

Guide de dépôt

Ce guide est mis à disposition des personnes qui déposent ou accompagnent au dépôt des données dans l'entrepôt NAKALA.

NAKALA est un entrepôt de données de recherche pour les sciences humaines et sociales. Son accès et son utilisation sont décrits dans la documentation d'Huma-Num : <https://documentation.huma-num.fr/nakala/>

La qualité et la richesse de la description des données sont des critères centraux des principes FAIR. Cela constitue un moyen d'atteindre les objectifs visés (faire en sorte que les données soient faciles à trouver, accessibles, interopérables et réutilisables).

La qualité se met en œuvre, par exemple :

- en utilisant des référentiels standardisés ;
- en respectant les mêmes normes intellectuelles de description pour un ensemble de données ;
- en choisissant des champs de métadonnées les plus adaptés à l'information donnée.

La richesse se met en œuvre en complétant le plus grand nombre possible de champs afin d'optimiser la compréhension des données. Dans NAKALA, la description est basée sur un ensemble minimal de cinq informations qui peuvent être enrichies de manière étendue et cumulative.

Ce guide propose un ensemble de conseils et bonnes pratiques pour les fichiers et les champs de métadonnées obligatoires et complémentaires de premier niveau.

Pour aller plus loin, un guide de description est disponible : <https://documentation.huma-num.fr/nakala-guide-de-description/>

Contact : nakala@huma-num.fr

Les fichiers

	Critères	Commentaires et propositions d'amélioration
Nommage	Il est recommandé d'établir un plan de nommage (préfixe commun pour tous les fichiers associés au projet) afin que les noms des fichiers soient cohérents et réguliers. Remarque : pour un dépôt multifichiers, NAKALA permet de trier les fichiers par ordre alphabétique, après chargement.	
Description	La description des fichiers peut être saisie : <ul style="list-style-type: none"> • pour des informations spécifiques à un fichier, dans le champ « Description » associé à chaque fichier. Ces informations sont indexées par NAKALA et sont affichées dans la visionneuse. • pour des informations globales sur l'organisation interne, le contenu, la définition des variables du ou des fichiers de la donnée, dans un document à part (type README, manifest IIIF, TEI-header, métadonnées métier, METS, etc.). 	
Formats	NAKALA accepte tous les formats, mais quand c'est possible, privilégier un format ouvert compatible avec FACILE, le service de validation des formats : https://facile.cines.fr/ Pour les fichiers au format texte, l'encodage de caractères en UTF-8 est recommandé, si applicable. La liste des formats supportés par l'API IMAGE IIIF dans NAKALA : « tif », « tiff », « jpg », « jpeg », « jp2 », « png », « pdf », « geojp2 ». Ces formats sont recommandés pour bénéficier des fonctionnalités du protocole IIIF.	

