

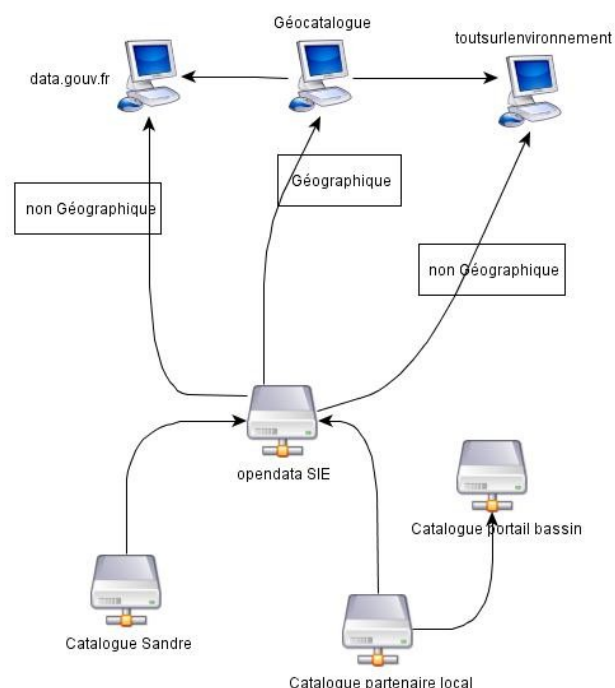
Titre	Règles de saisies des métadonnées de l'opendata du SIE
Créateur	Onema
Éditeur	Office national de l'eau et des milieux aquatiques
Version	2013-1
Diffusion	
Date	Octobre 2013

Le site data.eaufrance.fr a pour vocation à référencer les données publiques sur l'eau, réutilisables gratuitement, des acteurs du SIE : services de l'Etat, établissements publics, collectivités locales ...

Ce site répond aux obligations de la directive sur la réutilisation des données publiques ; le principe de sa mise en œuvre a été présenté en GCiB, puis au comité permanent des usagers du SIE (CPU SIE) du CNE, qui en ont approuvé le principe et a été inscrit dans le schéma web de diffusion des données.

Pour améliorer la saisie des métadonnées, et leur cohérence et bonne lisibilité, des règles de bonne pratique doivent être adoptées par les acteurs de l'eau. Elles sont largement inspirées du document du « Guide de saisie des éléments de métadonnées INSPIRE » publié par le CNIG en décembre 2011.

Ces métadonnées seront visibles sur le site data.eaufrance.fr, mais aussi par moissonnage sur les sites data.gouv.fr, et tousurlenvironnement.fr, et pour les données géographiques, sur le site geocatalogue.gouv.fr, selon le schéma de moissonnage ci contre.



Règles proposées

Règle générale de regroupement

Les données proposées en téléchargement sont souvent proposées sur Internet en plusieurs fichiers, correspondant par exemple à des territoires différents (régions, bassins, ...), des périodes temporelles différentes (un par année, ...) et des formats différents, ceci pour faciliter le chargement des données (limitation du volume des fichiers) et pour répondre à des besoins spécifiques.

Ces fichiers seront regroupés dans une seule fiche de métadonnée, donc avec un titre et une description unique, chaque fichier étant indexé dans cette fiche comme un lien de téléchargement particulier.

Identification des données

Titre

Donnée obligatoire, non répétable.

Il est recommandé de ne pas faire figurer les parties responsables c'est-à-dire les organisations exerçant une quelconque responsabilité (producteur, ...).

Il est recommandé de ne pas indiquer uniquement l'acronyme de la ressource (à moins de l'expliquer dans le résumé).

Les règles de regroupement proposées par data.gouv.fr font qu'un ensemble de ressource regroupe les fichiers correspondant à la même donnée diffusée selon des découpages territoriaux, temporels ou des formats différents. Les informations sur le territoire couvert sont à indiquer dans le titre, pour indiquer la somme des territoires couverts ; de même une notion de date ne doit être présente dans le titre que pour indiquer la période complète proposée, ou le début de la période proposée, si elle est mise à jour régulièrement.

