



GUIDE DU CURATEUR DE
L'ESPACE INSTITUTIONNEL

Table des matières

| | |
|---|---|
| Rôle et responsabilité du curateur..... | 2 |
| Vérifier les métadonnées..... | 3 |
| Valider un jeu de données..... | 6 |
| Retourner à l’auteur..... | 6 |
| Publier..... | 6 |
| Gérer les versions supérieures à la 1.0..... | 7 |
| Créer des modèles de saisie de jeux de données..... | 8 |

De DOREL à DOREL

DOREL (DONnées de la REcherche Lorraines) est, jusque fin 2023, l’entrepôt institutionnel pour les données de la recherche de l’Université de Lorraine et du site lorrain.

Le dépôt dans l’espace institutionnel DOREL de l’entrepôt national Recherche Data Gouv, qui le remplace, est recommandé dès à présent.

Il est accessible ici : <https://entrepot.recherche.data.gouv.fr/dataverse/dorel-lorraine>

Rôle et responsabilité du curateur

Il est proposé à chaque laboratoire de définir une ou deux personnes en charge de la curation des données de la collection dataverse attribué à l’unité de recherche. Dans ce cas, l’équipe Données de la recherche ne sera plus directement en charge de la curation ; et la procédure de création et de publication se modifie légèrement, pour intégrer la phase de validation dans le flux de travail du laboratoire.

- Le curateur doit :
 - **vérifier les métadonnées d’un jeu de données avant la publication** selon une check list précisée plus loin dans ce document ;
 - procéder à la publication du jeu de données ;
 - valider les éventuelles versions suivantes du jeu de données, si celui-ci évolue.
- Le curateur peut créer et éditer à la demande des **modèles de jeux de données**.
- **Le curateur peut demander à l’équipe Données de la recherche de créer des liens entre collections.**
- Le curateur peut demander à l’équipe Données de la recherche de créer de nouvelles collections (cas couverts : projets de recherche, équipes de recherche, plateformes expérimentales)

Le présent guide couvre ces différents points.

Vérifier les métadonnées

Lorsqu'un déposant crée des jeux de données, vous recevez une alerte mail en tant que curateur pour la publication de ces jeux. Vous devrez alors vérifier les points ci-dessous. Ceux-ci correspondent aux champs « obligatoires » et « recommandés » du guide du déposant.

⇒ **Dans le cas où une métadonnée obligatoire est manquante ou incorrectement formulée, le dépôt ne peut être accepté.**

