42

DDT\_42

DDT de la Loire

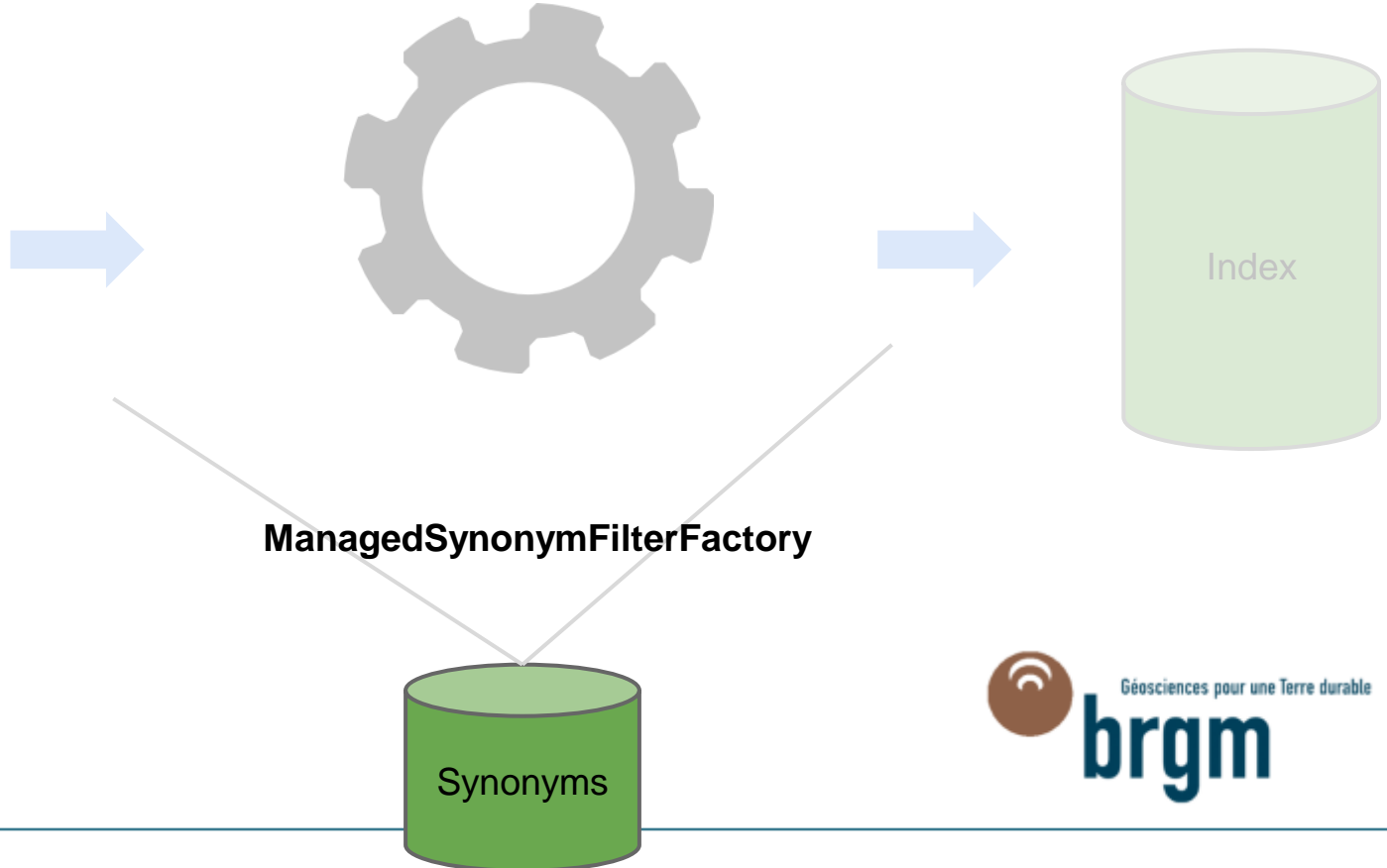
DDEA 42

DIREN de la Loire

➡ Direction Départementale des  
Territoires de la Loire (DDT 42)

➡ Département de la Loire (42)