Description

Donnée obligatoire, non répétable.

Il est attendu un texte significatif décrivant la ressource. Tout texte vide (ensemble de caractères d'espacement) ou de type « Non renseigné », ... ne convient pas.

Les premières lignes du texte sont celles qui seront vues dès la sélection des ressources : il faut donc qu'elles soient significatives, et ne pas être une simple répétition du titre.

Le résumé doit décrire la ressource de façon compréhensible par l'utilisateur. Pour un producteur, il s'agit en particulier de définir au mieux l'information ou le phénomène représenté dans la donnée. On va donc y trouver des éléments de définition, mais aussi éventuellement une indication sommaire de la zone couverte ou le cas échéant, des

informations sur les particularités de la version de la ressource. Il devrait contenir une définition officielle quand elle existe, ou une définition commune, La simple référence à un texte législatif ou réglementaire n'est pas suffisante.

Il faut indiquer dans la mesure du possible dans la description le nom de la banque (Ades, Osur, ...).

Il faut indiquer autant que faire se peut les principales informations ou attributs présents dans la données (par exemple : prix de l'eau, nitrates, pesticides, ...), de manière à faciliter la recherche de ce type d'information.

Type de ressources

Donnée obligatoire, non répétable.

Une seule valeur possible parmi les suivantes :

- o collection de données (géographiques),
- o jeu de données (géographiques),
- o jeu de données non géographiques
- o service

Par jeu de données il faut entendre un lot, une couche ou une série de donnée géographique (au sens inspire).

Par collection de données il faut entendre un ensemble de jeu de données géographiques (par exemple les différentes versions d'une même donnée)

Par jeu de données non géographiques il faut entendre un lot, un fichier ou un ensemble de tables de données ne contenant pas de coordonnées.

Par service il faut entendre un accès à une donnée.

Langue

Donnée obligatoire, répétable.

La valeur de cet élément est en général français (codé fre).

Lorsque la ressource n'inclut pas d'information textuelle, il est recommandé de mettre par défaut la langue des métadonnées soit le français (fre).

Accès à la donnée

Donnée obligatoire, répétable.

On indique autant de modalité d'accès qu'il y a de découpage en fichier proposé pour une même donnée.

Type d'accès

Une seule valeur possible parmi les suivantes :

- o Accès en téléchargement depuis un site distant
- o Téléchargement depuis data.eaufrance.fr

Le cas d'accès le plus courant est « accès en téléchargement depuis un site distant » : il s'agit du cas où la donnée est déjà en ligne sur un site de la toile eaufrance, ou du cas plus rare de banques permettant le téléchargement de données sélectionnées selon un outil de requête (ADES.eaufrance.fr par exemple).

Le « téléchargement depuis le site data.eaufrance.fr » correspond au cas où la donnée n'est pas en ligne, et où le téléchargement sera assuré via un chargement de la donnée dans data.eaufrance.fr.

Encodage

A sélectionner dans la liste proposée.
Si le fichier distant est une archive (ZIP ou autre), c'est le format de la donnée dans l'archive qui doit être indiqué.
Les flux web sont considérés ici comme des formats particuliers de livraison de la donnée.

Titre

Le titre du fichier ne doit pas reprendre le titre de la donnée, mais préciser le territoire (si chaque fichier correspond à un territoire particulier) et la période (si chaque fichier correspond à une période) ou tout autre élément de découpage en fichier.

URL(s)

Ce champ de saisie correspond au choix de l'option « Accès en téléchargement depuis un site distant »

Indiquer ici l'adresse complète qui permet d'accéder au fichier de donnée ou au service web.

Sauf dans le cas particulier de banques permettant le téléchargement de données sélectionnées selon un outil de requête (ADES.eaufrance.fr par exemple), on ne devrait pas avoir d'adresse aboutissant sur une page web classique, mais bien des adresses pointant sur des fichiers de données.