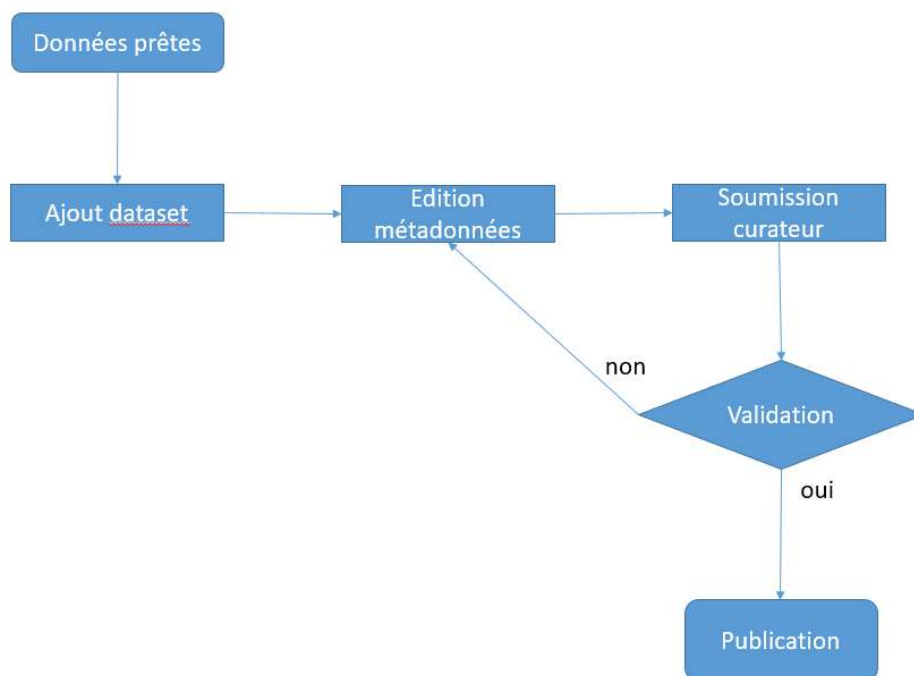
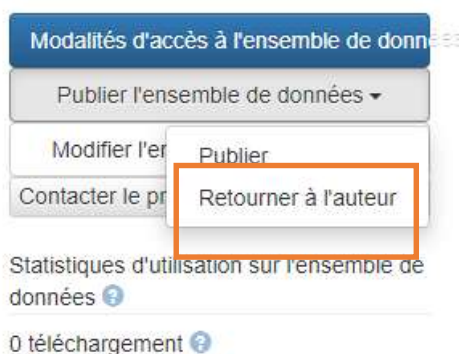
| Champ | Sous-champ | Consignes | Exemple | Statut |
|------------------------------|--------------------|--|---|--------------------|
| Titre | Titre | Vérifier que le titre est suffisamment descriptif (type de données, contexte, période de collecte et/ou localisation géographique...). | Consumer perceptions of infant food safety in France. | Obligatoire |
| Auteur | Auteur | Vérifier que le nom est sous la forme Nom, Prénom (ex. Hugo, Victor). | Rigobert, Elisa | Obligatoire |
| | Affiliation | Vérifier que l'affiliation est renseignée avec le sigle du laboratoire sans espaces ni ponctuations (CRULH, IJL...). | CREM | Obligatoire |
| | Schéma identifiant | Type (ou « schéma ») de l'identifiant. | ORCID | Recommandé |
| | Identifiant | N° de l'identifiant. | 1111-2222-1234-5678 | Recommandé |
| Producteur | Nom | Pour la plupart des cas : nom de l'institution, à répéter autant de fois qu'il y a d'institutions responsables | Université de Lorraine | Recommandé |
| | Nom abrégé | Sigle ou autre | UL | |
| | Affiliation | Dans le cas d'unités affiliées à des institutions plus grandes. | | Optionnel |
| | URL | | https://www.univ-lorraine.fr | |
| | URL du logo | | https://www.univ-lorraine.fr/wp-content/uploads/2020/10/LOGO_UL.png | |
| Personne-ressource (Contact) | Nom | Vérifier que le nom est sous la forme Nom, Prénom (ex. Hugo, Victor). | Durand, Jacqueline | Recommandé |
| | Affiliation | Acronyme du laboratoire; Tutelle 1, Tutelle 2... ; Pays <i>Ex. : LMOPS, Université de Lorraine, Centrale Supelec ; France</i> | CREM | Recommandé |
| | Courriel | Vérifier que le courriel est correct | jacqueline.durand@univ-lorraine.fr | Obligatoire |
| Description | Texte | Court passage descriptif dans la langue de dépôt. L'objectif est globalement le même que pour un résumé d'article : ni trop court ni trop long, il doit inclure l'essentiel de la spécificité et de la caractérisation propre du jeu de données (objet, nature, portée, approche utilisée...). | | Obligatoire |
| | Date | S'il n'y a qu'une description, supprimer le champ. | | |
| Sujet | | Vérifier que le déposant a bien sélectionné un ou plusieurs sujets. | Sciences sociales | Obligatoire |

