

SODHA END-USER LICENSE AGREEMENT

1. Agreement overview

1.1 Introduction

This is an end-user license agreement ("EULA") between users of the platform or infrastructure of the **Social Sciences and Digital Humanities Archive** (hereinafter referred to as "users" or "the user") and the Social Sciences and Digital Humanities Archive ("SODHA," "the data archive"), a department of the State Archives of Belgium acting in the name and on behalf of the depositor. This document identifies the expected level of service provided to each and any user of the SODHA platform.

As a public, free-of-charge service offered by a federal institution, SODHA adheres to the ideals of long-term preservation heralded by its host organization for as long as its line ministry, the Belgian Science Policy Office (BELSPO), provides financial support for the continuation of operations.

1.2 Definitions

The **SODHA infrastructure** or **SODHA platform** relies on the free, open-source DATAVERSE software application for repository management and publication of research data for reuse developed by the Institute for Quantitative Social Science (IQSS) of Harvard University. SODHA encapsulates two Dataverse installations, a production instance located at https://www.sodha.be and a test instance located at https://test.sodha.be.

A **user** of SODHA is any natural person who has or has not been registered via institutional login or any other login and who accesses and makes use of one of the aforementioned websites or who contacts SODHA by any means.

A **digital object identifier (DOI)** is a permanent and unique identifier (PUID) delivered by a specific agency to digital objects such as traditional scientific publications (articles, books, book chapters...) as well as to datasets.

A **backup** is a copy of a (collection of) digital object(s) made to ensure that, in case the original file(s) is (are) lost, it (they) can be restored thanks to said copy.

A **data archive** is a repository which offers records management and long-term archiving services specifically for digital research data. Unlike archives in the traditional sense of the word, data archives curate digital objects that were recently produced and who qualify as 'dynamic' or 'semi-static' records, i.e. they are still likely to be used either by their creators or other users and it is too early to document and rework them (e.g. conversion to open source file formats) for long-term archiving processing, as with traditional archive documents emanating from government agencies.

2. Service agreement

2.1 Types and levels of services

The key purpose of SODHA is to receive collections of research data (datasets) and make them available for reuse. To this end, our essential services are:



1) Providing users with a deposit interface

SODHA commits to upholding an online platform on which logged in users can document datasets (by the input of metadata) and upload data files.

The deposit interface of the open source software SODHA relies upon, i.e. Dataverse, is based on the international metadata standard for social science data, Data Documentation Initiative (DDI). In other words, the metadata fields which depositors must/may fill out (depending on whether a field is mandatory or not) are directly drawn from an authoritative document type definition (DTD) which was put together by experts in social sciences who determined that these information items were essential for properly documenting datasets.

2) Providing users with a data catalogue

The Dataverse software presents users who access its homepage with a list of published datasets right from the go.

By default, users are shown all published datasets starting with the most recent publications. They can browse the standard list of results, or reverse the display order chronologically speaking, or they can display the list of datasets in alphabetical or anti-alphabetical (A-Z or Z-A) order. An advanced search feature as well as search facets (narrow results with e.g. a specific author, keyword, author affiliation...) are also available.

3) Preserving research data

SODHA makes a point of providing users with a safe, dependable data storage solution. As a service of the State Archives of Belgium, SODHA enjoys the expertise in digital preservation of its host institution.

The backup policy of SODHA is to make a full copy of all data deposited in the repository at the beginning of the week and then incremental backups every other workday until the end of the week. In their 'live' state, data deposited on the SODHA platform are stored in one server locale while the backup copies are stored in another. However, both locales are located on the same site, without the 200-kilometer distance recommended for an optimal, maximum security backup procedure. The State Archives is currently working to integrate offsite storage in its infrastructure.

Eventually, data collections published on the SODHA platform for which the State Archives becomes responsible at some point — typically data emanating from federal organizations, which fall within the remit of the law on archives of 24 June 1955 [DE] [FR] [NL] — will be further documented and potentially integrated to a larger archive collection by an archivist who will write a dedicated finding aid. Once this is done and the data have been transferred to the State Archives' management system, they will (alongside their metadata) be copied on LTO tapes and sent off to the Long-Term Preservation platform (LTP), an infrastructure financed by BELSPO. The LTP platform is a state-of-the-art archiving solution which brings all guarantees of long-term preservation.

