



HAL
open science

Données à référence spatiale et entrepôts de données

Yuji Kato

► **To cite this version:**

Yuji Kato. Données à référence spatiale et entrepôts de données. Open Access Week en Bourgogne Franche-Comté, Dat@UBFC; Maison des sciences de l'homme et de l'environnement; Institut agro Dijon; Service commun de la documentation de l'université de Bourgogne; Service commun de la documentation de l'université de Franche-Comté; Maison des sciences de l'homme de Dijon; Presses universitaires de Franche-Comté; Éditions universitaires de Dijon, Nov 2023, Belfort, Besançon et Dijon, France. hal-04858252

HAL Id: hal-04858252

<https://hal.science/hal-04858252v1>

Submitted on 29 Dec 2024

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.



Distributed under a Creative Commons Attribution 4.0 International License

SEMAINE DE LA
**SCIENCE
OUVERTE**
EN BOURGOGNE
FRANCHE-COMTÉ
13 AU 17 NOVEMBRE 2023

Visios & Ateliers



DONNÉES à RÉFÉRENCE SPATIALE & ENTREPÔTS de DONNÉES

Yuji KATO
MSHE Besançon

Jeudi 16 novembre 2023 – 13h00



Plan

1. Science ouverte et données de la recherche
2. Entrepôts de données de la recherche
3. Données spatiales
4. Infrastructures de données spatiales

Science ouverte

“ Diffusion sans entrave des résultats, des méthodes et des produits de la recherche scientifique. Elle s’appuie sur l’opportunité que représente la mutation numérique pour développer l’accès ouvert aux publications et – autant que possible – aux données, aux codes sources et aux méthodes de la recherche. ”

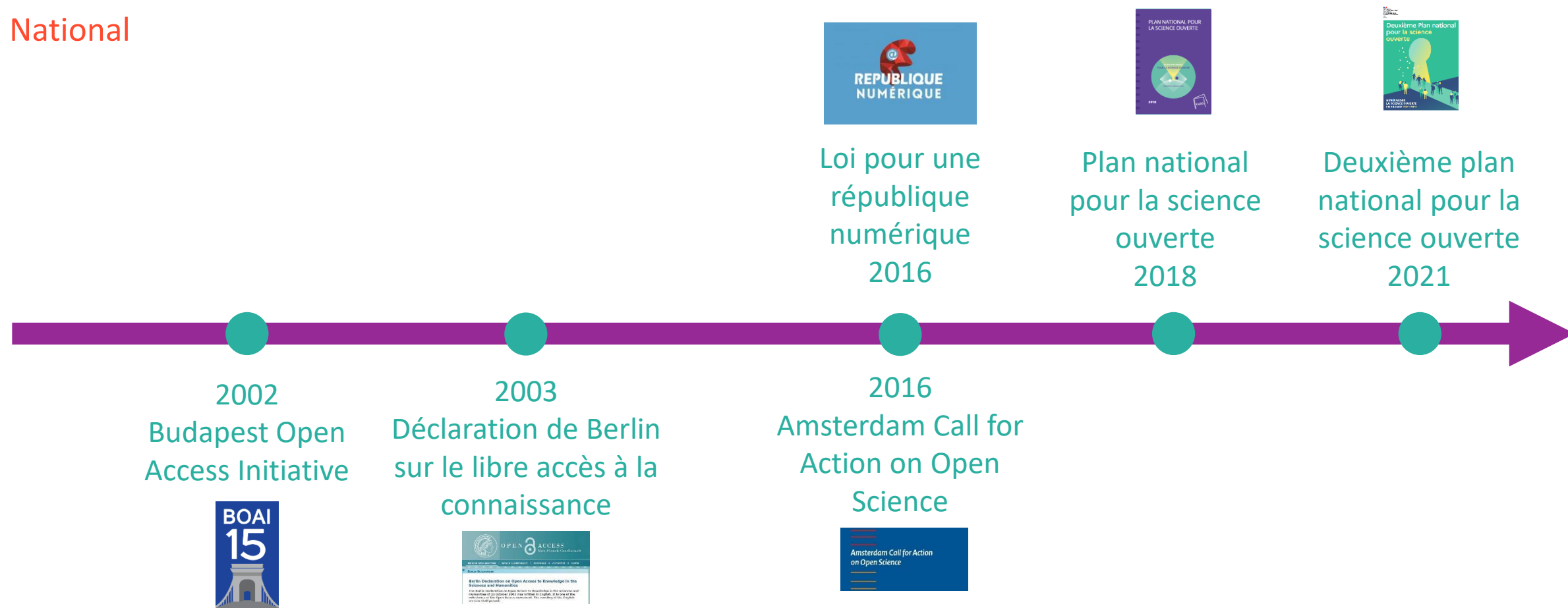
Deuxième plan national pour la science ouverte, 2021

Objectifs

- Augmentation de l’efficacité et de la qualité de la recherche
- Démocratisation des savoirs
- Construction d’un écosystème de transparence
- Favorisation de la confiance des citoyens dans la science

Science ouverte : quelques dates

National



International

Données de la recherche

“enregistrements factuels (chiffres, textes, images, son, vidéo...) utilisés comme sources primaires pour la recherche et qui sont habituellement acceptés par la communauté scientifique comme étant nécessaires pour valider les résultats de la recherche ”

OCDE, 2007

Quelques catégories de données :

- données d'observation
- données expérimentales
- Données de simulation numérique

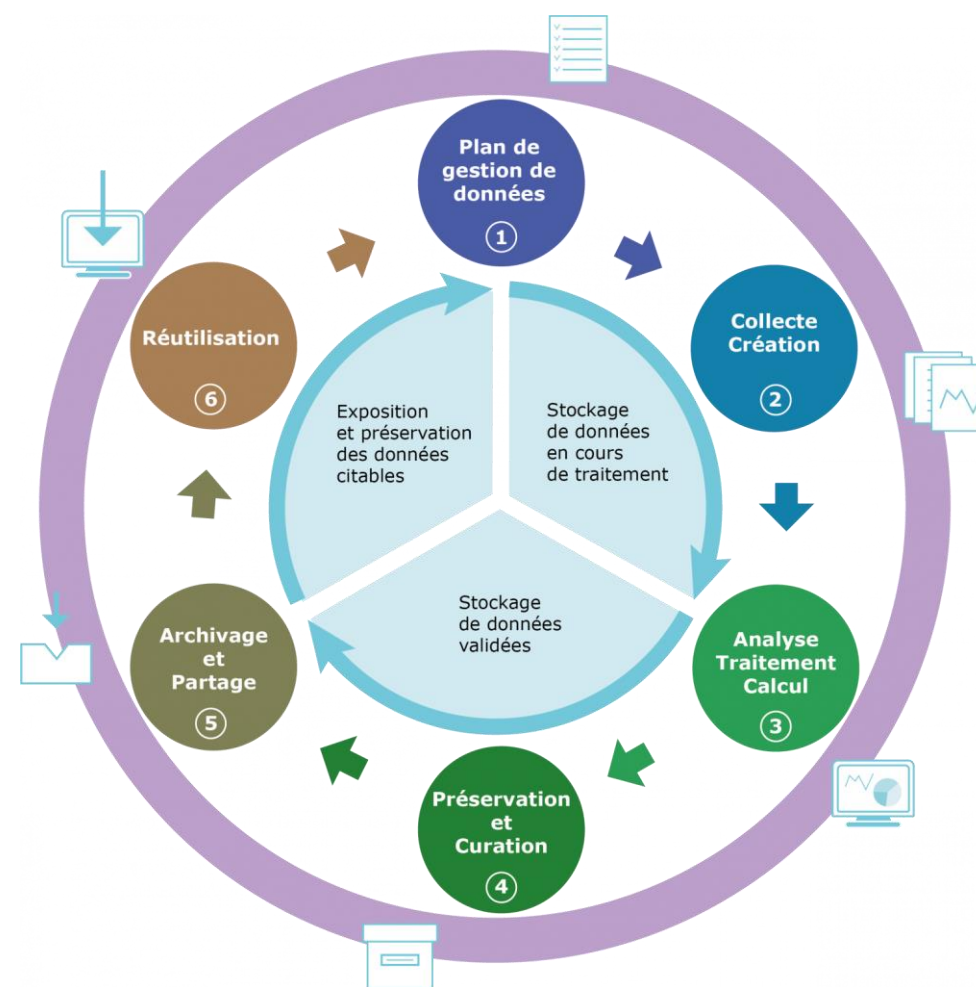
Gestion des données de la recherche

Faciles à trouver (métadonnées, DOI, ...)

