

BESTANDSFORMATEN AANBEVOLEN DOOR SOHDA

Aangezien het hoofddoel van SODHA is om te zorgen voor de verspreiding en het hergebruik van onderzoeksgegevens en omdat het [SODHA-Bewaaringsovereenkomst](#) het eigenaarschap van de opgeladen gegevens niet overdraagt aan het Rijksarchief verplicht SODHA niet om de opgeladen gegevensbestanden te converteren naar open / niet-gepatenteerde bestandsformaten.

Anderzijds is interoperabiliteit – een van de “FAIR”-principes (vindbaarheid, toegankelijkheid, interoperabiliteit en herbruikbaarheid) – de sleutel voor toegang tot en hergebruik van onderzoeksgegevens en daarom moedigen we u aan volgende bestandsformaten te overwegen. We nodigen de deponenten uit hun gegevens op te laden in **open formaten** zoals deze die vermeld staan op de tweede bladzijde van dit document. Open formaten garanderen ook dat uw gegevens op lange termijn beschikbaar blijven aangezien ze niet afhankelijk zijn van bedrijven die failliet kunnen gaan (en dat daarbij technologische geheimen over gepatenteerde bestandsformaten mee verdwijnen).

Maar indien u weet of vermoedt dat bepaalde elementen van uw werk zouden kunnen verloren geraken of beschadigd worden indien u ze bewaart in een open / niet-gepatenteerd formaat raden we u aan zowel het originele bestand op te laden als een kopie in een van de formaten volgens de standaarden die hieronder vermeld staan.

Opgelet: Dit hoeft u niet te doen voor **gegevensbestanden in tabelvorm**. Gegevens in tabelvorm kunnen in de SODHA Dataverse worden opgeladen in nagenoeg alle bekende tabelbestandsformaten (met inbegrip van SPSS .sav, .sas...) gezien Dataverse automatisch een .tab versie aanmaakt van de gegevens die werden verwerkt en de gebruikers daarna toelaat om het bestand te downloaden in dit formaat of in het originele formaat.

Tenslotte raden we aan om geen gecomprimeerde bestanden op te laden aangezien ze erg beschadigd kunnen geraken: de beschadiging van slechts 1 bit binnen het bestand kan leiden tot de onleesbaarheid van grote (of zelfs alle) delen van de gegevens.

↓ Ga hieronder naar bladzijde 2 om de lijst te raadplegen ↓

Types digitale objecten	Aanbevolen formaten
Spreadsheets / Gegevens in tabelvorm	.csv – .tsv/.tab – .ods – .odf
Teksten	.pdf (PDF/A-2a idealiter ¹) – .odt – .txt – idealiter in UTF-8 ²
Presentaties	.odp – .odt
Databases	.xml – .csv – idealiter enkel voor documentatie in .xml of SQL over de relatie tussen tabellen, triggers, enz.
Beelden	.tiff (niet gecomprimeerd TIFF 0.6) – .jpg (JPEG or JPEG 2000) – .png
Geluid	.mp3 – .flac – .ogg – .oga
Video's	.dpx
Kaarten, technische tekeningen	.dwg – .svg
Geografische gegevens	.gml (/ .xml/ .kml) – GeoTIFF (.tiff)
Webpagina's	XHTML 1.0 – .css – .js – .jsp – .php

¹ PDF/A3 kan niet-genormaliseerde objecten bevatten, en de lange termijn leesbaarheid van bestanden in dit formaat kan dus niet altijd worden gegarandeerd.

² UTF-8 garandeert dat de meeste, zo niet alle, speciale karakters (bijvoorbeeld é, ç, ü...) die eventueel in uw gegevens voorkomen niet beschadigd worden tijdens het opslaan (wat bijvoorbeeld gebeurt wanneer het woord *façade* in een tekst verschijnt als "fa◊ade").

Historiek van dit document

Datum	Versie-nummer	Aangebrachte wijzigingen
30-07-2021	1.0	Eerste versie
05-08-2021	1.1	Correctie van typefouten