ISO19139  
metadata  
record



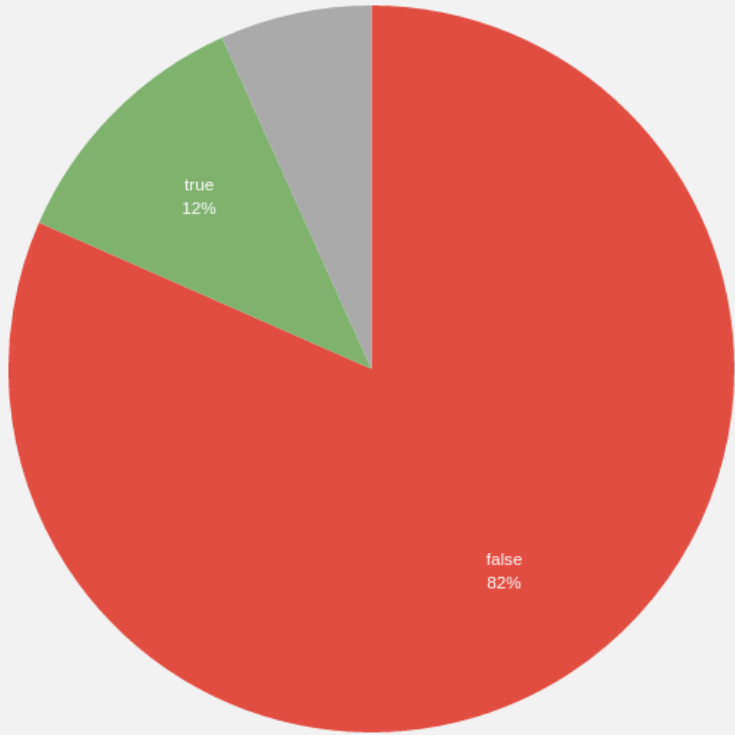
# Principes du dashboard / #2 Analyses

**Couplage avec d'autres outils :**

- >ETF pour la validation des services**
- >Valdateur du JRC ou du géocatalogue.fr pour la validation des métadonnées**

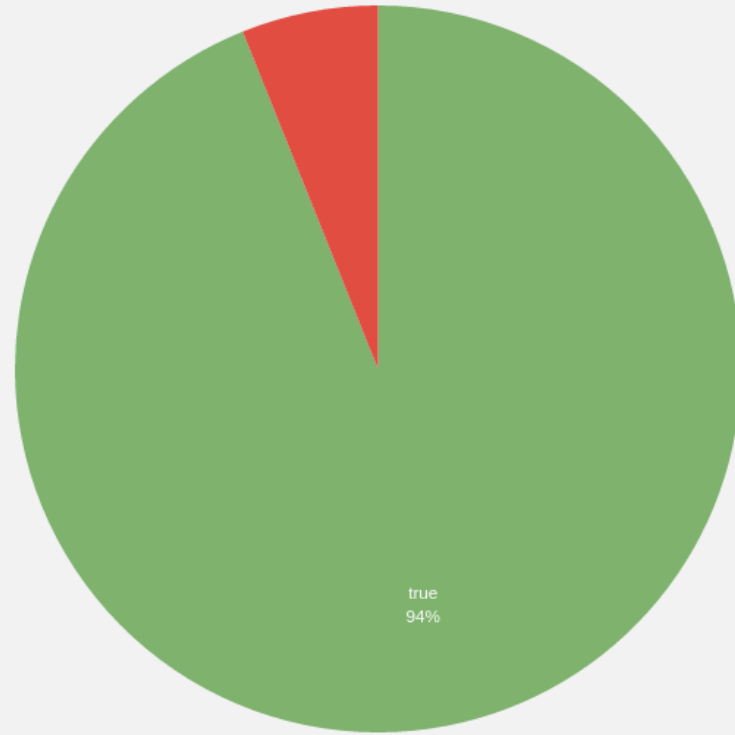
INSPIRE VALIDATION REPORT

● false (110,184) ● true (15,738) ● Missing field (9,119)