Accessibles (protocoles d'accès, ...)

Interopérables (vocabulaires contrôlés, ...)

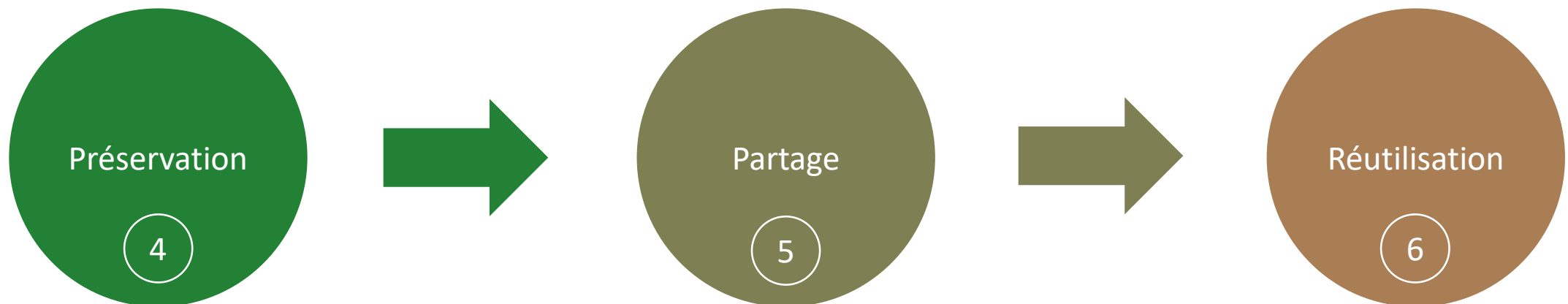
Réutilisables (licences, standards, ...)



Entrepôts de données de la recherche

“base de données destinée à accueillir, conserver, rendre **visibles et accessibles** des données de recherche. Son rôle est de permettre le dépôt ou la **collecte** de données, leur **description**, leur **accès**, et leur **partage** en vue de leur **réutilisation**. Chaque entrepôt dispose généralement d’une politique de dépôt, de description et de diffusion des données. ”

CoopIST, Cirad



Quelques exemples d'entrepôts de données

- Pluridisciplinaires

- Zenodo, Dryad, Nakala



- Disciplinaires

- Genbank (génétique), Néotoma (Paléo-écologie), Telma (manuscrits anciens)



- Institutionnels

- Cirad Dataverse, Datasuds (IRD), Datapartage (INRAE)



- Liés à des revues

- GigaDB (Gigascience Press), Dataverse Ubiquity Press



Annuaire d'entrepôts

Il existe des milliers d'entrepôts de données pour la recherche. Des annuaires permettent de s'y retrouver.

• Re3data

- Créé par le consortium Datacite en 2012
- recense plus de 3000 entrepôts

• Cat OPIDoR

- wiki (hébergé par le CNRS) des services français dédiés aux données de la recherche
- identifie 168 entrepôts.

The screenshot shows the re3data.org search interface. On the left is a 'Filter' sidebar with various categories like Subjects, Content Types, Countries, etc. The main area displays search results for 'Unidata's RAMADDA' and 'ZACAT'. Each result includes the repository name, a brief description, and a 'Country' field. The 'Unidata's RAMADDA' entry is for the United States, and the 'ZACAT' entry is for Germany.

The screenshot shows the Cat OPIDoR website. At the top, there is a search bar and a 'Rechercher sur Cat OPIDoR' button. Below the search bar, the page title is 'Entrepôt de données'. There are links for 'Page', 'Discussion', 'Lire', 'Voir la liste source', and 'Voir l'historique'. A list of data services is displayed in a table format.