However, data that are not owned by the State Archives — typically data deposited by university researchers — cannot be transferred to the LTP platform because, as per the SODHA Deposit Agreement, they remain the property of their original owners. Ownership of the data can be transferred to the State Archives by means of an *ad hoc* or a framework deposit agreement, typically signed between the State Archives and a specific research unit within a university.

Scolha

Consequently, while SODHA endeavors to become a trusted repository, we do not recommend that users keep their data *only* on SODHA because our means are still limited. SODHA adheres to the LOCKSS principle ("Lots Of Copies Keep Stuff Safe") and thereby advises users to consider SODHA as a secondary backup/recovery option. These contraindications are even more relevant in case users upload data with restricted access.

4) Providing advice to users

To the extent of the expertise within the data archive and, more broadly, the State Archives of Belgium, SODHA will gladly answer user queries regarding data archiving, general data management practices, file formats, data management plans, or data documentation, among other subjects.

However, because SODHA is still a young data archive, we cannot yet provide *ad hoc* guidance for specific data management issues, e.g. problems with data collection, social science software, research methodologies, ethical dilemmas, or data anonymization. As a general principle, if we receive requests for help for which we do not feel qualified, we will endeavor to redirect users to other services or organizations that are likely to be of assistance.

The SODHA Guide comprises a section for External resources with a list of tools, services, and organizations that can be of relevance to researchers in social sciences and digital humanities. If you know about interesting tools, services, or organizations that could be added to it, please contact us and we will gladly expand the list.

2.2 Access to Web-based services and applications

SODHA commits to enabling uninterrupted online access to its Web-based services and resources to the fullest extent of the means — hardware, software and human resources — made available to the archive by its funding agency.

In case of service outage, all efforts are made to restore access to Web-based services as speedily as possible. However, because SODHA is fully part of the State Archives of Belgium, which is a federal institution, possibility of intervention is limited to daytime working hours from Monday to Friday.

2.3 Digital object identifiers (DOIs)

The State Archives of Belgium has subscribed to DataCite for the provision of DOIs for datasets published on the SODHA platform. DataCite is a worldwide, non-for-profit consortium dedicated to promoting research outcomes. DataCite pursues this objective by providing DOIs, which boost the online visibility of digital objects, especially to search engine crawlers like that of Google, Google Scholar, Bing, etc.

On the SODHA platform, datasets receive a DOI once published. Files within datasets and new versions of already published datasets do not receive a new DOI.

Only published datasets receive a DOI. Datasets that are still marked as "Draft" and "Unpublished" have not actually received a DOI yet, as can be seen when attempting to access the dataset's webpage through its DOI URL before it has been submitted for review and published.



2.4 Communication and feedback

SODHA commits to reacting to requests from its users in a timely manner. SODHA endeavors to provide its users with feedback, relevant information and/or technical assistance for the use of the services of the data archive in a both pro-active and reactive manner.

As in the case of power outages, it should be noted that SODHA's reactivity to user correspondence is limited to daytime working hours from Monday to Friday.

When users specifically request access to administrative documents relevant to the workings of the data archive, SODHA, as part of the federal scientific institution that is the State Archives of Belgium, commits, in compliance with article 32 of the Belgian Constitution [DE] [FR] [NL] and the federal law on the publicity of administration of 11 April 1994 [DE] [FR] [NL], to act on a user's request within a 30-day span starting on the date when the request for access was delivered to SODHA.

SODHA makes no distinction between its users and strives to treat each and all individuals and requests equally.

2.5 Proprietary rights

As previously stated, the Dataverse software was developed by the Institute for Quantitative Social Science (IQSS) of Harvard University. Dataverse is made available for reuse under an Apache 2 license which can be consulted on GitHub. See also [1] and [2] for further information.

3. Version history of this document

Date	Version number	Changes made
30-07-2021	1.0	First version