IS ABOVE 75% FOR COMPLETENESS INDICATOR

● true (118,267) ● false (7,655)



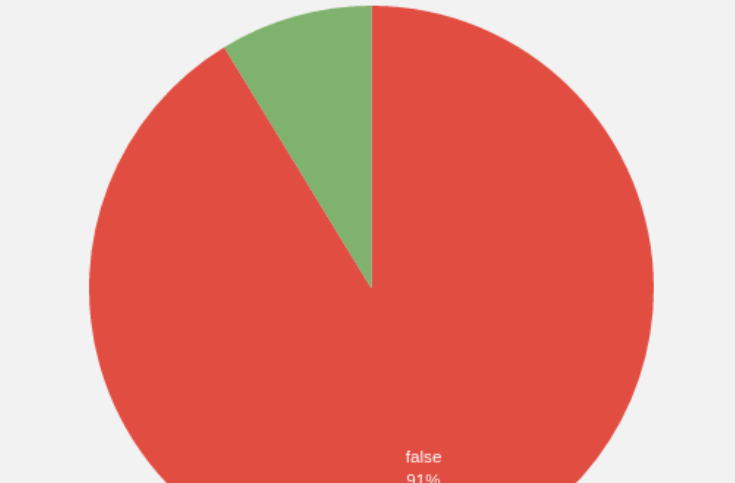
Validation is computed using the [INSPIRE validator](#).

A record is considered valid for INSPIRE if completeness indicator is above 75% (see `isAboveTheshold` field).

Even if completeness indicator is 100%, a record may be displayed as invalid in the INSPIRE validation chart due to INSPIRE validation warnings (eg. if the URL of a coupled resource point to an invalid link).

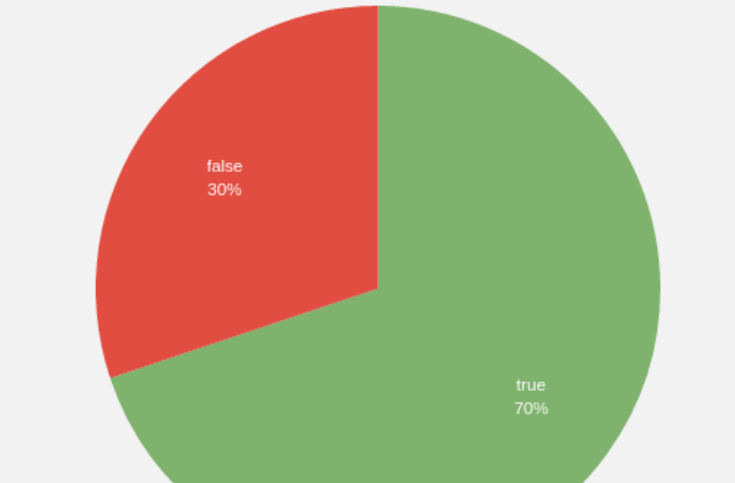
CONFORMANT DATA SETS

● false (40,211) ● true (3,847)



CONFORMANT SERVICES

● true (357) ● false (154)



Conformity of datasets and services is declared in the metadata records.

*pour l'ensemble des états membres*



# Validation des services

## Service validation report

### SEARCH

QUERY

TOTAL HITS

598

Load

HITS

range must

field : etfCompletenessIndicator

from : -1

to : 100

range must

field :  
etfCompletenessIndicatorOptional

from : -1

to : 100

FILTERING

ETF MANDATORY

● false (264) ● true (35)  
● Missing field (0)



ETF OPTIONAL

● false (292) ● true (7)