Services	Collection/ Sous-ensemble de :	Structure	Domaine scient
2P2IDB	-	CRDM	Ve & Santé
AB0db	-	LMGM	Ve & Santé
AGRHYS	-	OSUR SAS eLTER-France OZCAR Geosciences Revis	Sciences & Tech
AIEA Bases de données	-	AIEA	Sciences & Tech
AMMA-CATCH-DB	-	OSUG	Sciences & Tech

Données spatiales

“ensemble reliant une information relative à un objet ou à un phénomène (attribut) et la localisation de cette information sur la terre.”

Georis-Creuseveau, 2014

On parle aussi de donnée géolocalisée, d'information géographique, ...

La localisation s'exprime sous forme de coordonnées, dans 2 ou 3 dimensions, dans un système de projection

Données spatiales

Les différents types d'objets géographiques

Vecteur

Points



Exemple : adresses

Polylignes



Exemple : réseau routier

Polygones



Exemple : parcelles

Rasters



Exemple : Orthophoto

Diffusion des données spatiales : quelques dates

National



Transposition
d'INSPIRE dans le
droit français
2010



2016



2018



2021

1994

2002

2003

2007

2016

Fondation de
l'Open Geospatial
Consortium



Directive INSPIRE



International

Directive européenne INSPIRE

“INSPIRE est une directive européenne établissant une infrastructure européenne d'échange de données géographiques dans l'union européenne. Cette directive, transposée en droit français par une ordonnance du 21 octobre 2010, favorise la diffusion des données géographiques et de leurs métadonnées par la mise à disposition de services de recherche, de diffusion et de téléchargement.”

GéoBretagne

Obligations

- la fourniture des données selon des règles de mise en œuvre communes ;
- la constitution de catalogues de données (métadonnées) ;
- l'application de règles d'interopérabilité ;
- l'accès gratuit aux métadonnées ;
- l'accès aux données pour les acteurs réalisant une mission entrant dans le cadre d'INSPIRE ;
- les services pour permettre ces accès ;
- l'existence d'une organisation adaptée pour s'assurer de la bonne mise en œuvre de la directive.

Directive INSPIRE : Standards et services

Standards

- ISO 19115 : métadonnées de données géographiques
- ISO 19119 : Métadonnées de services
- ISO 19139 : Transcription en XML des ISO 19115 et 19119

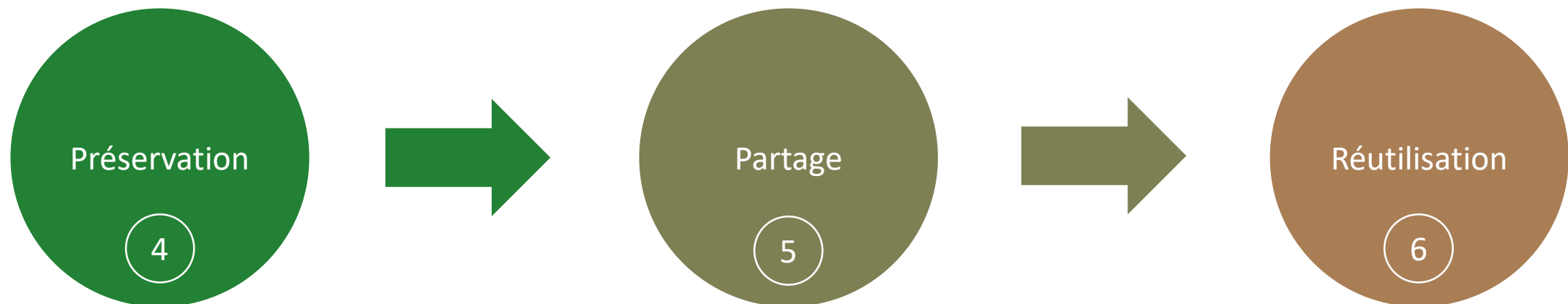
Services

- Catalogue Service for the Web (CSW) : catalogage
- Web Map Service (WMS) : consultation
- Web Feature Service (WFS) : téléchargement

