

OUTILS DE CRÉATION DE MÉTADONNÉES

À QUEL MOMENT DÉCRIRE SES DONNÉES ?

Pour bien gérer vos données, il est recommandé de les décrire au fur et à mesure de l'avancée du projet de recherche. Cela vous évitera d'avoir à tout décrire une fois le projet terminé. Ainsi vos métadonnées seront déjà prêtes lorsque vous aurez à les renseigner au moment du dépôt dans l'entrepôt.

QUELLE MÉTHODE UTILISER POUR DÉCRIRE SES DONNÉES ?

Il n'y a pas de méthode unique pour créer ses métadonnées. Les informations à renseigner dépendent du contexte dans lequel elles s'inscrivent, par exemple :

- des pratiques de la discipline,
- des caractéristiques de l'objet décrit par les données (un gène, une fleur, une montagne...),
- des caractéristiques techniques des données (format de fichiers...),
- du ou des services envisagés pour la diffusion, la sauvegarde et l'archivage des données,
- du degré de visibilité et de réutilisabilité souhaité,
- d'éventuelles obligations réglementaires (ex : directive INSPIRE).

Vous pouvez utiliser des schémas ou standards existants.

Si les schémas existants ne permettent pas de décrire assez précisément vos données, vous pouvez créer votre propre schéma, spécifique à votre projet. Pour cela, utilisez comme modèle un schéma existant et établissez une table de correspondances (mapping) entre vos métadonnées et ce schéma de référence.

EXEMPLES D'OUTILS POUR CRÉER SES MÉTADONNÉES

Des outils peuvent vous aider à créer vos métadonnées. Certains sont génériques, d'autres spécifiques à un usage ou une discipline. Nous allons en aborder deux : le premier est un modèle de champs à remplir préconisé par l'[Observatoire Terre Environnement Lorraine](#) (OTELo) et l'[InistCNRS](#). Le deuxième est un outil fourni par [Data Cite](#). Ils utilisent respectivement les formats CSV et XML qui sont adaptés pour l'échange de données numériques.

LE MODÈLE OTELO – INIST-CNRS



OTELo et l'Inist-CNRS ont rédigé un guide de bonnes pratiques incluant une partie sur les métadonnées. Vous pouvez télécharger le guide complet ici :

[GUIDE DE BONNES PRATIQUES : Gestion et valorisation des données de recherche](#)

Voici un extrait tiré de ce guide, dans lequel sont données quelques préconisations :

- Associer un fichier de métadonnées à chaque fichier de données.
- Préférer un format tabulé (p. ex. CSV) à un format texte (Read-me file : p. ex. PDF) si possible pour une exploitation automatique des données et le dépôt des données dans des entrepôts.

- Utiliser des [normes, standards, vocabulaires contrôlés](#) (lexique, thésaurus) pour faciliter le partage et l'intégration de données avec votre communauté scientifique.
- Privilégier la langue d'usage de votre communauté de recherche pour faciliter la publication de vos données.

LE MODÈLE OTELO INIST-CNRS EN CSV

A	B
Métadonnées	Format
1 Nom du projet	Texte libre
2 Titre développé	Texte libre
3 Nom du porteur du projet	Nom prénom
4 Institution	Texte libre
5 Titre du jeu de données	Texte libre. Le titre doit être explicite.
6 Identification du jeu de données	Nom du fichier
7 Description du jeu de données	Description simplifiée du contexte de production de données
8 Date de création	YYYY-MM-DD (Norme ISO8601, W3CDTF)
9 Responsable du fichier	Nom prénom
10 Laboratoire de rattachement du responsable de fichier	Texte libre ou liste contrôlée
11 Adresse mèl du responsable de fichier	
12 Producteur des données	Nom prénom
13 Laboratoire de rattachement du producteur de données	Texte libre ou liste contrôlée
14 Adresse mèl du producteur de données	
15 Langue des métadonnées	fr ou en (Norme ISO639-1)
16 Thématique scientifique	Catégorie issue des Thèmes INSPIRE (cf. Annexe I) Mots-clés issus de thésaurus ou classifications
17 Nom du point de collecte ou d'observation	Texte libre
18 Géolocalisation du point de collecte/prélèvement	Coordonnées GPS
19 Système de coordonnées utilisé	Lambert
20 Date/heure de collecte	YYYY-MM-DD ou YYYY-MM-DDThh:mmTZD (TZD : désigne zone Horaire. Pour la France +01 :00 en hiver, +02 :00 en été) Norme ISO8601, W3CDTF
21 Nom du protocole*	
22 Version du protocole*	
23 Description du protocole*	Description ou référence bibliographique
24 Paramètres du protocole*	Paramètres appliqués dans la méthode
25 Composants du protocole*	Instruments, logiciels ou scripts
26	
27	
28	

