

Atelier Geoflow - CR

Animateur : Wilfried Heintz (ZAPYGAR)

Participants : Marie Silvestre (ZAS), David Eschbach (ZAS), Cécile Pignol (ZAA), Eric L'elebrellec (ZAS), Isabelle Charpentier (ZAEU), Virginie Girard (eLTER FR)

Date : session 1 (22/10/20, 9h30-16h), session 2 (13/11/20, 11-14h)

SOMMAIRE

Consignes au préalable de l'atelier	2
1. Données	2
2. Installation et configuration R Studio	2
Présentation de geoflow	3
Pratiques	3
1. Exemple "Forêts anciennes des Pyrénées"	3
2. Applications	4
Rappel : Notions clefs	4
Questions / Tests à faire	5
Perspectives	6

Consignes au préalable de l'atelier

1. Données

Prévoir un jeu de données et sa description via 2 fichiers :

- (1) **geoflow_entities** (cf. onglet Listing avec 1 ligne commenté + commentaires onglet sheet 2) : https://docs.google.com/spreadsheets/d/1DH3xqsrPkIGr5AO0rBjGhgcp5ij-Rh-NAupk3W_vyel/edit?usp=sharing
- (2) **geoflow_contact** : https://docs.google.com/spreadsheets/d/1juMFEIIFHv-NDZ2_ryAubBUkAlnD6V3WoFI_BZ8qWtg/edit?usp=sharing

Rem. Gsheet Commenté par E. Blondel

https://docs.google.com/spreadsheets/d/1u2MwmA7u6QG97dhJS3wr0qqcWAnKx_cVmKDr0_XOXo/edit#gid=0

IMPORTANT!!! les fichiers doivent être des gsheet et non xlsx. Si vous importez un fichier xlsx, enregistrez sous gsheet.

2. Installation et configuration R Studio

***Installer R Studio (ou R) avec la configuration** décrite sur le doc suivant :

https://docs.google.com/document/d/1-sfMmByB-Wrya9BBJD9YyRkw7jtx3HgK_D4suVRCSwY/edit?usp=sharing

- En théorie, il suffit d'installer le package « remotes » et de lancer la commande `install_github("eblondel/geoflow", dependencies = c("Depends", "Imports"))`
 - Sous linux, il manque souvent des paquets
 - Si ça ne marche pas pour certains, utilisation du serveur RStudio de Wilfried / organisateur de l'atelier
- Si vous possédez déjà un geoserver / geonetwork / dataverse ou un compte zenodo, pensez à rassembler vos login/mdp.

***Lire les ressources dans le github**

<https://github.com/eblondel/geoflow>

<https://github.com/eblondel/geoflow/wiki> == copier/coller les blocs de json pour avoir la configuration json global sous R studio / R

Fichier json configuré sur le geonetwork RZA (attention à la diffusion, car codes accès bed_elter) :

https://drive.google.com/file/d/1_XhgtC_73GUk4sopWLh56hjmRQddjiJ2/view?usp=sharing

Présentation de geoflow

Geoflow :

- philo initiale = de garder un outil simple accessible à un grand nombre d'utilisateurs
- permet de produire des fiches de MD aux normes ISO ou EML sans se poser de questions au préalable

Documents supports :

- (1) https://drive.google.com/file/d/1c0KbiPhim2mqToE9jWEX5ILr7XT_WdDa/view?usp=sharing
- (2) https://drive.google.com/file/d/1IST2iiWdUEW10B2RuyuHS_S8bIEkXrTA/view?usp=sharing

Metacat : un outil équivalent de geonetwork avec la norme EML

Pratiques

1. Exemple “Forêts anciennes des Pyrénées”

Exemple créé par Wilfried (doi:10.15454/BNEUSM)

Source :

<https://docs.google.com/spreadsheets/d/1u5icSPedTu5UHDFmmQJ3s0s2MJ3mlkw2QOZjwLaNKFs/edit?usp=sharing>

Résultat dans Geonetwork :

<https://dynids.toulouse.inra.fr:8443/geonetwork/srv/fre/catalog.search;jsessionid=73C85A5D892715DF222B10FE41BC52B2#/metadata/pyrenees>

Résultat dans le visualisateur cartographique Open Faire Viewer :

https://dynafor.toulouse.inrae.fr/dataviewer/?&baseview=World%20Imagery&views=%5B%22pid%3Dpyrenees%2Cstrategy%3Dogc_filters%2C%22%5D&extent=-3.0647675384868647,41.343502497685755,3.9664824615131353,44.760250544560755¢er=0.45085746151313544,43.051876521123255&zoom=9

Résultat dans dataverse : <https://doi.org/10.15454/BNEUSM>

Résultat dans Google dataset search :

<https://datasetsearch.research.google.com/search?query=cartographie%20de%20l'occupation%20ancienne%20des%20sols&docid=NmG7FUCIcvIPEjS%2FAAAAAA%3D%3D>

2. Applications

Configurer le json avec TRUE ou FALSE suivant les objectifs.