| | | | | |
|---|----------------------------|---|---|--------------------------|
| Mot-clé | Terme | Vérifier que le déposant a choisi un ou plusieurs termes pour décrire son jeu de données (thématique, objet, contenu). Ces termes peuvent être libres ou, de préférence, être issus de thésaurus ou de vocabulaires existants. Dans ce cas, il est recommandé de préciser le vocabulaire employé ainsi que l'adresse à laquelle il est accessible en ligne. | Milk yield Rendement laitier | Obligatoire |
| | Vocabulaire | Si l'un des deux champs « Vocabulaire » ou « Adresse URL du vocabulaire » est rempli, l'autre doit l'être également. Attention ce n'est pas l'URL du mot précis qui est demandée mais celle du vocabulaire global. | Agrovoc | Recommandé |
| | Adresse URL du vocabulaire | | http://www.fao.org/agrovoc/fr | Recommandé |
| Publication connexe | Référence | Citation bibliographique de la publication utilisant le jeu de données. | | Recommandé |
| | Type d'identifiant | Renseigner le type d'identifiant concerné. DOI (le plus fréquent) Pour les dépôts HAL, laisser vide. | | Recommandé |
| | Numéro d'identification | Copier-coller l'identifiant (le DOI par exemple). HAL : copier l'identifiant sous la forme hal-xxxxxxx (ex. hal-00083055) | | Recommandé |
| | Adresse URL | Copier-coller l'URL exacte de la publication. | | Recommandé |
| Informations sur le projet | Acronyme | Si les données sont produites dans le cadre d'un projet. Acronyme ou sigle du projet. | | Recommandé si applicable |
| | Titre | Acronyme ou nom développé | | |
| | Identifiant | | | |
| | Site web du projet | | | |
| | Work Package | Si applicable | | |
| | Tâche | | | |
| Langue | | Choisir une ou plusieurs langues dans la liste déroulante. Ce champ doit servir à qualifier la ou les langue(s) utilisée(s) dans les fichiers de données. A distinguer de la langue des métadonnées. | | Obligatoire |
| Date de production | | Date à laquelle les données ont été produites. Si les données ont été produites sur plusieurs mois (voire années) mettez la date de début ? | 2020-11-01 | Recommandé |
| Endroit de production | | Lieu où les données ont été produites. | | Recommandé |
| Contributeur | Type | Choisir un rôle dans la liste déroulante. | | Recommandé |
| | Nom | Collaborateurs, différents des auteurs, ayant contribué au jeu de données. Nom sous la forme Nom, Prénom OU renseigner l'institution contributrice. | | Recommandé |
| Renseignements sur la subvention | Organisme subventionnaire | Vérifier que les organismes sont écrits de la bonne manière : <ul style="list-style-type: none"> Financement ANR : Agence nationale de la recherche | | Recommandé si applicable |

| | | | | |
|-----------------|------------------|---|--|--------------------------|
| | | <ul style="list-style-type: none"> Financement Horizon 2020 ou Europe : Commission Européenne Autre financement (DRAC, collectivité,...): indiquer le plus clairement possible le nom du financeur, en développant. Ex. : Caisse des dépôts et consignations. | | |
| | N° de subvention | Le numéro de subvention ou de contrat lié au financement. | | Recommandé si applicable |
| Type de données | | Vérifier la conformité : survey data, census/enumeration data, aggregate data, clinical data, event/transaction data, program source code, machine-readable text, administrative records data, experimental data, psychological test, textual data, coded textual, coded documents, time budget diaries, observation data/ratings, process-produced data, or other. | | Recommandé |
| Logiciel | Nom Version | Nom et version du logiciel utilisé. Ex. : OriginLab 2021b. | | Recommandé |

Valider un jeu de données

Une fois les vérifications faites, vous pouvez valider la publication ou renvoyer à l’auteur pour révision.



Retourner à l’auteur

A noter : l’interface ne permet pas de préciser les raisons de ce « refus » provisoire, il est donc fortement recommandé d’envoyer un mail séparé avant le renvoi automatique à l’auteur.

Publier

La publication du jeu de données entraîne plusieurs conséquences immédiates.

- Le jeu de données est visible par tous, y compris les utilisateurs non connectés.
- Il passe du statut « Draft » (Brouillon) au statut « Published » dans Datacite et est automatiquement passé en version 1.0.
- Un DOI pérenne est attribué au jeu de données, ce qui permet notamment de lier le jeu de données à des publications scientifiques.

Gérer les versions supérieures à la 1.0.

Deux règles régissent la numérotation des versions au-delà de la 1.0.

Ces règles peuvent se ramener à une seule très simple : « si la citation du jeu de données change, le n° de version majeure doit changer aussi ».

1. **Modification des métadonnées ou des conditions d'accès (licence) sans modification de données** : le choix est donné au moment de la publication de mettre à jour le jeu de données :
 - a. avec le même n° de version
 - b. avec un n° de version mineur (1.1, 1.2...)
 - c. avec un n° de version majeur (2.0, 3.0...)

2. **Modification des données avec ou sans modification des métadonnées** : une version majeure (2.0, 3.0...) est imposée.

A chaque modification de version, le curateur reçoit une notification par mail.

Le circuit de validation est le même que lors de la création d'un jeu de données. Il peut publier le jeu ou le retourner à l'auteur pour corrections.