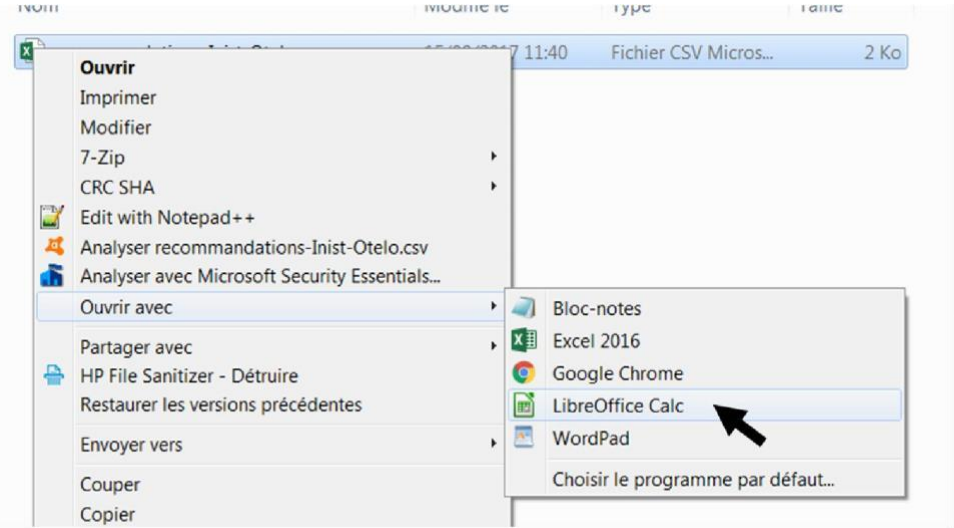
Voici le modèle que vous pouvez utiliser. Sélectionnez les champs de métadonnées qui semblent adaptés à votre contexte, à l'entrepôt dans lequel vous déposez vos données et qui apporteront une information minimale et suffisante pour comprendre et reproduire vos données.

[Télécharger le modèle](#)

UTILISER LIBREOFFICE

Nous vous conseillons tout d’abord d’utiliser LibreOffice Calc pour ouvrir correctement ces fichiers.
La suite bureautique LibreOffice est [disponible ici](#).
Une fois LibreOffice installé, ouvrez le modèle Otelo – Inist-CNRS en suivez les instructions ci-dessous.

1. OUVRIR LE FICHIER AVEC CALC



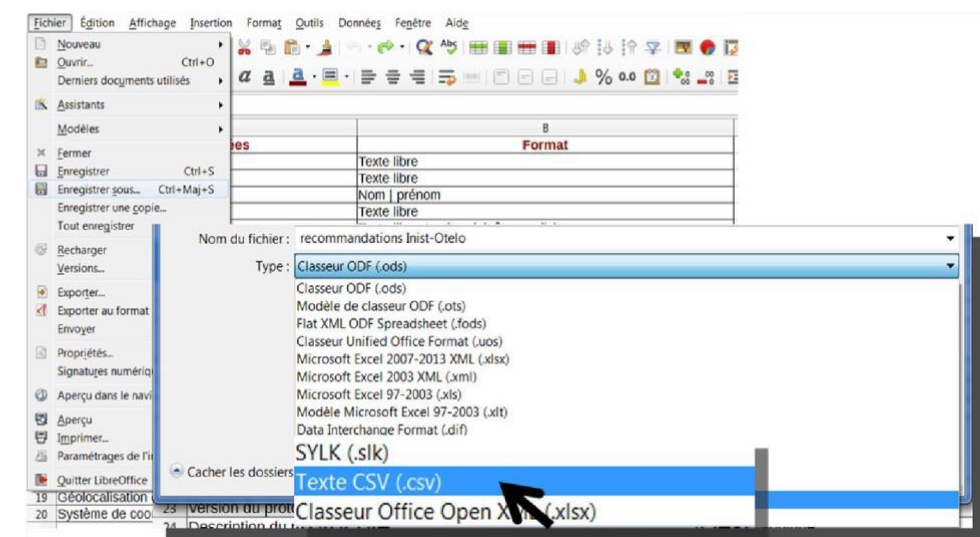
2. REMPLIR LES CHAMPS DE LA COLONNE FORMAT