Les métadonnées

		Critères	Commentaires et propositions d'amélioration
Métadonnées principales du dépôt (obligatoires)	Type de dépôt	Vérifier que le type sélectionné est adapté à la donnée. Remarque : dans le cas d'une donnée multifichiers, le type sélectionné rend compte du type principal de la donnée.	
	Titre	Renseigner un titre explicite et signifiant (comme le serait un titre d'article). Remarque : si un titre non signifiant est nécessaire (identifiant ou nom de fichier), il peut être ajouté dans un second champ « Titre ».	
	Auteur	Vérifier que le nom des auteurs respecte la forme : Prénom Nom. Pour toutes autres formes d'auteurs, cocher « Anonyme » et les renseigner dans <i>dcterms:creator</i> ou <i>dcterms:contributor</i> (zone « Ajouter d'autres métadonnées »). Si l'auteur dispose d'un identifiant ORCID, renseigner cet identifiant. Remarque : si l'auteur est déjà enregistré et que son ORCID n'est pas renseigné, demander la mise à jour en écrivant à nakala@huma-num.fr .	
	Cas d'un auteur dit « anonyme »	Vérifier s'il est possible de renseigner au moins une personne ayant participé à la production, à la collecte ou au traitement des données dans <i>dcterms:creator</i> ou <i>dcterms:contributor</i> (zone « Ajouter d'autres métadonnées »). Remarque : les noms renseignés dans <i>dcterms:creator</i> s'affichent dans le bloc de citation.	
	Date de création	Renseigner la date de création de la source originale ou de la ressource numérique. Seuls les formats de date suivants sont possibles : YYYY-MM-AA (année-mois-jour), YYYY-MM (année-mois), YYYY (année). Pour toutes autres formes de dates, cocher « Inconnue » et les renseigner dans <i>dcterms:created</i> ou <i>dcterms:date</i> (zone « Ajouter d'autres métadonnées »).	
	Licence	La licence associée à la donnée est de préférence libre (Creative Commons, Etalab, GNU, etc.) et correspond aux droits de réutilisation possibles. Remarque : vérifier ce qui justifie le recours à une licence non libre.	
Informations complémentaires	Description	Décrire la donnée du point de vue scientifique : contexte de production et/ou de collecte ; méthodologie ; objectifs, etc. Décrire le programme de recherche dans un autre champ « Description ».	
	Résumé	Saisir le résumé du projet dans <i>dcterms:abstract</i> (zone « Ajouter d'autres métadonnées »).	
	Table des matières	Saisir une liste ou une table des matières dans <i>dcterms:tableOfContents</i> (zone « Ajouter d'autres métadonnées »).	
	Mots-clés	Renseigner au moins 2 à 3 mots-clés précisant le contenu intellectuel et le sujet de la donnée en français est fortement recommandé. Le mot-clé doit être sous forme d'un mot, d'un groupe de mots ou d'une URI qui renvoie vers un référentiel de mots-clés.	
	Ajout de mots-clés dans une langue secondaire	Si des mots-clés sont renseignés dans une autre langue, les renseigner dans un autre champ « Mots-clés ».	
	Langues	Pour des fichiers vidéo, son et texte, renseigner la ou les langue-s du contenu des fichiers.	
Ajouter d'autres métadonnées	Métadonnées géospatiales <i>dcterms:spatial</i>	Renseigner la couverture géographique du contenu intellectuel des données dans <i>dcterms:spatial</i> si pertinent. Renseigner le positionnement hiérarchique de la localité est recommandé : « France, Auvergne-Rhône-Alpes, Rioupéroux ». Dupliquer le <i>dcterms:spatial</i> pour renseigner certains types d'informations spatiales normalisées : <ul style="list-style-type: none"> pour un concept issu du Thesaurus of Geographic Names (Getty Thesaurus of Geographic Names - TGN) : saisir l'identifiant du concept issu du TGN et sélectionner dans « Type » : <i>dcterms:TGN</i>. pour une zone géographique : saisir les informations de la zone géographique et sélectionner dans « Type » : <i>dcterms:Box</i>. Ex. Pour la zone géographique qui correspond à l'Australie occidentale, l'information doit respecter la forme suivante : <code>name=Western Australia; northlimit=-13.5; southlimit=-35.5; westlimit=112.5; eastlimit=129</code>. pour des coordonnées géographiques : saisir les coordonnées géographiques et sélectionner dans « Type » : <i>dcterms:Point</i>. Les coordonnées géographiques doivent être saisies dans la forme suivante : <code>name=Perth, W.A.; east=115.85717; north=-31.95301</code>. pour un code pays issu de la norme ISO3166 : saisir le code pays issu de ISO3166 (code Alpha-2) et sélectionner dans « Type » : <i>dcterms:ISO3166</i>. Pour des précisions sur les formes à respecter lors de l'utilisation des « Types », se référer au guide de description : https://documentation.huma-num.fr/nakala-guide-de-description/	
	Métadonnées temporelles <i>dcterms:temporal</i>	Renseigner la couverture temporelle du contenu intellectuel des données dans <i>dcterms:temporal</i> . Si l'information est sous la forme d'une période (temporelle ou historique) avec un intervalle temporel délimité, sélectionner dans « Type » : <i>dcterms:Period</i> . L'information doit respecter la forme suivante : « name=Haut empire Romain; start=-0030; end=0099 » ; saisir quatre chiffres est obligatoire pour renseigner les années ; l'année « start » doit être antérieure à l'année « end ». Pour des précisions sur les formes à respecter lors de l'utilisation des « Types », se référer au guide de description : https://documentation.huma-num.fr/nakala-guide-de-description/	
	Lien vers la publication <i>dcterms:bibliographicCitation</i> et/ou <i>dcterms:isReferencedBy</i>	Si une publication est associée aux données, il est fortement recommandé de l'indiquer. Renseigner la référence bibliographique complète dans <i>dcterms:bibliographicCitation</i> (vérifier que l'identifiant pérenne de la publication est renseigné ou l'ajouter, si applicable). Saisir l'identifiant pérenne de la publication dans <i>dcterms:isReferencedBy</i> , si applicable. Aucune information de langue ne doit être saisie.	
	Droits <i>dcterms:rights</i>	Si un embargo sur un ou des fichiers a été apposé, il est recommandé d'en préciser la raison dans un <i>dcterms:rights</i> .	
	Éditeur <i>dcterms:publisher</i>	Saisir le nom du responsable de la mise à disposition de la ressource dans <i>dcterms:publisher</i> (le nom du laboratoire dans sa forme actuelle, par exemple).	
	Relation vers des ressources externes : données hors NAKALA, base de données, etc. <i>dcterms:relation</i>	Si des données déposées dans d'autres entrepôts (projet interdisciplinaire) ou dans des ressources sur le web (base de données, sites, bibliothèques numériques, etc.) sont en lien avec la donnée, il est recommandé de les indiquer dans <i>dcterms:relation</i> . Saisir l'identifiant pérenne de la ressource externe à mettre en relation. Lorsqu'une URI est renseignée, sélectionner dans le champ « Type » <i>dcterms:URI</i> . Aucune information de langue ne doit être saisie.	

Remarques

Il est fortement recommandé de préciser la langue de chaque métadonnée renseignée sous forme textuelle.

Si vous avez uniquement renseigné une URI dans une métadonnée, sélectionner dans le champ « Type » *dcterms:URI* et ne surtout pas renseigner d'information de langue.