Irish Social Science Data Archive

File Format Policy

The Irish Social Science Data Archive (ISSDA) is Ireland's centre for quantitative data acquisition, preservation, and dissemination. Based at UCD Library, its mission is to ensure wide access to quantitative datasets in the social sciences, and to advance the promotion of international comparative studies of the Irish economy and Irish society.

Introduction

ISSDA's preferred and acceptable formats are widely used by our Designated Community[1]. ISSDA's Designated Community consists of educators and researchers across multiple sectors and jurisdictions with broad relevance to the Social Sciences and Public Health.

Preferred formats are formats most suitable for sharing, reuse and preservation, while acceptable formats are less suitable for data preservation but are widely used within the social sciences community. While we recommend that Data Depositors deposit data in preferred formats, acceptable formats will also be accepted. Other formats may still be considered by ISSDA in discussion with the Depositor.

ISSDA uses a software system called NESSTAR for publishing data on the web, to allow users to find, browse, visualise and analyse data online. To be included in NESSTAR data need to be provided in SPSS format.

ISSDA also recommends that multiple formats are deposited, where possible, to allow as broad a range of researchers to access and use the data as possible. For example, researchers outside of Higher Education institutions may not have access to expensive software for analysing data or may only have access to one type of software.

If the Data Depositor is unable to deposit data in the preferred formats, an effort will be made to derive files in SPSS portable format where the source file format can be reliably converted.

---

[1] The term 'Designated Community' is an OAIS terms and refers to an identified group of potential data consumers who should be able to understand a particular set of information from an archive.

ISSDA Preferred and Acceptable Formats

Type of data	Preferred formats	Acceptable formats
Quantitative Datasets (statistical file formats)	• SPSS Portable (.por) • SPSS (.sav) • STATA (.dta) • Tab-delimited file (.tab) with setup file (for SPSS, Stata, SAS, etc.)	• SAS (.7bdat; .sd2; .tpt)
Documentation (text documents)	• PDF/A (.pdf)	• PDF (.pdf) • ODT (.odt) • MS Word (.doc, .docx) • RTF (.rtf)
Metadata	• DDI (.xml)	

For long term archiving and preservation ISSDA recommends SPSS portable format (.por). This is a proprietary format, however as SPSS is widely used among the quantitative social sciences community it can be expected to remain accessible in the future.

A recent report (February, 2017) from the Open Science and Research Initiative (ATT)[2] identified SPSS as one of the most popular statistical analysis packages whose file formats SAV and SPSS Portable have become de facto standards in the field of social sciences research. ATT found that SPSS Portable is downwards and upwards compatible and therefore be recommended for preservation.

As a result of our close relationship with our Designated Community we are aware of the importance of availability of data in these formats to them. We are also aware that the use of R for statistical analysis is becoming increasingly popular but as our preferred formats can be read into R and so largely meet the current needs of our Designated Community using R.

Documentation is converted to PDF for dissemination purposes, where possible.

In preparing these preferred and acceptable formats we have consulted the recommendations of the following organisations: Digital Archiving and Networked Services (DANS)[3], Leibniz-Institut für Sozialwissenschaften (GESIS)[4], Inter-university Consortium for Political and Social Research (ICPSR)[5] and the UK Data Archive[6].

Review

This Policy will be reviewed regularly in light of any changes in technology or the requirements of our Designated Community.

---

[2] http://avointiede.fi/documents/10864/12232/Research_data_file_formats_final_report.pdf/
[3] https://dans.knaw.nl/en/deposit/information-about-depositing-data/DANSpreferredformatsUK.pdf
[4] http://www.gesis.org/en/services/archiving-and-registering/data-archiving/preparing-data-for-submission/
[5] https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/policies/colldev.html
[6] https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats