

# **Data storage specification of WDC-RRE**

**WDC-RRE**

**Institute of Geographic Sciences and Natural Resources Research,**

**Chinese Academy of Sciences**

**March 2016**

## Content

Chapter I General Provisions.....	1
Chapter II Terms and Definitions .....	1
Chapter III Data Storage Contents .....	1
Chapter IV Data Storage Process .....	2
Chapter V Data Privacy and Security.....	3
Chapter VI Representations and Warranties .....	3
Chapter VII Supplementary Provisions.....	4

WDC-RRE

## Chapter I General Provisions

To standardize the storage and use of WDC-RRE data and strengthen the management of scientific data, this specification is specially formulated. This specification is applicable to the storage management of World Data Center for Renewable Resources and Environment (WDC-RRE) Data, including basic geographic data, nature resources data, population and social economy data, disaster risk reduction data, land use and land cover data, loess plateau agriculture and environment data, temperature spatial data, regional and thematic data, etc. Please note that we are a public good institution, and we can share your data online as open data, or as data for research use only. WDC-RRE data providers and users need to accept this clause.

## Chapter II Terms and Definitions

The following definitions shall apply to terms used herein:

- **Data Package:** the content files associated with one document plus the metadata describing each file and the combined set of files.
- **Submit/Submission:** to deposit/depositing content in the WDC-RRE.
- **Submitter:** the individual who publishes content in the repository. It is anticipated that, in the majority of cases, the submitter will be an author of the associated document or a researcher collaborating with the author(s).
- **User:** someone who views, downloads, or otherwise reuses content from the WDC-RRE.
- **Institutions:** the institutions (e.g. universities) and/or units (e.g. departments/faculties) associated with this research
- **Version/Versioning:** updates to data packages after publication to correct errors in the published files or to provide additional data supporting the findings in the associated document.
- **License:** a range of copyright licenses is available to be selected

## Chapter III Data Storage Contents

WDC-RRE aims to maintain the integrity, availability and reusability of its content over the long term. Note that it remains the responsibility of submitters to provide sufficient metadata and background information in order to facilitate users' understanding of the data. Deposits should

include all data and documentation necessary to independently read and interpret the data collection.

(1) Data File

Each variable in the data collection should have a set of exhaustive, mutually-exclusive codes. Variable labels and value labels should clearly describe the information or question recorded in that variable. Missing data codes should be defined.

(2) Documentation

Documentation files are integral to interpreting a data collection and should thoroughly explain the data collection. Examples of documentation files include codebooks, data collection instruments, summary statistics, and bibliographies of publications pertaining to the data.

(3) Preferred file formats:

- **Text: plain text** (ASCII, UTF-8), **Open Office formats** (.doc, .docx, .xls, .ppt, etc.)
  - Comma-separated (or otherwise delimited) values for tabular data
  - Semi-structured plain text formats for non-tabular data.
- **Images:** PDF/A, JPEG/JPEG2000, PNG, Esri Grid, Geo TIFF, Shp, lrp, KMZ/KML, DXF/DWG
- **Audio:** FLAC, AIFF, WAVE
- **Video:** AVI, M-JPEG2000
- **Compressed/archived formats:** GZIP/TAR, ZIP, RAR
  - Files may be compressed and/or archived when necessary due to large file size, a large number of files (more than 10), or the need to gather files together in a particular directory structure in order for them to be understood.

## Chapter IV Data Storage Process

The data storage includes the following four steps:

Step 1: The data submitter should arrange the data in accordance with the requirements, and prepare to upload the data entities and documents after standardization.

After the users log in and upload the data, the system will prompt the uploading step including how to upload your data files and supplemental files, including description documents and any associated papers to the system. Please follow the detailed submission guidelines. You will also be

asked to answer a few data provider questions about the data.

Step 2: WDC-RRE reviews the data submitted by users. When accepting data storage, examine the integrity and normalization of data and determine the level of data preservation. Revise opinions on unqualified data and resubmit them within a time limit.

WDC-RRE staff will also prepare metadata and comprehensive documentation about your data. If any questions arise during the submitting your dataset, you must respond to inquiries in a timely manner. If you do not respond, the WDC-RRE will place the data submission on hold.

Step 3: Store the data that pass the audit. The WDC-RRE will prompt the uploader to submit the final data package before the release of data and make a backup of the data.

step4: After a data collection is deposited at WDC-RRE, a data archivist reviews the data and documentation, builds a study description, enhances the documentation, approves the data collection for distribution on the WDC-RRE website, and archives the data for long-term preservation.

## **Chapter V Data Privacy and Security**

WDC-RRE respects the privacy of our users and submitters and works to protect all personally identifiable information we collect.

Confidentiality in the research environment basically means planned and careful processing of personal data. Personal data should only be collected and processed to the degree necessary for the research, and unauthorized access to the data must be prevented.

Periodic verification that stored content remains uncorrupted. If data objects are found to be corrupted, affected data will be restored from known-good copies as appropriate.

Data security means keeping personal information collected, as well as computer systems, data files and transfers of data safe. Making back-ups of data files is this integral parts of data security.

## **Chapter VI Representations and Warranties**

Submitter represents and warrants that s/he is an author of the document associated with the submitted content or has been granted authority by the authors to submit and publish content in the repository.

By submitting content to the repository, submitter grants WDC-RRE irrevocable permission to make the data package available to the public. Datasets can be viewed and downloaded from the

WDC-RRE website by end users.

Submitter represents and warrants that the content does not contain any information (i) which identifies, or which can be used in conjunction with other publicly available information to personally identify, any individual; or (ii) which is false or misleading.

Ordinarily, one data package will consist of no more than 20 Gigabytes of content, but data packages containing more than 20 Gigabytes of content may be accepted.

The primary language of the content must be English. WDC-RRE curators cannot effectively review or curate submissions in other languages at this time.

## **Chapter VII Supplementary Provisions**

WDC-RRE reserves the right to alter services offered by WDC-RRE and/or these terms of service at any time in order to ensure their appropriateness to WDC-RRE's mission. WDC-RRE will post any changes on the website or a link therefrom and make reasonable efforts to communicate material changes to submitters; it is the responsibility of submitters and users to periodically check the website for such changes.

The right of interpretation of the relevant provisions of this specification belongs to WDC-RRE.

If there are any suggestions or comments about this document, please contact us:

Tel:+86-10-64889048-8006

E-mail: [wdc-rre@lreis.ac.cn](mailto:wdc-rre@lreis.ac.cn)

Address : 11A, Datun Road, Chaoyang District, Beijing, 100101, China,  
Institute of Geographic Sciences and Natural Resources Research, CAS.

WDC-RRE