Créer des modèles de saisie de jeux de données

Le logiciel Dataverse ne permet pas d'ouvrir les droits sur l'édition de modèles de métadonnées sans ouvrir également des droits d'administration sur la collection.

Sur l'ensemble des droits accessibles, **merci de n'utiliser QUE la fonction de conception de modèles**. Les autres manipulations sur la collection doivent passer par l'équipe DOREL.



Les modèles permettent de ne pas ressaisir les métadonnées récurrentes pour des jeux de données différents ex : auteur identique, date identique, mots clés identiques

Sur l'écran d'accueil, choisir la collection dataverse concernée :



Depuis la page de la collection, cliquer sur « Modifier » puis sur « Modèles de jeux de données » dans le menu déroulant. La case « Intégrer des modèles provenant... » est cochée par défaut : les modèles listés sont ceux de la collection dataverse active, plus ceux de la collection parente ou de l'ensemble des collections parentes.

Intégrer des modèles provenant de Entrepôt Bac à Sable + Créer un modèle de jeu de données

| Nom du modèle ↕ | Date de création ↕ | Utilisation(s) ↕ | Action |
|--------------------|--------------------|------------------|--|
| Kinetic mechanisms | 1 juil. 2021 | 0 | Définir par défaut <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Modifier le jeu de données <input type="checkbox"/> |
| Spectrographe | 14 déc. 2020 | 0 | Définir par défaut <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Modifier le jeu de données <input type="checkbox"/> |

Si aucun modèle ne correspond à votre besoin, cliquez sur « Créer un modèle de jeu de données ».

Toutes les métadonnées disponibles peuvent être pré-remplies.

Nom du modèle *

*Les astérisques indiquent les champs de métadonnées que les utilisateurs doivent obligatoirement remplir lors de l'ajout d'un jeu de données à ce dataverse.

Métadonnées de la référence bibliographique ^

Titre *

Sous-titre

Autre titre

Autre URL

Autre identifiant

Organisme **Identifiant** +

Auteur *

Nom * **Affiliation** +

Schéma de l'identifiant **Identifiant**

Personne-ressource *

Nom **Affiliation** +

Courriel *

Description *

Ce champ prend en charge uniquement certaines balises HTML.

Texte *

Il vous suffit de choisir celles que vous souhaitez voir pré-remplies pour les jeux de données qui utiliseront le modèle.

- Note : bien que toutes les métadonnées avec une étoile rouge soient obligatoires lors de la création d'un jeu de données, cette obligation de remplissage ne s'applique pas ici : vous pouvez les laisser vides, ou les pré-remplir si besoin.
- Veillez à donner à votre modèle un titre qui soit parlant pour vos collègues. Exemple : « données du projet XXXX » ; « photographies du projet XXXX » ;...

Astuce : si vous souhaitez qu'un champ n'apparaissant normalement que lors de la finalisation du dépôt par le déposant soit présent dès la première saisie, comme par exemple le logiciel utilisé, il vous suffit de vous placer sur ce champ et d'y écrire « A remplir ».

Une fois les métadonnées récurrentes renseignées, cliquer sur **Sauvegarder et ajouter des conditions d'utilisation**.

Ne PAS modifier les conditions d'utilisation dans la plupart des cas : merci de vous référer aux consignes dans le guide du déposant sur ce point.

Ne pas oublier d'**Enregistrer** en fin de parcours.

Ajouter la partie sur les liens entre collections qui est actuellement dans le guide du déposant.

Gérer les restrictions ou des embargos sur des fichiers

Pour la procédure générale concernant la mise en place d'embargos ou de restrictions, merci de consulter le guide du déposant. Le présent chapitre en rappelle toutefois les principes et donne quelques conseils relatifs à la curation.

Dans certains cas (données à caractère personnel non anonymisées, données de santé, données tierces...), certaines données ne doivent ou ne peuvent être diffusées librement ou immédiatement.

DOREL permet de répondre à ces cas de figure, tout en permettant la conservation, la gestion de l'accès et la citation de ces jeux de données.

Il est possible de restreindre l'accès à des fichiers de manière très fine (fichier par fichier) :

- en précisant une date d'ouverture (**embargo**)
- en posant une **restriction** d'accès sans date de fin.

