

FORMATS DE FICHIERS RECOMMANDÉS POUR L'ARCHIVAGE PÉRENNE ¹

	KOST-CECO ²	ETHZ	Library of Congress	University of Cambridge	UNIL	EPFL/DLCM (formats ouverts)
Source	CECO 2021a	ETHZ, [s. d.]a	LIBRARY OF CONGRESS 2021³	UNIVERSITY OF CAMBRIDGE [s. d.]	UNIL 2019	EPFL 2019
texte	PDF/A-2u	PDF/A 2b, plain text PDF/A 2u, xml	xml, PDF/UA, PDF/A, <i>Attention à l'encodage, par exemple : UTF-8, UTF-16, US-ASCII</i>	xml, txt, html, PDF/A	XML, PDF/A, HTML, ASCII, UTF-8	txt, PDF, ODT, ODM, tex, MD, HTM, XML, EXtXT2, ODF
image	Si nativement analogique : <ul style="list-style-type: none"> PDF/A-2 avec compression modérée JPEG (si originalement jpeg) PNG que pour les contenus web importants TIFF et éventuellement JPEG2000 (même si peu répandu) DNG si accès aux données raw 	<ul style="list-style-type: none"> TIFF (pérenne) (eciRGB_v2 AdobeRGB(1998) + JPEG (diffusion)) JPEG2000⁴ + JPEG (diffusion) Autres options éventuelles : <ul style="list-style-type: none"> TIFF (compressé) PNG (non compressé) JPEG2000 (.jp2) loseless compression + Digital-Negative (en complément d'un TIFF) SVG (without javascript binding) 	Photo : TIFF, JPEG2000, PNG, JPEG/JFIF, BMP Autre image : idem + SVG	TIFF, PNG, jpeg ⁵	TIFF, JPEG 2000, PDF, PNG, GIF, BMP	TIFF, PNG, JVG, JPEG, FITS

¹ Dans le tableau ci-après, vous trouverez les différents formats de fichiers recommandés pour l'archivage par les institutions étudiées en décembre 2022 : [École polytechnique fédérale de Lausanne](#) (EPFL), [Eidgenössische Technische Hochschule Zürich](#) (ETHZ), [Université de Genève](#) (UNIGE), [Université de Lausanne](#) (UNIL), [University of Cambridge](#), [Library of Congress](#), [Digital Curation Centre](#) (DCC), [Centre de coordination pour l'archivage à long terme de documents électroniques](#) (CECO).

² Centre de coordination pour l'archivage à long terme de documents électroniques

³ La *Library of Congress* propose dans ce même document également des recommandations pour les supports analogiques.

⁴ Format peu répandu

⁵ Jpeg : Réduit la qualité de l'image

	KOST-CECO ²	ETHZ	Library of Congress	University of Cambridge	UNIL	EPFL/DLCM (formats ouverts)
audio	<ul style="list-style-type: none"> wav (wave pcm) non comprimé flac (avec compression sans perte) 	WAV (uncompressed pulse-code modulated)	format final, PCM WAVE, WAVE	flac, wav, mp3	WAVE, AIFF, MP3, MXF	FLAC, wav, ogg, mxl, midi, moi, hundrum
vidéo	<p>Attention :</p> <ul style="list-style-type: none"> Le moins de conversion possible (laisser au maximum le format d'origine lors de la première étape de préservation) Le moins de migration possible : privilégier le format de départ (même si c'est "temporaire") Stockage à disposition → compression avec perte ou sans perte ? Evaluation archivistique : le format dépend des finalités 	FFV1 codec (version 3 ou +) in matroska container (.mkv)	<ul style="list-style-type: none"> IMF (avec MXF, composition playlist, XML) .mov + 4444 (XQ) ou 422HQ codec MPEG 2 (ok FFV1 in mkv : if no closed caption or no timecode information) 	Pas d'informations	MOV, MPEG, AVI, MXF	mp4, mj2, avi, mkv
feuille de calcul	Le format d'origine accompagné (ods, xlsx, xls) + un format plus pérenne	CSV	CSV, TSV	csv	CSV	csv, tab, ods, sgl, tsv, xml + dtd, xlsx
Base de données	SIARD (csv peut être une fausse bonne idée car il n'y a plus de liens entre les informations : tout dépend des finalités)		<ul style="list-style-type: none"> format using well known schemas with public validation tool line oriented (CSV, TSV...) Plateforme indépendante qui utilise du open format (db ou db3) xls/xlsx 	xml, csv	XML, CSV	csv, hdf5, txt, html, tex, fastQ, por
hypertexte	WARC (mais migration difficile), PDF/A		WARC	<i>Renvoi à ukdata service</i>	WARC	Pas d'information