Fichier

Ce champ de saisie correspond au choix de l'option « Téléchargement depuis data.eaufrance.fr »

Rechercher le fichier à télécharger dans data.eaufrance.fr ; une fois celui-ci téléchargé, l'adresse d'accès sera automatiquement renseigné.

Producteur

Donnée obligatoire, répétable

Producteur

Le producteur (ou organisme responsable) est celui qui assume le référencement dans data.eaufrance.fr, comme indiqué dans la note « Règles de gouvernance et bonnes pratiques pour l'alimentation de l'opendata du SIE »

Le nom à indiquer est le nom de l'organisme, voire du service au sein de l'organisme.

Il ne faut pas indiquer un nom de personne.

Si la donnée est produite par un nombre important d'acteur (données dans Ades par exemple) on donne comme nom de l'organisme un nom générique, correspondant au mieux à l'ensemble des organismes concernés (Partenaires du SIE, services de bassin, ...)

Il est déconseillé de fournir une adresse de messagerie électronique nominative, ceci afin de pouvoir plus facilement gérer les changements de fonctions des personnes impliquées.

Les adresses trop génériques de type « accueil@macomcom.fr » ou « contact@macomcom.fr » sont également à éviter

Mots clé

Catégorie ISO

Donnée obligatoire

Liste fermée propre à ISO 19115
En général on indique « environnement » ou « eaux intérieur, hydrographie », ou « biologie, faune et flore ».

Thème INSPIRE

Donnée répétable

Liste fermée propre à la directive Inspire.
A ne remplir que si la donnée est concernée par la Directive Inspire (donc principalement des données géographiques). Il faut alors indiquer le thème Inspire principal correspondant à la donnée concernée.

Domaine Etalab

Donnée répétable

Liste propre à Etalab, basée sur le thésaurus Eurovoc (voir annexe 1).
Ne sont proposés que les éléments d'Eurovoc ayant un intérêt dans le domaine de l'eau.

Thème eau

Donnée répétable

Liste fermée propre à data.eaufrance.fr (voir annexe 2).
Cette liste facilite l'accès aux données dans data.eaufrance.fr. Elle permet de caractériser rapidement une donnée, par rapport au type de donnée, au milieu, ou à l'approche état-pression-réponse.

Situation géographique

Emprise

Donnée propre aux informations géographiques

L'emprise est donnée en longitudes est et ouest et latitudes sud et nord en degrés décimaux, avec une précision d'au moins deux chiffres après la virgule.
L'emprise doit englober l'étendue du territoire pour lequel le producteur garantit la saisie de l'information fournie, une partie du territoire pouvant être vide du type d'objet saisi (par exemple, dans le cas de la localisation des sites Ramsar en Loire-Bretagne, l'emprise sera le rectangle englobant le bassin Loire-Bretagne, même si le rectangle englobant les sites Ramsar est plus petit ; en effet, l'absence de site Ramsar est une information en soi sur le reste du territoire du bassin Loire Bretagne).

Le rectangle de délimitation doit être le plus ajusté possible, afin de délimiter le plus fidèlement possible la ressource décrite (ne pas donner un rectangle couvrant la France entière pour des données limitées à une commune).

Si la ressource couvre la France métropolitaine et/ou un territoire d'outre-mer, elle contiendra autant d'emprises géographiques que de territoires couverts (pas d'emprise « mondiale »). Une seule emprise est définie pour le cas d'une ressource couvrant la France métropolitaine.

Territoire / Bassin

Donnée répétable

Liste fermée des bassins DCE plus France Métropolitaine et France entière.
Permet une recherche simple par principaux territoires sur l'eau.

Référence temporelle

Date

Donnée obligatoire

La date peut être celle de création, de révision, ou de publication. Il faut fournir si possible la date la plus récente (en particulier s'il y a révision ou mise à jour de la donnée).

