



**Métadonnées de Collection**  
**et Jeux de Données**

**V0.3**

*Document de travail*

*Le 18 novembre 2019*

epointal	Métadonnées Documents de travail	
----------	-------------------------------------	---

## **Objet**

Ce document précise les métadonnées et paramètres retenues par ForM@Ter pour la création de son MétaCatalogue.

V0.1	Premières spécifications	epointal	Septembre 2019
V0.2	Précisions et premiers code listes	epointal	Novembre 2019
V0.3	Détails des champs complexes	epointal	18 Novembre 2019

## Introduction

Pour la création du catalogue de métadonnées, nous avons les objectifs suivants:

- pouvoir exporter dans différents formats les métadonnées.

- ISO19139 valide INSPIRE<sup>i</sup>
- DataCite<sup>ii</sup>
- Format Pivot DataTerra
- UMM NASA<sup>iii</sup>
- compatible avec OGC:O&M

Pour répondre à ces impératifs, nous utilisons l'implémentation ISO19139 à laquelle nous ajoutons 3 thésaurus (variables, plateformes et instruments).

## Définitions

### Collection (ou series)

Une collection est un regroupement de jeux de données réalisés dans les mêmes conditions:

- même plateforme,
- même variable
- et si possible même instrument

### Jeu de données ou granule (ou dataset)

En simplifiant, il s'agit du fichier de données.

**Un jeu de données ne peut appartenir qu'à une unique collection.**

## Les champs essentiels

Titre	Type	Card	Description
hierarchyLevel	series ou dataset	1	Type de fiche
dateStamp	Date iso8601	1	Date de maj des métadonnées (automatique)
metadataContact	<a href="#">#Contact</a>	1..n	Contact pour les métadonnées
<b>dataIdentification</b>			
identifier	URI, URN, DOI, ...	1..n	Identifiant des données
title	Texte	1	Titre du jeux de données ou de la collection
abstract	Texte ou html	1	Description du jeu de données ou de la collection
graphicOverview	url + titre + description	1..n	capture d'écran, image d'illustration
dataContact	<a href="#">#Contact</a>	1..n	Contact pour les données (organisation + email )
date	Date iso8601	1..3	Date de création ou publication ou de maj des données
Variable/productType	<a href="#">#descriptiveKeyword</a>	1..n	Mot clés d'un thésaurus à réaliser
platform	<a href="#">#descriptiveKeyword</a>	1	Thésaurus à réaliser
spatialExtent	Bounding Box (WGS84)	1..n	Etendue(s) spatiale(s).
temporalExtent		1..n	Etendue(s) temporelle(s)
resourceConstraints	<a href="#">#resourceConstraints</a>	1..n	Contraintes d'accès et d'utilisation, license...
onlineResource	<a href="#">#onLineResource</a>	0..n	lien vers services divers (téléchargement, visualisation, répertoire ftp, page d'information, service WMS, service WFS)
<b>dataQuality</b>			
lineage	texte	0..1	description de la généalogie des données

## Liste détaillée des champs

M pour Mandatory  
 C pour Conditional  
 R pour Recommended  
 O pour Optional  
 N/A pour Not applicable

La liste suivante récapitule les champs d'une fiche de métadonnées, ainsi que son exigibilité pour ForM@Ter, la norme ISO19139+INSPIRE et DataCite.

Les dénominations sont le plus possible reprise des dénominations ISO19139.

Les champs plus complexes sont précisés ensuite.

Dénomination	Type	Défaut	ForM@Ter	Card	ISO19139+INSPIRE	DATA CITE	Description
<b>Metadata</b>							
fileIdentifier	URI, URN ...		<b>M</b>	1	M		Identifiant <b>unique et permanent</b> de la fiche de métadonnée
parentIdentifier	identifiant de fiche de métadonnée parent		<b>C</b>	0..1	C	N/A	l'identifiant de la fiche parent si elle existe. Un seul parent possible.
metadataLanguage Code	fre et/ou eng	eng	<b>M</b>	1..2	M		Langue de la fiche de métadonnées (peut être bilingue fre et eng)
dateStamp	Date ou DateTime ISO8601		<b>M</b>	1	M		Date de mise à jour de la fiche de métadonnées
pointOfContact	<a href="#">#Contact</a>		<b>M</b>	1..n	M		Contact avec au moins un mail pour les métadonnées
metadataCharacterSet	<a href="#">MD_CharacterSet Code</a>	utf-8	<b>M</b>	1	M		Encodage de la fiche de métadonnées
hierarchyLevel	series   dataset	dataset	<b>M</b>	1	M	M	series, dataset, service...
referenceSystemInfo	Liste à établir <a href="#">Specification INSPIRE</a>	ESPG:4326	<b>M</b>	1..n	M		Le(s) système(s) de coordonnées de référence utilisé(s) pour les données
temporalCRS	Liste à établir		<b>C</b>	0..n	C	N/A	Référentiel temporel des données si on utilise un référentiel particulier autre que UTC (datetime) ou calendrier grégorien (date)
verticalCRS	Liste à établir		<b>C</b>	0..n	C	N/A	Si a du sens pour les données
<b>identificationInfo - Identification de la collection ou du jeu de données</b>							
identifiant	URI, URN,		<b>M</b>	1..n	M	M	Identifiant unique pour la

		<b>DOI, ...</b>					collection de jeux de données. Peut avoir plusieurs identifiants La réf dans la DB + un DOI ....
status		<a href="#">MD_ProgressCode</a>	completed	<b>O</b>	0..1		
title		Texte		<b>M</b>	1	M	M
abstract		Texte		<b>M</b>	1	M	M
date	maj	Date ISO8601		<b>M</b>	0..1	M au moins 1 des 3	Date de maj du jeu de données ou de création ou de la collection
	création				0..1		
	publication				0..1		
pointOfContact		<a href="#">#Contact</a>		<b>M</b>	1..n	M	M
descriptiveKeyword		<a href="#">#Keyword</a>		<b>R</b>	0..n	R	R
GEMET Inspire Keyword		<a href="#">GEMET INSPIRE</a>		<b>M</b>	1	M	O
topicCategory		<a href="#">TopicCategories</a>	geoscientific information	<b>M</b>	1..n	M	O
spatialRepresentationInfo		<a href="#">MD_SpatialRepresentationTypeCode</a>	raster	<b>M</b>	1..n		
variables productType		Liste à établir Thésaurus ForM@Ter		<b>M</b>	1..n	O	
platform		Liste à établir Thésaurus ForM@Ter		<b>R</b>	1..n	O	N/A
instrument		Liste à établir Thésaurus ForM@Ter		<b>O</b>	0..n	O	N/A
dataCenter		Liste à établir Thésaurus ForM@Ter		<b>M</b>	1	O	N/A
spatialResolution		Nombre (voir <a href="#">#spatialResolution</a> )		<b>O</b>	0..n		N/A
timeResolution		liste à établir <a href="#">Thésaurus GCMD</a>		<b>C</b>	0..n		N/A
dataLanguageCode		eng ou fre	eng	<b>O</b>		M	O
dataCharacterSet		<a href="#">MD_CharacterSetCode</a>	utf-8	<b>C</b>	0..1	C	

epointal	Métadonnées Documents de travail	
----------	-------------------------------------	---

