

D1.3: System requirements

Requirements (as collected in the interviews)

- Simple top-layer (web interface) facility to upload file, must be easy to understand for newcomers and external users, which means few metadata fields.
- Command-line tools for uploading and managing data
- Access as a file server/local drive
- Simple search facility, web-based (few clicks to get to the desired files).
Examples: "find all tracks with cross-talk and a female voice in the multi-track data sets"
- To achieve simple search, need for custom tags/keywords/terminology to be saved as metadata, specific for the data sets, as well as free-field descriptions
- Ability to select (using for example tags) and download entire data sets/subsets as zip files. If data is linked to software, also download the correct version of the software automatically
- Ability to dump/retrieve metadata
- Ability to link data sets to software (see OFF ontology), preferably from sound software.ac.uk
- Versioning of certain kind of data (e.g. manual annotations, complete data sets)
- Incremental backups
- Data encryption for sensitive data
- Support for RDF
- Use of DOIs, with specific reference to the version of the data set
- Ability to create simple websites to expose a data set
- Data must be available in the longer term (i.e. internet archive)
- Online as soon as possible (at least basic upload and search functionalities)
- Short learning curve and short overhead for using it compared to only having the files around on the HDD (PhD students don't have time)
- Clear policies/guidance on how to publish sensitive data (e.g. anonymized data from human-based studies)

Results of the online survey (19 respondents)

- Upload via web: 12
- Virtual drive: 11
- Zip download: 9
- Command line tools: 9
- Advanced web search: 8
- Links to software: 8
- Versioning: 8
- Publication website: 7
- Customisable metadata: 6
- RDF support: 4
- DOI support: 1