

SODHA-BELEID VOOR HET PUBLICEREN VAN DATASETS

Dit beleidsdocument bepaalt onder welke voorwaarden datasets die werden ingediend voor review effectief zullen gepubliceerd worden door SODHA.

Hoofdbedoeling is om het publicatiebeleid van SODHA toe te lichten.

1. Overzicht van de publicatiecriteria

Datasets ingediend voor review moeten beantwoorden aan volgende criteria om te worden gepubliceerd:

- Indien toegang tot een of meerdere bestanden beperkt wordt heeft de deponent de specifieke toegangsvoorwaarden gespecificeerd in het veld “Voorwaarden voor toegang”. ([Bewaar-gevingsovereenkomst](#), art. 10; [Toegang en Hergebruik Beleid](#), b.)
 - Het onderzoek dat leidde tot de dataset is relevant voor de sociale wetenschappen of de digitale humanewetenschappen. (Beleid voor het publiceren van datasets, 2.1)
 - De gegevens bevatten geen gevoelige informatie (bijvoorbeeld informatie waarmee personen kunnen geïdentificeerd worden). (Beleid voor het publiceren van datasets, 2.2)
 - De metadata bevatten geen dubbelzinnige informatie. (Beleid voor het publiceren van datasets, 2.3)
 - De bestanden van de dataset kunnen succesvol worden gedownload en geopend.
 - Wanneer een nieuwe versie van een reeds gepubliceerde dataset wordt gemaakt, moeten wijzigingen worden gedocumenteerd in een extra exemplaar van het veld “Beschrijving”.
 - De dataset bevat geen uitvoerbare / binaire files.
-

2. Criteria voor de publicatie van de datasets

SODHA volgt **fundamentele** (inhoudelijke) en **formele** criteria om te bepalen of ingediende datasets geschikt zijn voor publicatie. Ook het aspect “**gevoelige informatie**” wordt in overweging genomen.

In essentie moeten gegevenscollecties voortkomen uit wetenschappelijke onderzoeksdomeinen die relevant zijn voor de objectieven en opdrachten van SODHA.

Formeel moeten de metadata die de datasets beschrijven beantwoorden aan een aantal basisregels.

Inzake veiligheid mogen datasets geen gevoelige informatie bevatten, zoals persoonlijke gegevens.

2.1 Fundamenteel niveau (inhoud)

SODHA is een **gegevensarchief** voor de **sociale wetenschappen** en de **digitale humanwetenschappen**. Het platform is bedoeld voor gegevens die relevant zijn voor een of meer vakgebieden in deze twee wetenschappen.

Om te bepalen wat wordt verstaan onder sociale wetenschappen en digitale humanwetenschappen verwijst SODHA naar gezaghebbende classificaties zoals de BELSPO-databank voor projecten [FEDRA](#), ECOOM-Hasselt [Flemish Research Discipline Standard](#), en het [verslag over digitale humanwetenschappen](#) van het Massachusetts Institute of Technology (MIT).

Ondanks dit thematisch criterium volgt SODHA een “neutrale” en “agnostische” visie voor de inhoud van de datasets: de rol van ons gegevensarchief is niet het valoriseren van de kwaliteit van het onderzoek van de deponenten, noch van de omvang van de datasets, noch van de onderwerpen die werden uitgekozen door de onderzoekers.

In het geval een dataset niet geschikt wordt bevonden voor opladen in SODHA, conform de normen die hierboven werden vermeld, kan SODHA zorgen voor begeleiding voor een andere publicatie of voor andere opslagruimtes (voor zover onze expertise dat toelaat).

2.2 Gevoelige informatie

Op dit ogenblik kan SODHA helaas geen **gegevens** aanvaarden die gevoelige informatie bevatten, met name (maar niet alleen) persoonsgegevens. De infrastructuur wordt aangepast om dit in de toekomst mogelijk te maken.

Deponenten mogen evenwel op het SODHA Dataverse platform **metadata** ingeven over gevoelige gegevens. Met andere woorden, indien ze onderzoeksgegevens niet kunnen publiceren omdat ze te gevoelig zijn, worden ze toch aangespoord om het bestaan van dergelijke gegevens te melden, aangezien dit voor de betrokken onderzoekers waardevolle informatie is. [Raadpleeg onze FAQ voor meer informatie.](#)

2.3 Formeel (metadata)

Wanneer deponenten een dataset willen opladen in SODHA wordt hen gevraagd een aantal verplichte metadata-velden in te vullen.

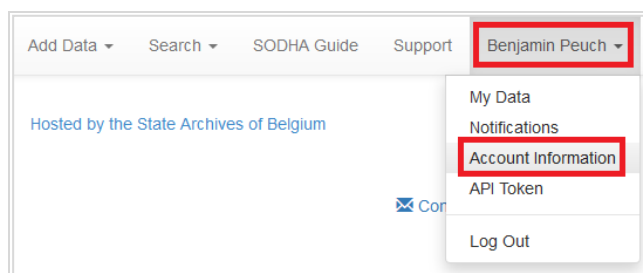
Door de formele samenhang te bewaren van de metadata die op SODHA gepubliceerd worden, vergroot de kans dat onderzoekers hun weg vinden bij het terugvinden van de gegevens waar ze naar op zoek zijn tijdens het navigeren of het ingeven van zoektermen op de SODHA webapplicatie. Daarom zijn volgende regels van toepassing:

- a. **Namen** van personen die worden vermeld in de velden Auteur, Contact en Producent moeten volledig worden uitgeschreven (vooral de voornaam, ondanks initialen in te midden), gescheiden door een komma en in deze volgorde: **[familienaam, voornaam, initialen in te midden]**

Voorbeelden	Pirlot, Mélanie Van den Broeck, Jan-Albert Márquez Gutiérrez, Emelina Soraya de la Vega, Christian Von Humboldt Zimmerman, Dietrich du Val, Anne-Sophie Angélique Smith, Christian A. Pham, Van Duc Zhang, Ziyi
-------------	---

Er kunnen uitzonderingen worden gemaakt, bijvoorbeeld voor oudere datasets waar de voor-
namen van de auteurs niet met zekerheid kunnen worden geïdentificeerd. Gelieve in dergelijke
gevallen contact op te nemen met de SODHA-beheerders via sodha@arch.be.