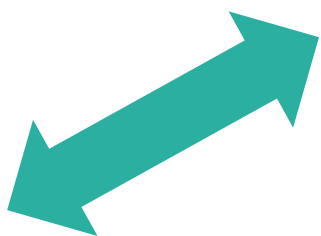
Les infrastructures de données spatiales (IDS)

“ Ensemble de composantes technologiques et non technologiques constituant un réseau intra et inter-organisationnel et ayant pour but de faciliter **l'accès**, **l'échange** et **l'utilisation** de l'information géographique ”

Vandenbroucke et al., 2009



Principaux composants d'une IDS



Visualiseur cartographique

Search filters include: Type de ressource, Type de représentation, Disponible via, et Métaclasse. Metadata cards show details for various datasets like 'ODIT - Mobiliers trouvés sur l'axe dans la forêt de Chailly en 2014' and 'Carte de représentativité des brigades antiterroristes de l'Age du Fer'.

Catalogue de métadonnées



The interface shows a table of data stores with columns for 'Type de données', 'Espace de travail', and 'Nom de l'entrepôt'. It also includes a sidebar with navigation options like 'Serveur', 'Données', and 'Services'.

Type de données	Espace de travail	Nom de l'entrepôt
<input type="checkbox"/>	geosync-ouvert	a1_georef
<input type="checkbox"/>	geosync-ouvert	a4_georef
<input type="checkbox"/>	geor	automated_tests
<input type="checkbox"/>	geosync-ouvert	besancon_a1
<input type="checkbox"/>	geosync-ouvert	besancon_a1_s
<input type="checkbox"/>	geosync-ouvert	geosync_pg_open
<input type="checkbox"/>	geosync-restrict	geosync_pg_rstct
<input type="checkbox"/>	geosync-ouvert	geosync_shp_sper
<input type="checkbox"/>	geosync-restrict	geosync_shp_rstct
<input type="checkbox"/>	geosync-ouvert	hillshade_lidar_be

Serveur de données géographiques

Quelques exemples d'IDS

- Nationales
 - Géoportail
- Régionales
 - IdéoBFC, GéoGrandEst, GéoBretagne
- Institutionnelles
 - Data.shom.fr, InfoTerre,
- De recherche
 - GéoSHERPA, indigeo,
- Annuaire d'IDS
 - Afigéo

The logo for Géoportail, featuring the word "géo" in green and "portail" in blue.The logo for GéoBretagne, featuring the word "Géo" in blue and "Bretagne" in black with a registered trademark symbol.

Conclusion : IDS VS entrepôts de données

Plusieurs démarches en faveur de l'ouverture des données :

- Multiplication des jeux de données accessibles

Mais construites en parallèle, ce qui induit des difficultés :

- Des standards différents utilisés
- Une faible interopérabilité entre ces systèmes

IDS

- + Disposent d'outils spécifiques aux données spatiales
- Ne permettent pas de publier d'autres types de données

Entrepôts de données

- + Permettent de publier tous types de données
- Pas d'outils dédiés aux données spatiales



MERCI
pour votre écoute !



Toutes les **INFORMATIONS** sur la Semaine de la Science ouverte en Bourgogne Franche-Comté :

<https://data.ubfc.fr/2023/10/18/semaine-de-la-science-ouverte-en-bourgogne-franche-comte/>



PROGRAMME COMPLET de la Semaine de la Science ouverte en BFC :

<https://data.ubfc.fr/wp-content/uploads/sites/44/2023/10/PROGRAMME-SEMAINE-SCIENCE-OUVERTE-WEB.pdf>



PROGRAMME des ATELIERS PRATIQUES :

Mardi 14 novembre :

- **14h00 - 16h30**
Dijon > BU Le Cortex

Jeudi 16 novembre :

- **09h30 - 12h00**
Dijon > MSH Dijon, salle des conseils
- **14h00 - 16h30**
Belfort > BU Belfort
- **14h00 - 16h30**
Besançon > OSU THETA



CONTACT :

dataubfc-doc@ubfc.fr



SOURCE des PICTOGRAMMES utilisés :

<https://iconmonstr.com/>