spatialExtent		bbox (WGS84)		<b>M</b>	1..n	M	M une seule	
temporal Extent	begin	Date ISO8601		<b>M</b>			N/A	
	end			<b>O (rien ou now si on going)</b>				
dataMaintenance Frequency		<a href="#">MD Maintenance FrequencyCode</a>	asNeeded	<b>O</b>	0..1	O		Fréquence de la maintenance des données
<b>resourceConstraints - legalConstraints</b>								
useLimitation			No conditons apply	<b>R</b>	1..n			Texte décrivant comment citer la resource par exemple nom de la license, url de la license
useConstraints		<a href="#">MD RestrictionC ode</a>		<b>C</b>	0..n			Mot clés: License copyright ...
accessConstraints		Liste INSPIRE	No limitations on public access	<b>M</b>	1..n			restriction d'accès suivant la liste INSPIRE
<b>distributionInfo</b>								
format				<b>M</b>	1..n	M		Format des données
onlineResource		voir onlineResource		<b>R</b>	0..n			Lien de téléchargement ou d'accès (surtout pour les dataset!)
distributor		Voir contact		<b>R</b>	1	O	M	responsable de la diffusion des données (publisher pour datacite dans geonetwork)
<b>qualityInfo</b>								
conformity				<b>M</b>	0..1		N/A	Conformité des données aux normes ?
lineage	source				0..n			Description de la généalogie des données
	statement	texte		<b>R</b>				

## Détails des champs complexes

### Les identifiants

### La résolution spatiale

2 possibilités pour la résolution spatiale: une distance en mètre ou un dénominateur.

## Exemples ISO19139

/MD\_Medadata/

## Les mots clés

Les mots clés sont triés par [MD\\_KeywordTypeCode](#).

MD_KeywordTypeCode
discipline
place
stratum
temporal
theme

## Mot clé libre

Un mot clé libre est défini par le mot clé + le type de mot clé ([MD\\_KeywordTypeCode](#))

## Mot clé de thésaurus

Les thésaurus sont des vocabulaires contrôlés, c'est à dire une liste de termes si possible avec leur définition. Ils peuvent être hiérarchiques ou non.

Ce vocabulaire sera disponible sous forme de fichier au format [skos-rdf](#).

Les thésaurus ForM@Ter, en cours d'élaboration, sont:

- dataCenter
- variable
- plateforme
- instrument

ainsi que:

- timeResolution

De nombreux autres thésaurus peuvent être utilisés:

- Geonetwork par exemple, utilise le thésaurus [naturalearth-and-seavox](#) comme thésaurus de localisation.
- La NASA a créé plusieurs thésaurus: [GCMD keywords](#), aussi bien de localisation, que de mots scientifiques ou d'instruments



- Nous pouvons aussi citer les «Climat and Forecast Standard Names » dont les différentes versions se trouvent sur la page <http://cfconventions.org/standard-names.html> ainsi qu'un autre thésaurus de localisation

		description	exemple
keyword	keyword	dénomination du concept	age_of_sea_ice
	keywordURI	uri du concept	<a href="http://mmisw.org/ont/cf/parameter/age_of_sea_ice">http://mmisw.org/ont/cf/parameter/age_of_sea_ice</a>
thesaurus	thesaurusName	dénomination du thésaurus (y compris version)	Climate and Forecast (CF) Standard Names (v.69)
	thesaurusURI	uri du thésaurus	<a href="http://mmisw.org/ont/cf/parameter">http://mmisw.org/ont/cf/parameter</a>
	thesaurusDateRevision	date de mise à jour du thésaurus	(si existe: généralement change de version à chaque révision)
	thesaurusDatePublication	date de publication du thésaurus	2019-10-17

## Cas particulier du GEMET1.0

La directive INSPIRE demande d'ajouter un mot clé [GEMET1.0](#)

## Exemples ISO19139

### Mots clés libres

Ici, ils sont du même type, donc ils sont dans la même balise gmd:descriptiveKeyword

```
<gmd:descriptiveKeywords>
  <gmd:MD_Keywords>
    <gmd:keyword>
      <gco:CharacterString>Spain</gco:CharacterString>
    </gmd:keyword>
    <gmd:keyword>
      <gco:CharacterString>Andalusia</gco:CharacterString>
    </gmd:keyword>
    <gmd:type>
      <gmd:MD_KeywordTypeCode codeListValue="place"
codeList="http://standards.iso.org/iso/19139/resources/gmxCodeLists.xml#MD_KeywordTypeCode">Place</
gmd:MD_KeywordTypeCode>
    </gmd:type>
  </gmd:MD_Keywords>
</gmd:descriptiveKeywords>
```