Qualité et Validité

Généalogie

La généalogie de la ressource décrit l'historique d'un jeu de données et, s'il est connu, le cycle de vie de celui-ci, depuis l'acquisition et la saisie de l'information jusqu'à sa compilation avec d'autres jeux et les variantes de sa forme actuelle.

Il s'agit d'apporter une description littérale et concise soit de l'histoire du jeu de données, soit des moyens, procédures ou traitements informatiques mis en oeuvre au moment de l'acquisition du jeu de données. Par exemple, la généalogie peut consigner l'échelle de saisie si cette information est importante pour l'utilisation du jeu de données.

La date ou la version de la donnée source est également un élément utile.

Un lien vers une ressource documentaire externe (ce peut être un PDF) décrivant par exemple un processus de transformation, pourra être inséré.

Résolution spatiale

Donnée propre aux informations géographiques

Est exprimé sous forme d'une valeur entière correspondant au dénominateur de l'échelle.

Point de contact

Nom de l'organisme

Le point de contact est en général celui qui assume le référencement dans data.eaufrance.fr, comme indiqué dans la note « Règles de gouvernance et bonnes pratiques pour l'alimentation de l'opendata du SIE »

Le nom à indiquer est le nom de l'organisme, voire du service au sein de l'organisme.

Il ne faut en général pas indiquer un nom de personne.

Email

Il est déconseillé de fournir une adresse électronique nominative, ceci afin de pouvoir plus facilement gérer les changements de fonctions des personnes impliquées.

Les adresses trop générique (de type « accueil@macomcom.fr » ou « contact@macomcom.fr ») sont également à éviter.

Cette adresse est bien évidemment relevée par une personne responsable.

Fonction

La fonction du point de contact est « point de contact », sauf cas particulier.

Annexe 1 : liste eurovoc propre au domaine eau

Thèmes eurovoc
aménagement hydraulique (6831)
assainissement
consommation alimentaire (2026)
consommation d'eau
cours d'eau (5211)
distribution de l'eau (2846)
eau (5211)
eau de baignade
eau potable (5211)
eau souterraine
eau superficielle
eaux intérieures (1231)
énergie hydraulique (6626)
engrais chimique (5626)
espèce marine
eutrophisation
gestion de l'eau (5206)
hydrogéologie (3606)
hydrologie (3606)
impact sur l'environnement
inondation (5216)
instrument économique pour l'environnement
irrigation (5621)
milieu aquatique
milieu marin
pêche en eau douce (5641)
pesticide (5626)
politique communautaire de l'environnement
polluant de l'eau
pollution de l'eau (5216)
pollution des cours d'eau (5216)
pollution d'origine tellurique (5216)
pollution marine
pollution par l'agriculture
protection de l'environnement
protection des eaux (5206)
qualité de l'environnement
redevance environnementale
ressource en eau
risque naturel (5206)
substance dangereuse
substance toxique
utilisation de l'eau (5206)

Annexe 2 : liste de thèmes propre à data.eaufrance.fr

Thèmes data.eaufrance.fr
Types par milieu
- Eaux intérieures de surface (rivières et lacs)
- Eaux souterraines (nappes d'eau)
- Eaux côtières (eaux maritimes)
- Eaux de transition (estuaires)
- Milieux remarquables - biodiversité
Types par missions
- Gouvernance de l'eau (Politique publique, gestion intégrée)
- Zonages techniques et réglementaires (zones protégées, réglementées)
- Données réglementaires (données issues des déclarations obligatoires, hors zonage)
- Organisation du suivi de l'eau (programmes de surveillance, réseaux de mesure)
- Rapportage (données issues des reportages)
Types par nature
- Etat du milieu (qualité et quantité)
- Usages et pressions (Prélèvements, rejets, activités ayant un impact sur l'eau)
- Ouvrages liés à l'eau (obstacles à l'écoulement, stations dépuración, ...)
- Actions sur l'eau (réponses, zonage réglementaire, redevance, actions volontaires)
- Risques liés à l'eau (naturels, sanitaires et technologiques)
- Données de références (référentiels, vocabulaire contrôlé)