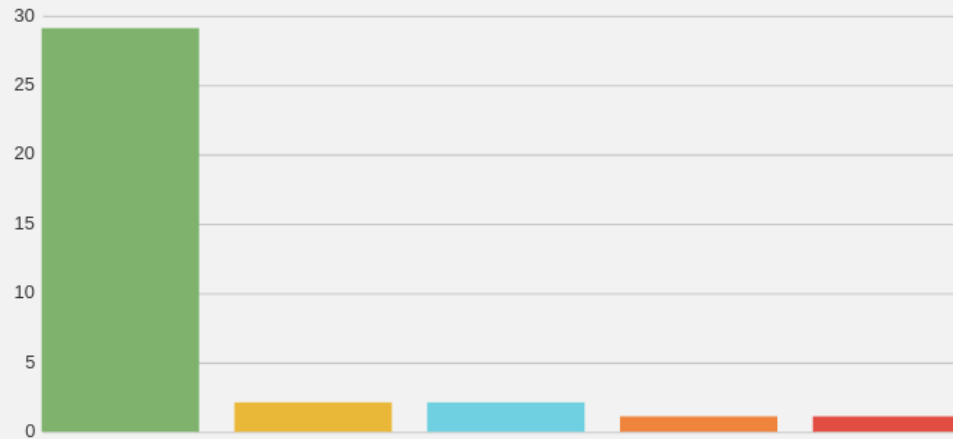


TEXT

### COMPLETENESS INDICATOR

Note that the completeness indicator currently weighs all tests equally. Please, consider that failures in some tests may have much more impact as failures in other tests. So use the completeness Indicator with caution