Pour lancer le programme (fenêtre R):

- library (geoflow)
- *library(RPostgreSQL) optionnel*
- initWorkflow ("config_geoflow_lter.json")
- executeWorkflow ("config_geoflow_lter.json")

Pour les **messages d'erreurs**, les renseigner au fichier ci-après :

<https://docs.google.com/document/d/1jOw6VUa1MM2sYv777oJxp1dOcowlyebRAoDlcAt13pk/edit?usp=sharing>

Recommandations :

- mettre dans le titre TEST pour distinguer les fiches sur le geonetwork. Penser à les supprimer si pas d'utilisation.
- **Attention un _ doit toujours être suivi d'un retour à la ligne et il ne doit jamais y avoir un espace après le _ !**

Constat sur la production des fiches :

- les étendues spatiales et temporelles ne s'affichent pas
- les vignettes ne s'affichent pas
- le vocabulaire AGROVOC n'apparaît pas
- quid dans guide de recommandation : mettre des mots clefs dans GENERAL du type ZAS, Zone Atelier Seine...
- SRID = 2154 pour Lambert 93 et 4326 pour WGS84

Rappel : Notions clefs

Fiche ISO 19 115 : norme conceptuelle pour les fiches de MD

Fiche ISO 19 139 : norme d'interprétation opérationnelle de l'ISO 19.115 pour publier des fiches MD dans Geonetwork/ implémentation XML

Fiche ISO 19 1109 : norme conceptuelle pour un dictionnaire des variables

à compléter si besoin

Questions / Tests à faire

- Agrovoc : mettre le thesaurus dans le geonetwork pour pouvoir voir les termes ?
- Développement :
 - Est-ce que EML4R est terminé ? Que faire d'un fichier EML ? pousser automatiquement vers PNDB ?
 - Peut-on avoir un aperçu du **rapport "validateur_inspire"** ? peut-on imaginer un affichage de ce genre (QA Tool de DEIMS) : <https://qa.deims.org/?id=e5cf07bb-f849-4034-babe-ff0a47cbc33a>
 - Peut-on / serait-il intéressant d'aller plus loin sur le contrôle qualité des MD générées ?
- Langues :
 - Est-il possible d'écrire **pour un même identifiant** d'une fiche de MD en anglais et en français pour actionner l'**option bilingue du geonetwork** ?
 - Peut-on avoir des données dans une langue (ex. fr) et des métadonnées dans une autre langue (ex. eng) ?
- De manière générale, quels sont les **attributs/champs dans chaque colonne** (ex. Description > abstract) et quelles sont les valeurs par défaut ? ok visible dans sa feuille guideline (abstract; purpose; info)...toutefois très peu de définition / explication dans le doc.
- **Colonne Type** : Améliorer les formats acceptés (plus que juste Dataset)

	resource_type	spatialRepresentationType = Type	resource_format
	<i>type de donnée :</i> séries de données géographiques, ensemble de séries de données géographiques, services de données géographiques.	<i>type de représentation</i>	<i>format de la ressource</i>
Datacite	resourceTypeGeneral (1er lettre en MAJUSCULE) Audiovisual, Collection, Dataset, Event, Image, InteractiveResource, Model, PhysicalObject, Service, Software, Sound, Text, Workflow, Other	n/a	n/a
Geoflow	<i>(1er lettre en MAJUSCULE)</i> Dataset <i>(Pour l'instant et par défaut, mettre "dataset"). Est-ce possible de se base aussi sur la nomenclature Iso (dans les cas de la generation de fiche Iso) ?</i>	<i>Type de données</i> vecteur, texte, points géographiques, tabulaires <i>meme question Pourquoi pas grid, TIN, StereoModel... ?</i>	n/a
Iso 19115	hierachyLevel (ressource type) (1ere lettre en minuscule) dataset, series, service, attribut, fieldsession, ... (liste est + longue)	spatialRepresentationType : vector, grid, text table TIN, Stéréomodel, vidéo	resourceFormat/ presentationForm : tableDigital, mapDigital, imageDigital, profilDigital, videoDigital, documentDigital, modelDigital
EML (PNDB / GBIF)	A completer	A completer	A completer
Pivot OZCAR-THEIA	<i>a verifier</i>	<i>(modèle pivot issu d'une "presque concaténation traduite" Iso19115 et Datacite</i> <i>'Type Données'</i> Numeric, Vector, Raster, Photo, Video, Text, Audio, Other	<i>a verifier</i>

- **Colonne Date** : préciser le format date/heure ou si c'est libre
- Tester s'il est possible de **réinjecter les infos d'une fiche de MD** déjà réalisée (ex. fiche gabarit (MetadataReasy) produit par Christine Plumejaud)
- **Format pivot** ?!
- **Catalogue des données** : décrire les variables et les modalités des variables; peut-on le faire indépendamment d'un fichier shape avec table attributaires
Existe-il un fichier CSV pour remplir un catalogue d'attribut Iso19.110 sans passer par le Catalogue ? Rattacher à un protocole ?
- **référentiels** :

Perspectives

*Plan de communication sur l'outil auprès des ZAs == 1 page + ppt 5 diapos, à transmettre avec le CR de l'AG == deadline 21 nov.

*Ecrire un article :

- Envoyer mail à Julien et Emmanuel pour avoir leur accord sur publication >> Wilfried Heintz
- Présenter geoflow et les applications à l'échelle du RZA == à discuter dans les grandes lignes lors de l'AG == fournir le déroulé des idées + la 10aine de publi clefs pour l'intro == pour E-ENVIR2021

Chaque use-case montrer un exemple

- Publier GeoFlow : <https://www.journals.elsevier.com/computers-and-geosciences/>
- Publier notre approche / visée eLTER : <https://www.journals.elsevier.com/science-of-the-total-environment>
- Portail Theia/OZCAR est maintenant publié dans Hydrological Sciences Journal.

*Rédaction du tutoriel, envisager :

- une entrée par type de format de données
- une entrée par type de données
- une entrée par objectif (production de fiche de MD, entrepôt...)

