

# Outils R pour créer et mettre à jour des métadonnées et des accès standardisés sur des jeux de données

**Objectifs:** présenter une méthode, entièrement implémentée en R, pour standardiser l'écriture des métadonnées et la publication des données associées (standards OGC, EML, RDF pour les formats de métadonnées et WMS et WFS pour les protocoles d'accès). L'atelier montrera des exemples de codes qui utilisent uniquement des packages R pour écrire les métadonnées à partir de différents types de jeux de données (stockées dans des bases de données SQL type PostgreSQL ou dans des fichiers NetCDF accessibles sur des serveurs Thredds) et des packages R pour les publier (basés sur les APIs de Geonetwork et Geoserver). Une chaîne de traitements (développée pour le projet H2020 BlueBridge) sera utilisée pour fournir des codes qui automatisent la mise à jour des métadonnées et des jeux de données associés (quand les sources de données ont elles-mêmes été mises à jour).

- **jour 1:** présentation orientée utilisateurs (objectifs généraux, présentations des standards et démonstrations simples d'utilisation des packages R),
- **jour 2:** présentation pour des profils utilisateurs du type gestionnaires de données / programmeurs R (présentation détaillée du workflow et manipulation avec des TD pour tester et adapter les codes en ligne à d'autres sources de données).

**Période et Lieu:** Jeudi 8 et vendredi 9 février 2018 à Montpellier, à [Agropolis International](#) (20 - 30 minutes depuis le centre ville en tramway et bus, arrêt de bus juste devant le centre: cf [plan d'accès](#)).

**Frais d'inscription:** gratuit, repas à l'IRD (à la charge des participants)

**Matériel:** prendre son PC, le travail se fera en ligne sur l'infrastructure fournie par le projet H2020 BlueBridge (RStudio server).

**Participants:** Formulaire de pré-inscription :

<https://questionnaire.umontpellier.fr/index.php/334681?lang=fr>

**Liens vers les packages R:**

- Standards OGC: [Geometa](#) (R tools to write, read & validate geographic metadata: OGC/ISO 19115, 19119, 19110, 19139), [Geonapi](#): R Interface to GeoNetwork API, [Geosapi](#): R interface to GeoServer REST API
- EML: package [EML](#)
- NetCDF: package [ncdf4](#)

## *Jeudi 8 Février*

### *Atelier métadonnées: exemple du projet BlueBridge*

09:00	<b>Accueil</b>		
09:30 - 10:00	<b>Introduction / objectifs: infrastructure, workflow, sources de données, VRE et matériel pour le groupe:</b>	<i>Infrastructure BlueBridge: VRE, inscription et accès à RStudio Server Sources de données: OPeNDAP (Thredds..), SQL (PostGIS..), tabulaires (CSV / google spreadsheet..)</i>	<i>E. Blondel / J.Barde</i>
10:00 - 10:30	<b>Présentation générale du workflow</b>	<i>Comment le configurer pour l'amorcer et l'utiliser pour faire des mises à jour des métadonnées...</i>	
<b>Pause Café</b>			
11:00 - 11:30	<b>Métadonnées statiques VS dynamiques: interfaces pour la saisie ou le calcul</b>	<i>Distinguer les éléments de métadonnées automatisables et ceux qui ne le sont pas.. Fournir les informations dans un google spreadsheet ou requêtes SQL / OPeNDAP</i>	
11:30 - 12:30	<b>Standards de métadonnées: choisir des standards, comment lire les normes, mapping et packages R</b>	<i>Exemple de la norme ISO 19115/110/119: diagrammes UML, schémas XML et exemples de métadonnées XML conformes. Autres exemples (EML..)</i>	
<b>Pause Repas</b>			
14:00 - 15:00	<b>TP: Utiliser les codes R et RStudio Server pour écrire une métadonnée simple..</b>	<i>Présentation d'échantillons de codes et l'environnement de compilation (IDE RStudio server)</i>	
15:00 - 15:30	<b>TP: Publier la métadonnée dans GN et la donnée dans Geoserver, Thredds, IPT Les outils associés: SDI components</b>	<i>Outils: Geonetwork, Geoserver, Thredds</i>	
<i>Pause Café</i>			
16:00 - 17:30	<b>TP: Introduction aux codes utilisés dans le workflow pour la gestion de (méta)donnée</b>	<i>Présentation générale: sans rentrer dans le détail des codes R, on présentera comment produire les métadonnées essentielles (titre, résumé, mots-clés, contacts...) à partir d'un google doc (ou table SQL) et on ajoutera quelques métadonnées plus fines à titre d'exemple.</i>	
18:00	<b>Fin de la journée</b>		

## Vendredi 9 Février

### Atelier métadonnées: exemple du projet BlueBridge

09:00 - 09:30	<b>Présentation détaillée du workflow</b>	Détail sur le fichier de configuration du workflow (Json) et sur les différentes étapes et codes associés	<i>E. Blondel / J.Barde</i>
09:30 - 10:30	<b>Ecrire une métadonnée 19115/19110 sur une base de données Postgis</b>	Adapter le workflow pour décrire une nouvelle base de données: connexion de la base, description des requêtes	
10:30 - 11:00	<b>Pause Café</b>		
11:00 - 12:00	<b>Ecrire une métadonnée 19115 sur un fichier NetCDF (accès OPeNDAP sur serveur Thredds)</b>	Adapter le workflow pour décrire des fichiers NetCDF ou NCML : accès OPeNDAP, description des jeux de données, créer des fichiers virtuels avec NCML pour modifier ou enrichir les métadonnées	<i>Julien Barde</i>
12:00 - 13:00	<b>Interagir en R avec Geonetwork et Geoserver</b>	Publier les métadonnées et créer des Web Services associés (WMS/WFS...)	<i>E. Blondel</i>
13:00 - 14:00	<b>Pause Repas</b>		
14:00 - 15:30	<b>TP: gérer son propre workflow</b>	Décrire une base PostGIS et le catalogue de requêtes associé ou un ensemble de fichiers NetCDF sur un serveur Thredds	<i>E. Blondel / J.Barde</i>
15:30 - 16:00	<b>Pause Café</b>		
16:00 - 17:30	<b>TP: étendre les codes</b>	Utiliser les codes pour écrire d'autres métadonnées (EML, RDF, NetCDF). Comment étendre l'approche à d'autres standards ou outils (Thredds, SOS..).	<i>E. Blondel / J.Barde</i>
18:00	<b>Fin de la journée</b>		