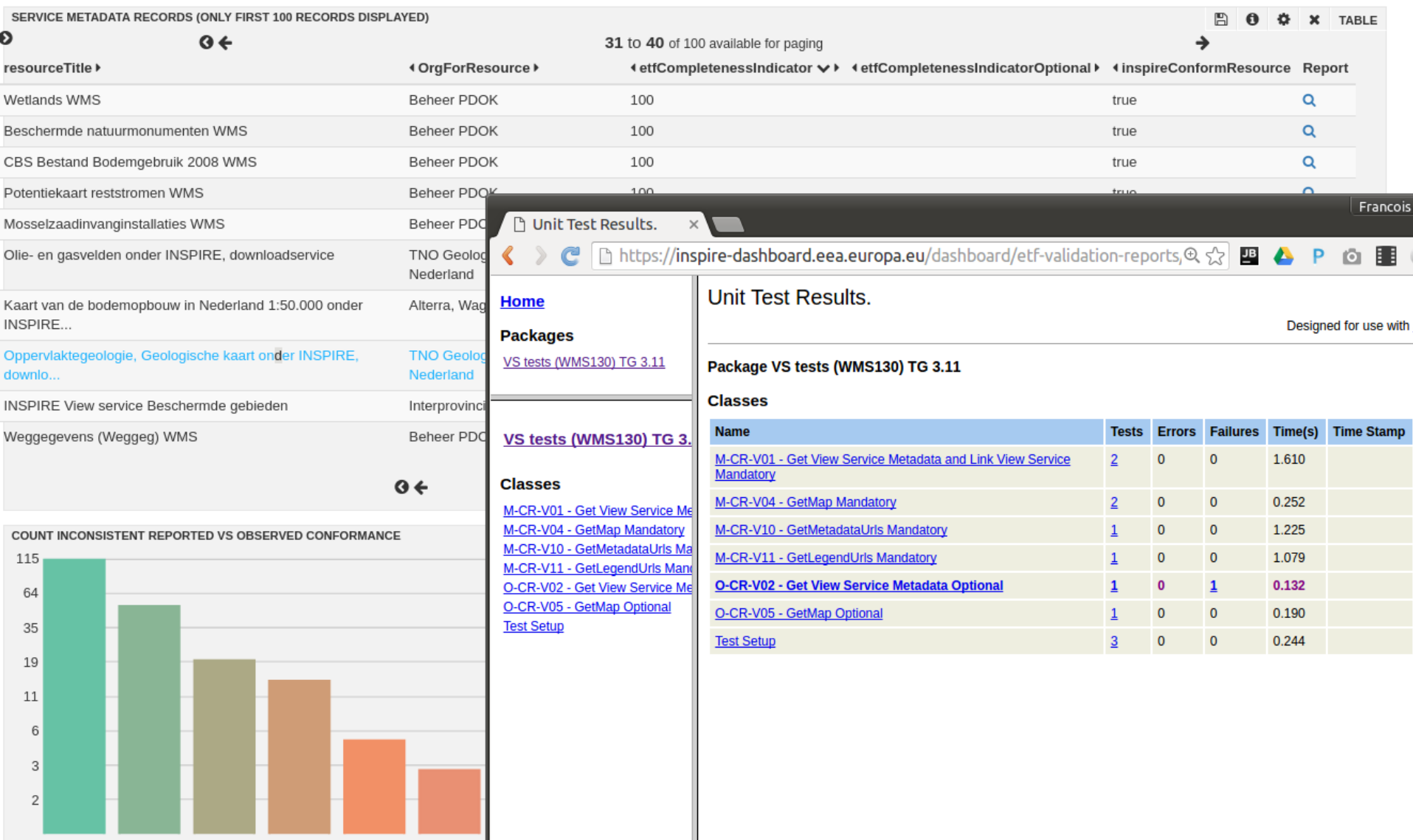
### VALID SERVICES PER ORGANISATION



- Beheer PDOK
- Royal Netherlands Meteorological Institute (KNMI)
- Alterra, Wageningen UR
- TNO Geologische Dienst Nederland
- Het Waterschapshuis

TERMS

# Validation des services



# Principes du dashboard / #3 Calcul d'indicateurs

## > Calcul d'indicateurs à partir de l'index Solr

## > Calculés selon des règles de calcul prédéfinies

- MDv11\_al\_t1 = Number of spatial data sets for theme Coordinate reference systems that have metadata = +(resourceType:dataset resourceType:series) +inspireTheme:"coordinate reference systems"

## > Possibilité de faire des filtres :

- Par point de moissonnage (=par IDG)
- Sur un champ de l'index (e.g. par producteur)