## Mot clé d'un thésaurus

```
<gmd:descriptiveKeywords>
  <gmd:MD_Keywords>
    <gmd:keyword>
      <gmx:Anchor xlink:href="http://inspire.ec.europa.eu/theme/lu">Land use</gmx:Anchor>
    </gmd:keyword>
    <gmd:thesaurusName>
      <gmd:CI_Citation>
        <gmd:title>
          <gmx:Anchor xlink:href="http://www.eionet.europa.eu/gemet/inspire_themes">GEMET -
INSPIRE themes, version 1.0</gmx:Anchor>
        </gmd:title>
        <gmd:date>
          <gmd:CI_Date>
            <gmd:date>
              <gco>Date>2008-06-01</gco>Date>
            </gmd:date>
            <gmd:dateType>
              <gmd:CI_DateTypeCode
codeList="http://standards.iso.org/iso/19139/resources/gmxCodeLists.xml#CI_DateTypeCode"
codeListValue="publication">Publication</gmd:CI_DateTypeCode>
            </gmd:dateType>
          </gmd:CI_Date>
        </gmd:date>
      </gmd:CI_Citation>
    </gmd:thesaurusName>
  </gmd:MD_Keywords>
</gmd:descriptiveKeywords>
```

## Les contacts

## Les ressources en ligne

## Les contraintes d'accès et d'utilisation

La directive INSPIRE est loin d'être claire concernant les contraintes d'accès et d'utilisation. On trouve sur le site wiki de georezo.net, la description des différents cas possibles:

[https://georezo.net/wiki/main/donnees/inspire/aide\\_a\\_la\\_saisie\\_des\\_metadonnees\\_inspire](https://georezo.net/wiki/main/donnees/inspire/aide_a_la_saisie_des_metadonnees_inspire)

## Code listes

Status

Rôles

CRS

topicCategory

## Exemples de fiches

### GeoJSON

### ISO19139

### DataCite

Documentations:

Exemple d'implémentation INSPIRE pour les métadonnées d'un jeu de données

[https://inspire.ec.europa.eu/sites/default/files/md\\_2.0\\_datasets\\_example.xml](https://inspire.ec.europa.eu/sites/default/files/md_2.0_datasets_example.xml)

Guide d'implémentation ISO 19139 INSPIRE

<https://inspire.ec.europa.eu/sites/default/files/documents/metadatas/inspire-tg-metadatas-iso19139-2.0.1.pdf>

DataCite métadonnées:

[https://schema.datacite.org/meta/kernel-4.2/doc/DataCite-MetadataKernel\\_v4.2.pdf](https://schema.datacite.org/meta/kernel-4.2/doc/DataCite-MetadataKernel_v4.2.pdf)

Code listes et thésaurus GCMD - NASA

<https://gcmdservices.gsfc.nasa.gov/static/kms/>

Pour les codes listes ISO19115

[https://geo-ide.noaa.gov/wiki/index.php?title=ISO\\_19115\\_and\\_19115-2\\_CodeList\\_Dictionaries](https://geo-ide.noaa.gov/wiki/index.php?title=ISO_19115_and_19115-2_CodeList_Dictionaries)

A propos de sos: <https://github.com/52North/SOS>

Spécification CRS

[INSPIRE\\_specification\\_CRS\\_v3.0.pdf](#)

- i La directive INSPIRE est une directive européenne sur le partage des données géographiques et comprend un ensemble de recommandations pour l'implémentation des métadonnées.
- ii DataCite est un organisme international de bibliothèques et autres services des sciences de l'information dont l'objectif est de faciliter l'accès aux données numériques. DataCite possède son propre [Metadata Schema](#)
- iii Le [modèle de métadonnées de la NASA](#)