Ces restrictions valent pour l'ensemble des personnes n'ayant pas de droits sur le jeu de données, qu'elles soient identifiées dans DOREL ou non.

Il n'est possible de mettre des restrictions ou des embargos que tant que le jeu de données n'a pas été publié.

La restriction :

- n'a pas de limite dans le temps et doit être levée par le déposant ou le curateur ;
- permet d'autoriser la demande de communication, pour un utilisateur connecté à DOREL.

L'embargo :

- est limité dans le temps (date à préciser au moment du dépôt) ;
- au bout de la durée spécifiée, le jeu est automatiquement ouvert.

Bien qu'il n'existe pas de règle générale et universelle concernant la pose d'embargos ou de restrictions, on peut avancer deux principes :

- les restrictions doivent autant que possible concerner uniquement les exceptions visées par les lois et règlements en vigueur ;
- les embargos peuvent être posés pour des raisons bien plus diverses : attente d'une publication, données liées à un brevet...

Les raisons des uns et des autres nécessiteront dans certains cas un échange avec le déposant pour définir la solution à retenir.

Générer une URL privée

Tant que le jeu de données n'est pas publié, il est possible de générer une URL privée permettant de donner accès aux fichiers à un tiers, sans que ce dernier ait besoin de disposer d'un compte sur DOREL.

Cette fonctionnalité peut être utile dans le cadre des procédures de revue par les pairs, ou pour tout autre besoin de communication des fichiers avant publication.

Deux types d'URL privée, anonymisée et non anonymisée, sont proposées.

La fonctionnalité de création d'URL privée n'est accessible qu'aux curateurs. Cette obligation est malheureusement liée aux limitations du logiciel Dataverse.

Créer une URL privée non anonymisée

Dans le cas d'une URL privée non anonymisée, l'ensemble des éléments du jeu de données (métadonnées et fichiers) sont accessibles à toute personne disposant de l'URL.

Depuis la fiche, cliquer sur Modifier le jeu de données > URL privée, puis « Créer une adresse URL privée ».

The screenshot illustrates the workflow for creating a private URL. On the left, a sidebar menu shows 'URL privée' selected under 'Modifier le jeu de données'. This leads to a modal window titled 'URL privée du jeu de données non publié'. The modal contains instructions: 'Utiliser une adresse URL privée pour permettre à ceux qui n'ont pas de compte Dataverse d'accéder à votre jeu de données non publié. Pour plus d'informations sur la fonctionnalité d'URL privée, reportez-vous au [guide d'utilisation](#).' Below this, a message states 'L'adresse URL privée n'a pas été créée.' and a button 'Créer une adresse URL privée' is visible. A second modal window shows the successful creation of the URL, displaying the text 'This Private URL provides access to the anonymized dataset' and the URL: 'https://bac-dataverse.univ-lorraine.fr/privateurl.xhtml?token=b3df5804-8513-4e21-be33-e911f57d2dd0'. A 'Copy to Clipboard' button is present next to the URL.

Créer une URL privée anonymisée

Dans le cas d'une URL privée anonymisée, un certain nombre de métadonnées sont masquées, permettant notamment la revue par les pairs en double aveugle.

La procédure est la même que ci-dessus ; cliquer cependant sur le bouton de droite « Créer une URL pour un accès anonymisé » au moment de la génération de l'URL.

Les personnes destinataires ne verront que les métadonnées ne comportant pas d'information d'authentification. En voici ci-dessous un exemple :

Citation Metadata ^

| | |
|--------------------------------|---|
| Dataset Persistent ID ? | doi:10.80427/FK2/OHVQEI |
| Title ? | Social media use among teens |
| Author ? | withheld |
| Contact ? | withheld |
| Description ? | The Social Media Use Among Teens survey was conducted by the Youth Communication Development Project to understand social media communication behaviours among youth in Canada. The survey collected responses from Canadian youth using an online questionnaire that asks about social media use including, platform type, frequency of use, activity type, and location of use. This information is supplemented with the respondent's demographic and household characteristics. |
| Subject ? | Social Sciences |
| Keyword ? | social media teens communication |
| Notes ? | This survey was organized online. Mode of interview has been found to impact results, therefore it is not recommended that these results are compared with other survey results where the interview mode was telephone based |
| Language ? | English |
| Depositor ? | withheld |
| Deposit Date ? | 2022-03-31 |