	KOST-CECO ²	ETHZ	Library of Congress	University of Cambridge	UNIL	EPFL/DLCM (formats ouverts)
SIG	Détails disponibles directement dans la source (cela dépend des contextes).		Design 3D : idem image + GIF SVG, .dxf, shapefile (vecteur)		SHP, DBF, GeoTIFF, netCDF (<i>détails disponibles dans la source</i>)	netcdf, tabula 415, attribute data, she, shx, dbf, faz, sbx, sen, post615, TIFF, tfw, GeoJson
CAO/FAO	IFC	AutoCAD drawing (dwg), dxf, exteasisle 3D, x3d	Scamed 3D object STL, RTI, polygon file, Wavefront OBJ			x3d, x3dv, x3d3, pdf3d, pov, pdbml
Autres			Partitions : XML (music XML, MEI), PDF/A, PDF/UA		Container : tar, gzip, zip Statistiques : ASCII, DTA, POR, SAS, SAV	xml, json, rdf
Programmes codés						m, r, py, ypn8, rstudio, rmd, netCDF, aiml

Bibliographie – Formats

CECO (Centre de coordination pour l'archivage à long terme de documents électroniques), 2021. Catalogue des formats de fichiers pour l'archivage. *KOST-CECO* [en ligne]. Version 6.2. [Consulté le 10 février 2022]. Disponible à l'adresse : <https://kost-ceco.ch/cms/formats-de-fichiers.html>

DIGITAL CURATION CENTRE, 2007. DCC- Digital Curation Manual : instalment on "Files Formats". Version 1.0. *Digital Curation Centre* [en ligne]. Octobre 2007. [Consulté le 3 février 2022]. Disponible à l'adresse : <https://dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/file-formats/file-formats.pdf>

EPFL (École polytechnique fédérale de Lausanne), 2019. EPFL Library Research Data Management Fastguides. *Zenodo* [en ligne]. 11 février 2019. [Consulté le 2 février 2022]. Disponible à l'adresse : <https://doi.org/10.5281/zenodo.3327830>

ETHZ (Eidgenössische Technische Hochschule Zürich), [sans date]a. File formats for archiving. *Eidgenössische Technische Hochschule Zürich* [en ligne]. [Consulté le 2 février 2022]. Disponible à l'adresse : <https://documentation.library.ethz.ch/display/DD/File+formats+for+archiving>

ETHZ (Eidgenössische Technische Hochschule Zürich), [sans date]b. Converting digitized line graphs into JPEG2000 or TIFF. *Eidgenössische Technische Hochschule Zürich* [en ligne]. [Consulté le 2 février 2022]. Disponible à l'adresse : <https://documentation.library.ethz.ch/display/DD/Converting+digitized+line+graphs+into+JPEG2000+or+TIFF>

ETHZ (Eidgenössische Technische Hochschule Zürich), [sans date]c. Converting image files into archival formats. *Eidgenössische Technische Hochschule Zürich* [en ligne]. [Consulté le 2 février 2022]. Disponible à l'adresse : <https://documentation.library.ethz.ch/display/DD/Converting+image+files+into+archival+formats>

LIBRARY OF CONGRESS, 2021. Recommended Formats Statement 2021-2022. *Library of Congress* [en ligne]. [Consulté le 10 février 2022]. Disponible aux adresses : <https://www.loc.gov/preservation/resources/rfs/RFS%202021-2022.pdf>(PDF)
<https://www.loc.gov/preservation/resources/rfs/TOC.html> (site Web)

UK Data Service, [sans date]. Recommended formats. *UK Data* [en ligne]. Consulté le 28 avril 2022]. Disponible à l'adresse : <https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/>

UNIGE (Université de Genève), [sans date]. Formats fichier. *Université de Genève* [en ligne]. [Consulté le 2 février 2022]. Disponible à l'adresse : <https://www.unige.ch/researchdata/fr/perserver/formats-fichier/>

UNIL (Université de Lausanne), [sans date]. Archivage & partage. Université de Lausanne [en ligne]. [Consulté le 2 février 2022]. Disponible à l'adresse : <https://www.unil.ch/openscience/fr/home/menuinst/open-research-data/gerer-les-donnees-de-recherche-research-data-management/archivage--partage.html>

UNIVERSITY OF CAMBRIDGE, [sans date]. Choosing Formats. *University of Cambridge* [en ligne]. [Consulté le 3 février 2022]. Disponible à l'adresse : <https://www.data.cam.ac.uk/data-management-guide/creating-your-data/choosing-formats>