

Titre	Règles de gouvernance et bonnes pratiques pour l'alimentation de l'opendata du SIE
Créateur	Laurent Coudercy
Éditeur	Office national de l'eau et des milieux aquatiques
Version	2013
Diffusion	
Date	décembre 2012

Dans le cadre du Schéma National des Données sur l'Eau, et du schéma web de diffusion des données, il est créé au sein du SIE un répertoire d'accès aux données sur l'eau appelé data.eaufrance.fr.

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 éviter les incohérences et les doublons dans les alimentations sur data.eaufrance.fr, une gouvernance et des règles de bonne pratique doivent être adoptées par les acteurs de l'eau.

Règles proposées

1. Les données environnementales participant au SIE sont libres et gratuites, sous réserve des secrets éventuels limitant leur possibilité de diffusion, conformément au schéma national des données sur l'eau. Sous cette réserve liées aux secrets prévus par les directives européennes sur l'accès et la réutilisation, elles sont donc librement réutilisables, y compris pour un usage commerciale, et ceci gratuitement. Seules les données sur l'eau non concernées par une limitation d'usage ont vocation à être référencées dans data.gouv.fr.
2. Référencement unique dans data.eaufrance.fr pour alimenter data.gouv.fr, et le géoportail : data.gouv.fr et le géoportail seront alimentés directement par moissonnage de data.eaufrance.fr : il faut donc privilégier le référencement dans data.eaufrance.fr, et ne pas saisir les références aux mêmes données dans data.gouv.fr ou le géoportail. Le schéma joint en annexe montre l'organisation des moissonnages envisagés.

3. Référencement des données nationales géré centralement : les données nationales disponibles sur la toile eaufrance seront référencées par l'Onema ou les administrateurs des banques nationales, et pas par chaque producteur. Par exemple les zonages réglementaires sur l'eau sont référencés au niveau national, en mettant le lien vers leur version publiée sur le site du Sandre. De même les données sur les eaux souterraines seront référencées nationalement, en pointant sur Ades.
4. Référencement des données locales par les acteurs concernés : à l'inverse, pour les données disponibles uniquement au niveau local, chaque producteur a la charge d'en assurer le référencement ; les données relevant du domaine de l'eau seront référencées dans data.eaufrance.fr, les autres pouvant être référencées directement dans les autres portails nationaux.
5. Référencement des seules données validées et publiées : les données à référencer dans data.eaufrance.fr doivent être limitées aux seules données validées et publiées sur Internet de manière publique, libre de réutilisation commerciale et gratuite.
6. Référencement sans déplacement des données : les données référencées ne seront pas copiées dans data.eaufrance.fr, mais seront disponibles via des liens profonds pointant vers les sites de la toile eaufrance ou des partenaires.
7. Référencement vers des URL fixes : autant que faire se peut, les données référencées seront accessibles par des URL fixes, qui seront utilisées pour le référencement.
8. Compléter le référencement par un lien vers les métadonnées : si des métadonnées, ou des documents expliquant la donnée existent, un lien sera fait vers ce document dans les métadonnées de data.eaufrance.fr.
9. Utiliser la licence etalab : les données sur l'eau mises à disposition et référencées dans data.eaufrance.fr doivent être gratuites, librement accessibles et les droits cédés doivent s'appuyer sur la licence etalab, conformément à la circulaire du premier ministre de 2011.
10. Les données seront proposées sous des formats facilitant leur réutilisation technique : seront ainsi privilégiés les formats libres (csv, xml, gml, txt, odt ...) ; pourront aussi être utilisés ceux qui sont couramment lus et utilisés par des logiciels libres (shp, MIF-MID, xls, kml, ...) ; des flux (wfs, flux xml, ...) devront être proposés dès que possible. Si le format de données suppose plusieurs fichiers ou si des documents annexes sont joints, les données sont livrées sous forme d'archive (zip, rar, ...).

Annexe

Schéma du chainage entre les différents catalogues

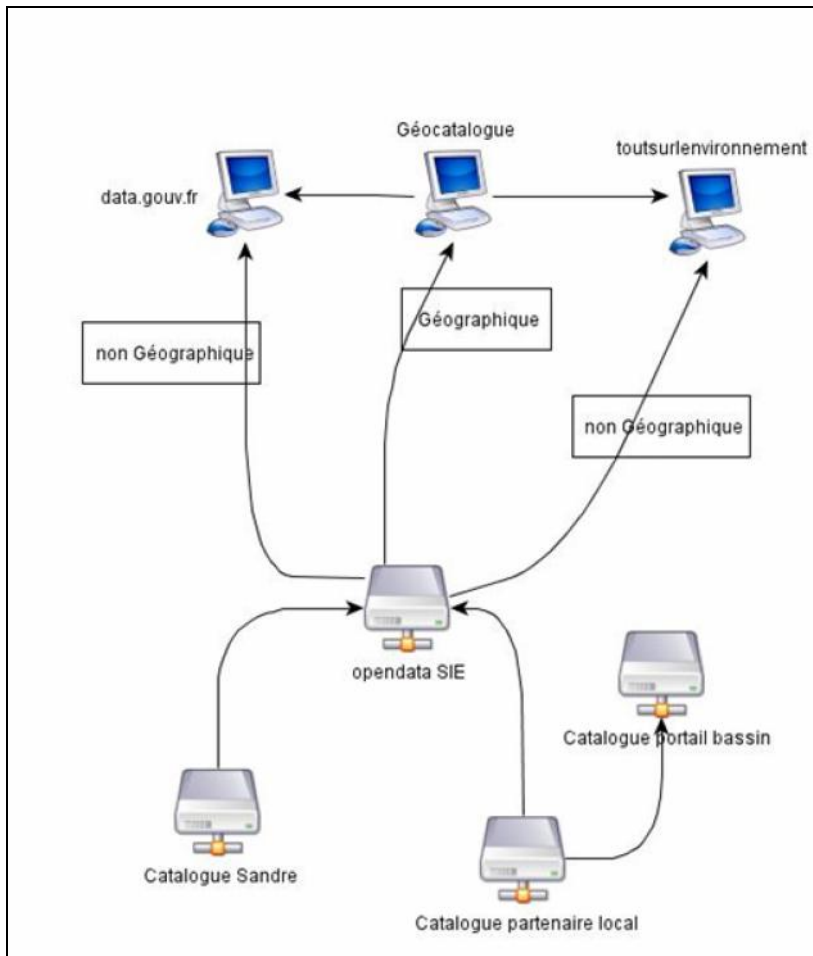


Schéma simplifié de moissonnage entre catalogues.

Ce schéma fait l'hypothèse que data.gouv.fr puisse moissonner et moissonne le géocatalogue. Ce schéma de moissonnage vise à limiter les risques de doublons dans les catalogues nationaux, en particulier en séparant le moissonnage sortant de l'opendata du SIE en deux lots : un géographique, qui est récupéré par le géocatalogue, un non géographique qui alimente data.gouv.fr et tousurlenvironnement.fr, ces deux derniers devant moissonner ou moissonnant déjà le géocatalogue.