

**Service ID** S00063



**Location** Poland, Remote

## Curation and provision of annotated datasets

### Provider service

Lukasiewicz Poznanski Instytut Technologiczny (L-PIT), Poznan Supercomputing and Networking Center (PSNC)

### Link to content

<https://agrifoodtef.eu/catalogue-of-services/curation-and-provision-annotated-datasets>

### Type of Sector

Arable farming, Food processing, Greenhouse, Horticulture, Livestock farming, Tree Crops, Viticulture

### Accepted type of products

Data, Design / Documentation

### Type of service

Collection of test data, Provision of datasets

### Description

The service provides access to annotated data collected from devices tested as part of other project services. It includes both data annotation services (manual and automatic) and the provision of annotated databases or data packages. In addition, we offer disc space for temporary archiving of the collected data along with its annotation in cases where this is not possible at the client's site. As a result of the service, procedures for both manual and automatic annotation will be identified that are appropriate to the type of measurement data and the way it is collected and assembled. The methods of labelling and assigning specific data classes collected by the agriFoodTEF technical infrastructure will be proposed. Possibilities for documenting and storing data in datasets will be indicated. In addition, procedures for extracting selected features from stored research data can be offered.

## **How can the service help you**

This service provides secure access to the data streams collected by the investigational devices tested as part of the services run in the agriFoodTEF programme. The service provider provides annotation of the data collected, either manually or automatically as required. The method of annotation will depend on the method of data collection or transmission and the devices, protocols, etc. used for this purpose. The scope of the service also includes the possibility to access the archiving of the collected data with its annotation in cases where the principal does not have this possibility. The archiving can be temporary or long-term depending on momentary needs so that there is no need to interrupt the experience, the test. In addition, the service offers an indication of the ways in which the data collected by the agriFoodTEF technical infrastructure can be labelled and assigned to specific classes. The service provider will also suggest ways to document and store the data in the datasets. Furthermore, procedures can be offered for extracting specific traits from previously held and annotated data.

## **How the service will be delivered**

Within the service, we employ a methodology in line with the guidelines of commonly applicable documents (standards, directives, etc.). Based on this, we will plan and carry out the process of curation and annotation of the dataset. We require a functional system to be tested, allowing us to generate the dataset. All information is treated confidentially, with the option to sign a Non-Disclosure Agreement (NDA). If the tests and the datasets need to comply with specific industry standards or regulations, kindly inform us in advance. The timeframe and costs are determined individually based on the scope and complexity of the tests and the relevant properties of the datasets in question. We ensure flexibility and professional support at every stage of the testing process, including both physical and virtual analyses. Please feel free to contact us to discuss the details and tailor the service to your unique needs.

## **Service customisation**

Our service to provide and curate annotated datasets is based on a structured workflow. The time, duration, and other details of the service can be adjusted according to the characteristics of a problem at hand and specific customer needs. The customer needs to provide the system to be tested used to generate the datasets. As at, the customer will obtain an annotated dataset and its curation.