

FORMATS DE FICHIERS RECOMMANDÉS PAR SODHA

SODHA n'impose pas la conversion des fichiers soumis par les déposants vers des formats ouverts / libres pour deux raisons : le but principal de SODHA est de permettre la diffusion et la réutilisation de données de recherche, et la [Convention de dépôt SODHA](#) n'occasionne pas un transfert de propriété des données déposées aux Archives de l'État.

Cela étant dit, l'interopérabilité (qui constitue l'un des principes FAIR) est essentielle pour permettre l'accès à et la réutilisation de données de recherche. C'est pourquoi nous encourageons les déposants à privilégier les formats recensés sur la 2^e page de ce document pour leurs fichiers de données. Il s'agit de formats **ouverts**, qui garantissent également la pérennité de vos données (puisque'ils ne sont pas couverts par le secret industriel).

Cependant, si vous avez des raisons de croire que certains aspects de vos données risquent d'être compromis voire corrompus à l'occasion d'un changement de formats, nous vous recommandons alors de déposer vos fichiers à la fois dans leur format (éventuellement propriétaire) original ainsi que dans un format libre, comme ceux recensés à la page suivante.

N.B. : Vous n'avez pas besoin de convertir vos **fichiers de données tabulaires** dans des formats libres. En effet, le logiciel Dataverse sur lequel s'appuie SODHA crée automatiquement une version .tab des fichiers de données tabulaires qu'il reçoit dans d'autres formats (y compris le .sav, le format de SPSS, ou encore .sas et d'autres).

↓ Veuillez vous rendre à la page 2 pour voir la liste des formats ↓

Types d'objets numériques	Formats recommandés
Classeurs / Données tabulaires	.csv – .tsv/.tab – .ods – .odf
Fichiers textes	.pdf (PDF/A-2a de préférence ¹) – .odt – .txt – idéalement avec un encodage en UTF-8 ²
Présentations	.odp – .odt
Bases de données	.xml – .csv – idéalement avec une docu- mentation au format .xml ou SQL concer- nant les relations entre les tables, les triggers, etc.
Fichiers images	.tiff (TIFF 0.6 non compressés) – .jpg (JPEG ou JPEG 2000) – .png
Fichiers sons	.mp3 – .flac – .ogg – .oga
Fichiers vidéo	.dpx
Cartes et plans	.dwg – .svg
Données géographiques	.gml (/ .xml/.kml) – GeoTIFF (.tiff)
Pages et sites Web	XHTML 1.0 – .css – .js – .jsp – .php

¹ Le format PDF/A3 peut contenir des objets non normalisés, c'est pourquoi la lisibilité à long terme de tels fichiers n'est pas toujours garantie.

² L'encodage en UTF-8 garantit que la plupart des caractères spéciaux (ex. : é, ç, ü...) susceptibles d'apparaître dans vos données ne seront pas corrompus au moment de la sauvegarde des données (c'est ce qui se produit quand, par exemple, le mot *façade* est reproduit sous la forme « faade »).

Historique de version de ce document

Date	Numéro de version	Changements effectués
30-07-2021	1.0	Première version