	A	B
1	Métadonnées	Format
2	Nom du projet	Texte libre
3	Titre développé	Texte libre
4	Nom du porteur du projet	Nom prénom
5	Institution	Texte libre
6	Titre du jeu de données	Texte libre. Le titre doit être explicite.
7	Identification du jeu de données	Nom du fichier
8	Description du jeu de données	Description simplifiée du contexte de production de données
9	Date de création	YYYY-MM-DD (Norme ISO8601, W3CDTF)
10	Responsable du fichier	Nom prénom
11	Laboratoire de rattachement du responsable de fichier	Texte libre ou liste contrôlée
12	Adresse mèl du responsable de fichier	
13	Producteur des données	Nom prénom
14	Laboratoire de rattachement du producteur de données	Texte libre ou liste contrôlée
15	Adresse mèl du producteur de données	
16	Langue des métadonnées	fr ou en (Norme ISO639-1)
17	Thématique scientifique	Catégorie issue des Thèmes INSPIRE (cf. Annexe I) Mots-clés issus de thésaurus ou classifications
18	Nom du point de collecte ou d’observation	Texte libre
19	Géolocalisation du point de collecte/prélevement	Coordonnées GPS
20	Système de coordonnées utilisé	Lambert

3. AJOUTER OU SUPPRIMER DES LIGNES

	A	B
1	Métadonnées	Format
2	Nom du projet	Texte libre
3	Titre développé	Texte libre
4	Nom du porteur du projet	Nom prénom
5	Institution	Texte libre
6	Titre du jeu de données	Texte libre. Le titre doit être explicite.
7	Identification du jeu de données	Nom du fichier
8	Description du jeu de données	Description simplifiée du contexte de production de données
9	Date de création	YYYY-MM-DD (Norme ISO8601, W3CDTF)
10	Responsable du fichier	Nom prénom
11	Responsable de	Texte libre ou liste contrôlée
12	er	
13	er	Nom prénom
14	ucteur de	Texte libre ou liste contrôlée
15	ées	
16		fr ou en (Norme ISO639-1)
17		Catégorie issue des Thèmes INSPIRE (cf. Annexe I)
18		Mots-clés issus de thésaurus ou classifications
19	vation	Texte libre
20	prélèvement	Coordonnées GPS
21		Lambert
22		YYYY-MM-DD ou YYYY-MM-DDThh:mmTZD (TZD : désigne zone Horaire. Pour la France +01 :00 en hiver, +02 :00 en été)
23	Date/heure de collecte	

4. ENREGISTRER AU FORMAT CSV



DATA CITE METADATA GENERATOR



[DataCite](#) a mis au point un outil permettant de générer ses métadonnées au format XML. Bien que spécifique au schéma DataCite, l'outil reste assez généraliste et peut vous être utile quel que soit votre contexte. Vous pouvez [utiliser la version originale qui est disponible et mise à jour sur Github](#).

TESTER DATA CITE METADATA GENERATOR

Une version réadaptée est disponible en ligne ici. Une fois que vous avez terminé de remplir les champs, vous pouvez exporter vos métadonnées en XML en cliquant sur « Ready to save » puis « Click here to save: metadata.xml ».

[TESTER DATA CITE METADATA GENERATOR](#)

DataCite Metadata Generator - Kernel 4.0

Mandatory Elements (éléments obligatoires)

DOI:
Vos données ont-elles déjà un DOI?

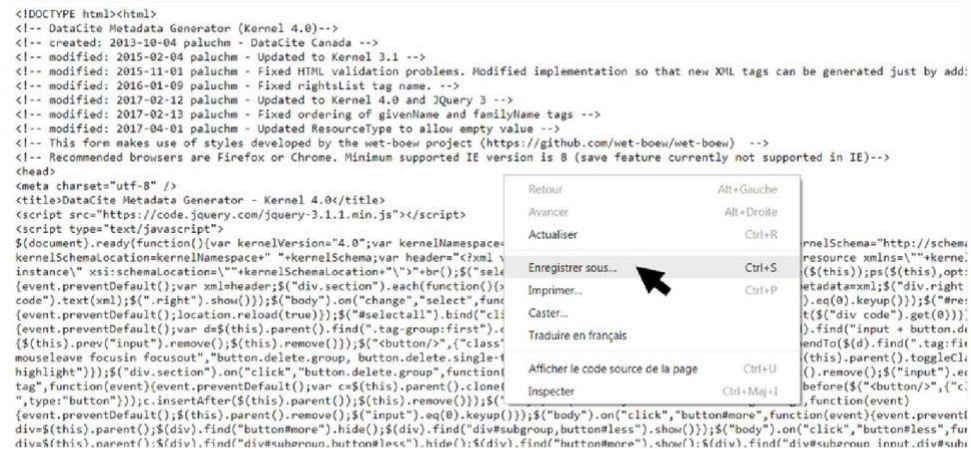
Title(s):
Indiquez le nom ou le titre de votre ressource. Si besoin, ajoutez des variantes au titre en précisant le type de titre.