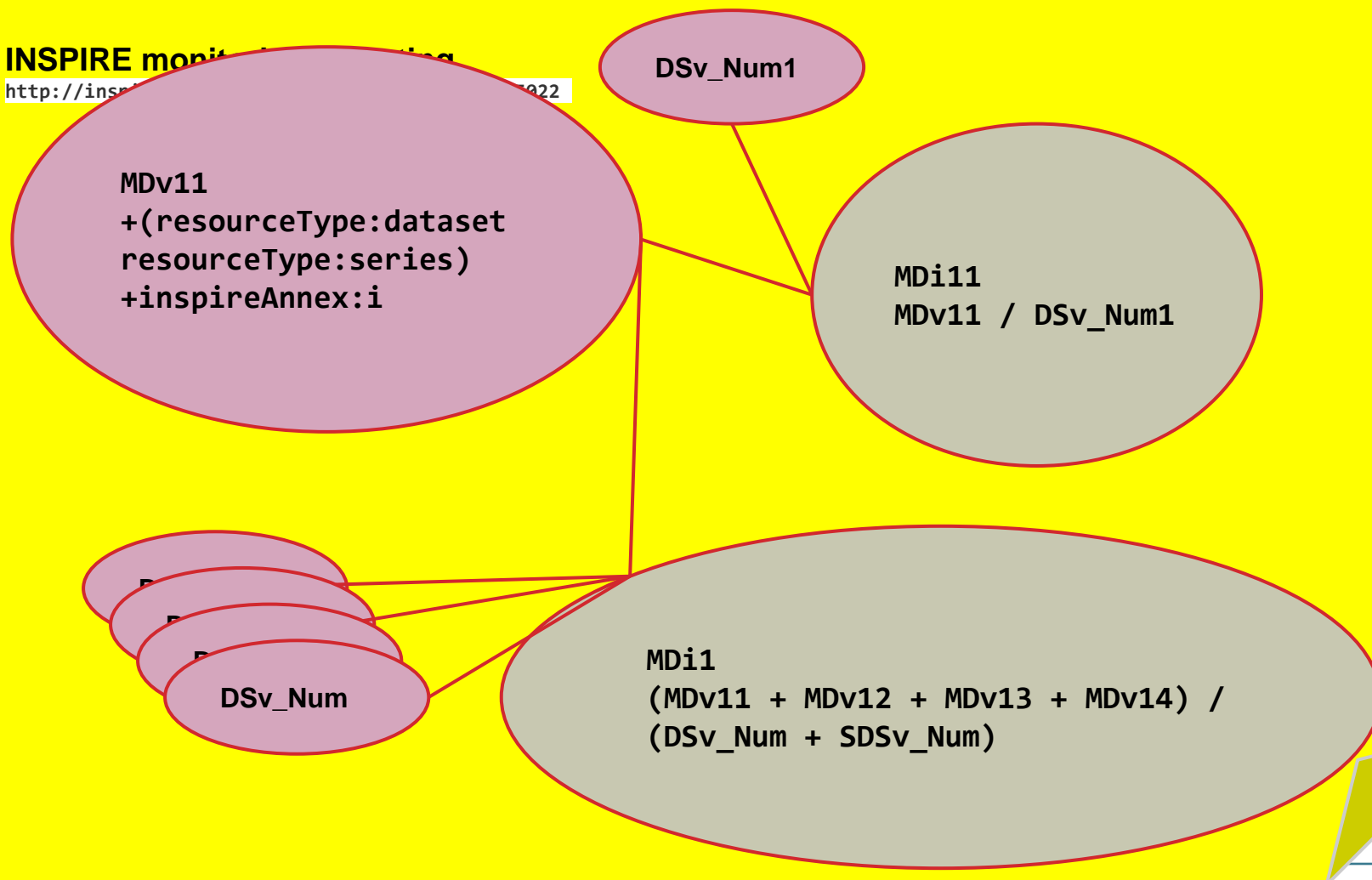
## > Export de rapport au format XML

- INSPIRE monitoring format (indicator only, with row data)
- JSON, ...

## > Import de rapport pour intégration dans les tableaux de bord « officiels »

# Rapport INSPIRE / Exemple d'indicateur

INSPIRE monitoring indicator  
<http://inspire.europa.eu/inspire/indicators/22>



Choose monitoring

inspire

View report rules

Choose reporting area

NL

Query filter

eg. to select only dataset use '+resourceType:dataset', to select based on point of contact use '+OrgForResource:RIVM'.

Contact details

org

Email

Contact informations are required when generating INSPIRE reporting to indicate report contact.

417 record(s) matching current filter.

Preview

Download

Schedule (not available yet)

## Report on Monitoring Indicators generated for NL based on 417 record(s)

The table below shows the list of variables and indicators computed for the selected area. It may also reports any errors encountered during computation.

	Name	Value
MDv		
MDv11	Number of spatial data sets for Annex I that have metadata	55
MDv12	Number of spatial data sets for Annex II that have metadata	7
MDv13	Number of spatial data sets for Annex III that have metadata	104
MDv14	Number of spatial data services that have metadata	255
MDv_DS	Number of spatial data sets that have metadata	166
MDv21	Number of spatial data sets for Annex I that have conformant metadata	0
MDv22	Number of spatial data sets for Annex II that have conformant metadata	0
MDv23	Number of spatial data sets for Annex III that have conformant metadata	0
MDv24	Number of spatial data services that have conformant metadata	0
MDv2_DS	Number of spatial data sets for all Annexes that have conformant metadata	0

Règles de configuration des rapports

INSPIRE Monitoring



Géosciences pour une Terre durable

brgm

# Principes du dashboard / #4 Visualisation

## > Utilisation au niveau national

- Publication des indicateurs INSPIRE
- Calcul de nouveaux indicateurs & analyse par plate-forme

## > Amélioration du projet

- Internationalisation de l'interface
- Analyse des producteurs
- Création de cartes au niveau national (régionales, départementales)
- Suivi des indicateurs à intervalle régulier

# Interface web de création de tableaux de bords

## Indicateurs INSPIRE 2016 - Accessibilité des données



### INDICATEURS INSPIRE 2016 - ACCESSIBILITÉ DES DONNÉES

Ce tableau de bord présente l'évolution de l'indicateur sur l'accessibilité des données entre 2013 et 2016

Autres indicateurs disponibles :

[Existence des métadonnées](#)

[Conformité des métadonnées](#)

[Conformité des données](#)

[Conformité des services](#)

[Utilisation des services en réseau](#)

[Statistiques globales](#)

