

# ***LaZAR: The Repository for Regional Research Data in the Humanities (handout)***



## **1. Why LaZAR?**

Do you do field work and also produce a lot of heterogeneous digital data?

Then you are surely often confronted with the following questions:

- How do I make my multimedia research data citeable?
- How can I ensure the permanent, stable and secure distribution of data that I make available online?
- What measures can I use to protect my data from distortion and manipulation?
- How do I link old data with new data?
- How can I accomplish the simple citeability of the data which have sensitive contents and are therefore not suitable for public release?
- How do I regulate my project partners' authorisations in the digital survey, processing and publication of the data?
- How do I deal with multilingualism in my research data?
- How do I protect my sources' personal information without violating data protection regulations?
- And finally, a very trivial but completely legitimate question: Can my multimedia research data be as researchable and accessible as data in archival collections or the books and periodicals in a library, systems we are already familiar with?

## **2. Our Service**

To solve these problems and more, the Institute of Slavic Studies and Caucasus Studies at Friedrich Schiller University Jena, along with its cooperation partners, the head office of the Common Library Network in Göttingen and the TIB – Leibniz Information Centre for Science and Technology and University Library in Hanover, created and developed a modern, high-performance web product called LaZAR.

## **3. A Look at How the Infrastructure Works**

At the moment, the following tasks can be completed with LaZAR:

- The processing of research data and making them accessible
- Researching using and re-using research data

## **4. Instruments**

The LaZAR infrastructure offers the following instruments to complete tasks and solve the problems mentioned above (see Image 1)

1. An interface for uploading research data
2. A research module
3. A list of the values or objects located in the sub-databases
4. A selection of three display appearances for the research modules
5. A tool for the selection and the individual organisation of the way data are displayed, including the function to alternate between an organised and a flat hierarchy of the data sets
6. A tool for sorting data
7. A reader / editor for the data sets
8. A filter
9. Folders for the aggregations
10. Language selection
11. A tool for setting up a personal user account
12. A navigator for the pools
13. A mask for the expert search
14. Various other functions available via a right mouse click, including:
  - A full-screen display
  - Downloading
  - Exporting
  - Processing
  - Editing (individually or using the group editor)
  - Deleting
  - Placing items in a folder

## 5. Functions

The database provides the following functions for the processing and re-use of research data:

- The large-scale uploading of research data in various formats
- The direct submission of live sound, video and image recordings
- The hierarchical classification of the data
- The automatic recognition and acceptance of technical metadata
- The addition of structural and descriptive metadata to the research data
- The use of interfaces for entering controlled values when adding metadata to research data (such as Integrated Authority Files [GND], geographical names and the names of the world languages)
- Minimisation of the workload through the completion function when entering values which have already been saved and through the large-scale data processing function
- Consultation of the collections
- The use of one or more filters, search functions and hierarchical management functions
- The downloading or export of entire collections and individual data files or digital objects
- The embedding of the research data in electronic as well as conventional publications
- The creation of folders and management of data in the aggregated data sets
- Much more

## 6. Publishing in LaZAR

### Creating, Editing, Managing and Providing Access to Research Data

It is only possible to use the services to process and make your own research data accessible in the LaZAR infrastructure after being granted the appropriate authorisation. To do so, you will need to register with LaZAR as a contributing user, free of charge.

Applicants need to download a form, fill it out, sign it and send it back via e-mail. Within two working days, the sender will confirmation about the change in their status or a notice explaining the reason for their rejection. Prior to applying for contributing user status, the applicant can work with the database as a self-registered user with a significantly limited authorisation.

To self-register, fill out an online form and accept the terms of use.

After receiving contributing user status, contributing users can upload their research data to the database, add metadata, and *provide them for publication*. This task leads the user to their own workspace, the *pool*, which they are authorized to use while they are contributing users to LaZAR. A pool can also be used by multiple contributing users.

All of the research data in LaZAR are, without exception, categorized in a collection. In accordance with this requirement, it is recommended that the users of a pool create an empty collection right when they begin their work. This is possible via the function **Menu / New Data Sets**. By clicking on **Next**, an empty template without any objects is generated. A characteristic image (photo, sketch, etc.) which represents the whole collection can be uploaded as an object. In the root collection, the hierarchy field remains empty.

When subordinating other data, it is sufficient to enter the name of the previously created collection in the **Hierarchy Level** field in the appropriate data set.

If a stable internet connection is available, it is possible to use the interface to directly submit live audio, video and image recordings. This applies to smartphones and tablets.

A special advantage of the database is the large-scale uploading / processing of the research data. To do so, template masks are available to users for the creation of the new data and for the processing of already created data.

Saving new entries in the template ensures that each of the data sets included in the template is labelled with the same information.

When processing the research data which are being considered for release in the LaZAR repository, the data must be labelled with the following required metadata: **Title**, **Originator** and **Language**.

The database offers the possibility to concurrently describe some of the digital object's characteristics in three languages. To do so, the fields with free text options are available: **Title**, **Text**, **Description** and **External Relationship**.

The valid languages are German, English and one of the other languages included in ISO 693-3.

## **7. Researching in LaZAR**

### **Searching, Aggregating, Embedding**

The LaZAR repository features a permanent, growing inventory of published research data collections. A collection is not comparable with an anthology.

The materials in the collections can actually be organised by any desired data characteristic. There are, however, some characteristics that may be redundant in terms of use for the systematisation of data. For example, this would include classifying data by their format (image, video, audio, Office document), because the format is one of the characteristics that is automatically recognised by the system and is separately featured in the filter. Therefore, it is more useful to organise the data in the collection by content or thematic characteristics.

Every collection is published under the authorship of individual people or groups of people. Usually the entire concept belongs to the author, from the initial survey to the indexing of the research data included in the collection. By doing so, the originator does not just account for data produced by them, but also the data from other originators, including data from amateur authors.

Research in the repository is possible using the search function. At any time, the display's appearance can be changed; however, the most comfortable appearance is the text view, because you can also read a short description next to the image if one has been provided.

Every appearance offers the possibility to view the data in a hierarchical as well as a flat context; that can be done by placing or removing a tick mark in a particular spot (Image 1, Point 5).

Image 1