UTILISER NOTEPAD++

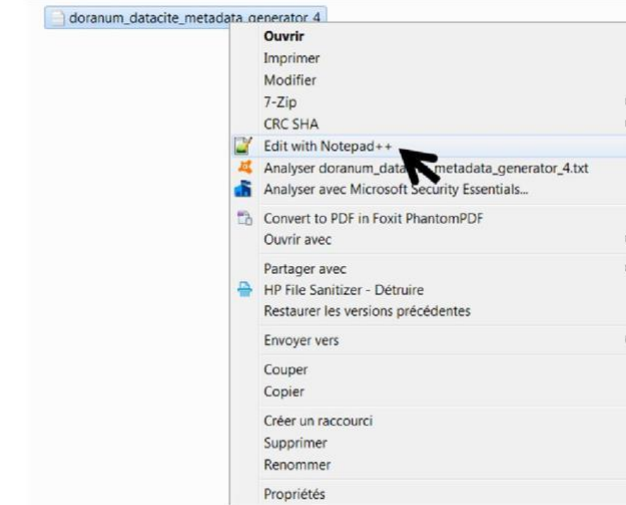
Si vous n'avez pas d'éditeur de texte, vous pouvez télécharger [Notepad++](#), un logiciel libre et qui vous permettra de gérer vos fichiers sous de nombreux formats (XML, TXT, HTML...).

Une fois Notepad++ installé, [téléchargez le fichier Datacite Metadata Generator TXT](#). Le fichier va s'ouvrir dans un nouvel onglet, suivez ensuite les instructions ci-dessous :

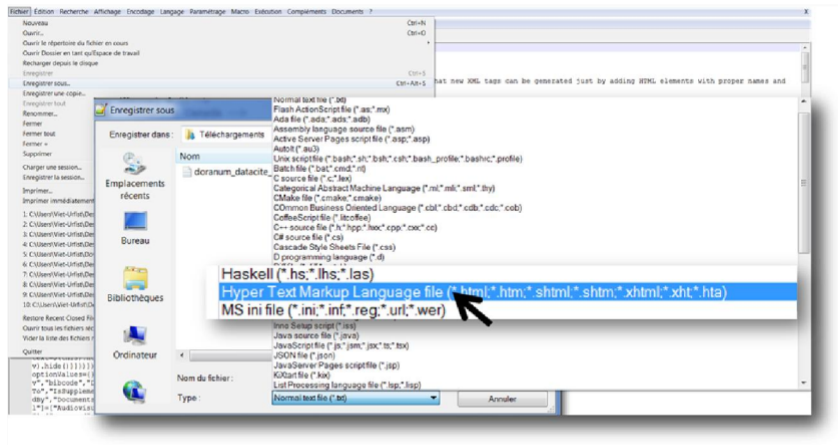
1. ENREGISTRER LA PAGE SUR VOTRE ORDINATEUR



2. OUVRIR AVEC NOTEPAD++



3. ENREGISTRER EN HTML



4. OUVRIR LE FICHIER HTML POUR LANCER L'OUTIL

DataCite Metadata Generator - Kernel 4.0

Mandatory Elements (éléments obligatoires)

DOI:
Vos données ont-elles déjà un DOI?
[DOI]

Title(s):
Indiquez le nom ou le titre de votre ressource. Si besoin, ajoutez des variantes au titre en précisant le type de titre.
[TITLE]

[titleType]

D'AUTRES OUTILS DE CRÉATION DE MÉTADONNÉES

Nous ne pouvons pas tous les citer, mais pour information, voici un répertoire collaboratif de schémas de métadonnées utilisés dans le monde de la recherche, mis en place par un groupe de travail de ResearchData Alliance : <https://rdamsc.bath.ac.uk/>

Vous connaissez ou utilisez d'autres outils pour créer des métadonnées ? Faites-nous en part en commentaire !