### FENÊTRE DE TEMPS

1y

2y

3y

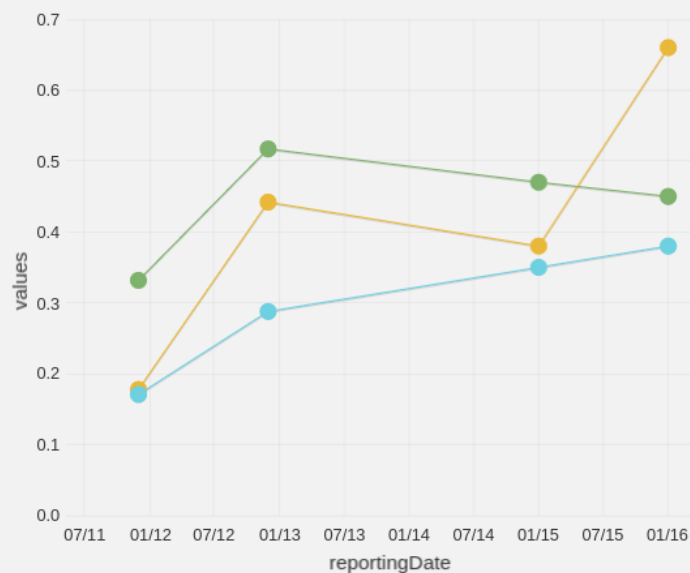
4y

5y

[Relative](#) | [Absolute](#) | [Since](#) | ☐ Auto-refresh

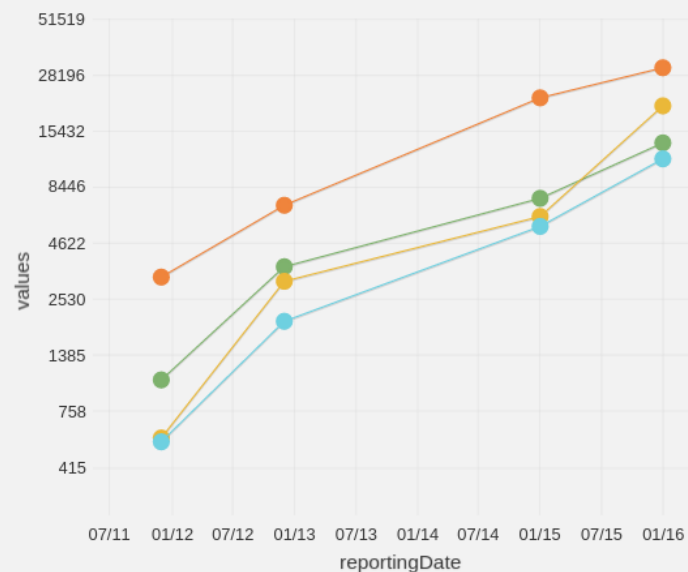
### POURCENTAGE DE DONNÉES CONSULTABLES (NSI21) ET/OU TÉLÉCHARGEABLES (NSI22)

● nsi21 (4) ● nsi22 (4) ● nsi2 (4) indicatorValue values per 1y | (12 Hits) | Time correction: browser



### NOMBRE DE DONNÉES CONSULTABLES (NSV21) ET/OU TÉLÉCHARGEABLES (NSV22) ET NOMBRE TOTAL DE DONNÉES (MDV1\_DS)

● mdv1\_ds (4) ● nsv21 (4) ● nsv22 (4) ● nsv23 (4) indicatorValue values per 1y | (16 Hits) | Time correction: browser





# Création à partir de composants configurables

General Panels Add Panel

## Row Settings

Select Panel Type

- bettermap
- column
- facet
- filtering
- fullTextSearch
- heatmap
- histogram
- hits
- map
- multiserries
- query
- rangeFacet
- scatterplot
- table
- tagcloud
- terms
- text
- ticker
- timepicker

Close

values

125

48

18

7

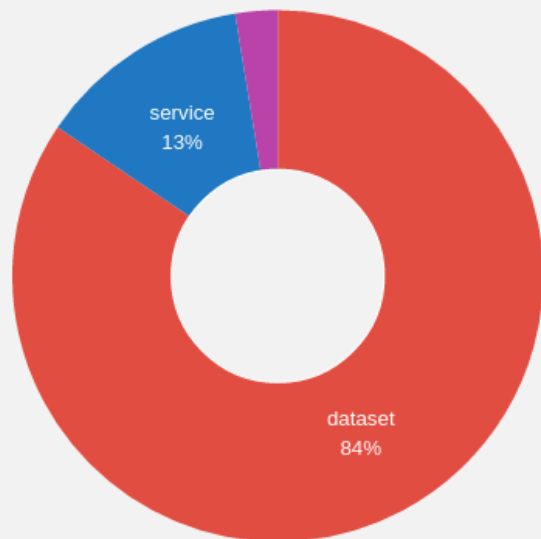
3

# Découpage par type de ressources, séries de données et services

TYPE OF RESOURCES

TERMES

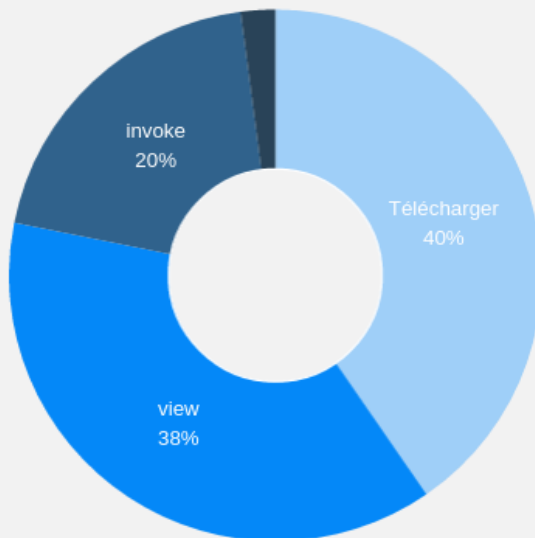
● dataset (25,700) ● service (3,950)  
● series (780) ● feature (2)  
● nonGeographicDataset (1)



TYPE OF SERVICES

TERMES

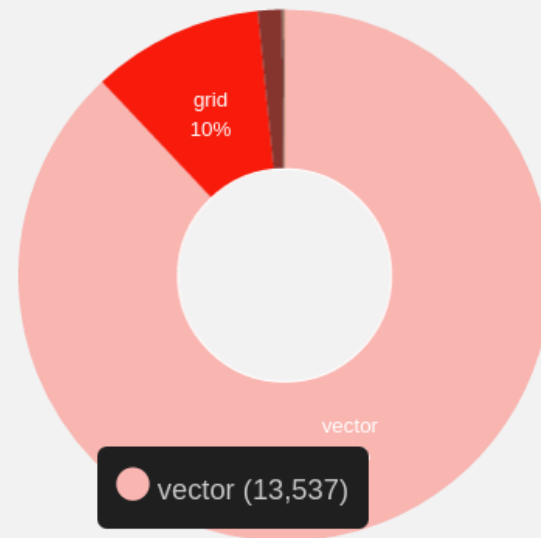
● Télécharger (1,531) ● view (1,432)  
● invoke (748) ● discovery (79)



TYPE OF DATASETS

TERMES

● vector (13,537) ● grid (1,602)  
● textTable (223) ● tin (18) ● Vector (6)



# Analyse par territoire – quelques exemples

Nb de fiches par région



Nb de fiches par département



Nb de données  
ouvertes par région



Nb de données  
téléchargeables par  
département



# Perspectives

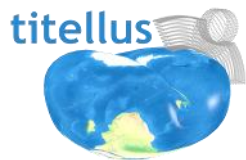
## **Amélioration du moissonnage**

- lancement à une fréquence prédéfinie**
- rapport de moissonnage / historique**

## **Calcul des indicateurs « en temps réel »**

## **Supervision des fluctuations de contenu entre les catalogues**

## **Recherche des doublons/similarités/différences entre les catalogues**



**Merci de votre attention**

Etienne Taffoureau – BRGM  
François-Xavier Prunayre - Titellus



# Technical overview