Aangezien de naam van de deponent automatisch wordt gedocumenteerd in het veld Depo-
nent op basis van de informatie in het account van de deponent wordt hen gevraagd hun voor-
naam voluit te schrijven, zowel wanneer ze zich registreren bij SODHA als later bij het wijzigen
van hun persoonlijke informatie (via het menu Accountinformatie):



- b. Om **regelafbrekingen** weer te geven in de metadatavelden moeten de deponenten gebruik
maken van de HTML-tag **
** en niet alleen de Entertoets. De te volgen methode is deze:

Om dit te bekomen:

Deze dataset werd gemaakt tussen 2012 en 2014.
De dataset werden ingezameld tussen 3 juli 2012 en 6 januari 2013.

moeten **
**-tags als volgt worden toegevoegd:

Deze dataset werd gemaakt tussen 2012 en 2014.**

**
De dataset werden ingezameld tussen 3 juli 2012 en 6 januari 2013

- c. **Paragrafen** worden gescheiden met een witregel, zoals in het voorbeeld hierboven.
- d. **Titels** van **gecteerde werken**, bijvoorbeeld in het veld Gerelateerde publicatie, kunnen worden
weergegeven in italic door gebruik te maken van de tags **<i></i>**. Dit is de te volgen methode:

Om het volgende te bekomen:

D'hont, Ekaterina, *Emotions: Psychology and the Social Sciences*, Brussels, Psykon Press, 2016.

moeten `<i></i>` tags als volgt worden toegevoegd:

D'hont, Ekaterina, `<i>Emotions: Psychology and the Social Sciences</i>`, Brussels, Psykon Press, 2016.

- e. Wanneer **meer dan één element** moet worden opgenomen in een **herhaalbaar veld** (zie de knop “+”), moet het veld zoveel keer worden herhaald als nodig, in plaats van alles op te slaan in een onderdeel van het veld:

Date of Collection ⓘ	Start ⓘ 2018-03-21	End ⓘ 2018-04-11	+	-
	Start ⓘ 2018-10-05	End ⓘ 2018-11-16	+	-

Door de velden te herhalen ontstaat meer indexering en worden uw gegevens meer zichtbaar.

- f. Omdat SODHA open **linked data** wil bevorderen, zal, in het geval de beheerders van het SODHA-platform bijkomende informatie vinden zoals de DOI van een geciteerd werk of de ORCID-code van een auteur, enz., gesuggereerd worden om die elementen toe te voegen aan de metadata van de deponent. Op die manier worden uw gegevens beter zichtbaar.

2.4. Versie management

Als een dataset al een keer is gepubliceerd en deponenten maken er een nieuwe versie van, dan moeten ze de wijzigingen die ze in de eerdere versie hebben aangebracht, documenteren in een nieuw exemplaar van het veld “Beschrijving”. De originele(n) versie(s) van dit veld dienen behouden te blijven. Door op de knop “+” te klikken wordt een nieuw veld gegenereerd:

Description * ⓘ	This field supports only certain HTML tags .		+
	Text * ⓘ This dataset contains the files related to a survey investigating the needs and practices of Belgian social science researchers in terms of research data management. The survey took place within the framework of the Social Sciences Data Archive (SODA) project, which aims to set up a data archive for social sciences that will also act as a CESSDA (Consortium of European Social Science Data Archives) service provider in Belgium.		
	Date of Description ⓘ 2019-06-01		

2.5. Uitvoerbare / binaire files

Omwille van veiligheidsredenen kan SODHA geen uitvoerbare bestanden accepteren.

2.6. Slotoverwegingen

Wanneer een deponent een dataset indient voor review zal de beheerder van SODHA de metadata eerst proeflezen alvorens ze te publiceren.

Indien de metadata formeel afwijken van de regels hierboven nemen de beheerders van SODHA de vrijheid om **minimale** correcties aan te brengen. **Afgezien van deze regels zullen de beheerders van SODHA de input van de deponenten geenszins wijzigen, en de deponenten worden altijd ingelicht over elke wijziging die de SODHA-beheerders aanbrengen aan hun metadata.**

Indien anderzijds sommige elementen in de metadata dubbelzinnig blijken (bijvoorbeeld door typfouten) zullen de SODHA-beheerders de deponenten contacteren om hen te vragen duidelijkheid te verschaffen vooraleer er gepubliceerd wordt.

3. Historiek van dit document

Datum	Versie-nummer	Aangebrachte wijzigingen
30.07.2021	1.0	Eerste versie
05.08.2021	2.0	<ul style="list-style-type: none">• De subsecties 2.4 “Versie management” en 2.5 “Uitvoerbare / binaire files” werden toegevoegd• Correctie van typfouten