

**Service ID** S00172



**Location** Austria, Remote

## Data augmentation to optimise AI model training and prevent the model f

### Provider service

Josephinum Research

### Link to content

<https://agrifoodtef.eu/services/data-augmentation-optimise-ai-model-training-and-prevent-model-overfitting>

### Type of Sector

Arable farming, Horticulture, Livestock farming, Tree Crops, Viticulture

### Accepted type of products

Data

### Type of service

Data augmentation

### Description

In preparing image or sensor data, pre-processing is crucial. We apply techniques like normalisation for data uniformity and correlation analysis to uncover patterns, enhancing interpretability. Dimension reduction optimises computational efficiency, while adding randomised noise and artificial expansion improves model generalisation. We implement anonymisation to protect data privacy and ensure precise data labelling for supervised learning. These methods, tailored to your needs, enhance your data's quality and usability.

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

This service enhances the quality and usability of customer data. Before the service, customers may face challenges such as insufficient or non-uniform data, data privacy concerns, or a lack of properly labelled datasets. After the service, they receive enhanced datasets that are better normalised, more diverse, anonymised for privacy, and precisely labelled. This improvement facilitates better performance in AI models and analytics while ensuring compliance with privacy requirements. The augmentation methods are selected based on the customers needs and vary from case to case.

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

Customization options include selecting specific preprocessing techniques, such as anonymization for privacy-sensitive data or targeted labelling for supervised learning. The service can be tailored to meet unique requirements, such as focusing on specific data patterns or ensuring compatibility with particular AI frameworks. Limitations may include the nature and quality of the input data and compatibility with the augmentation tools. Customers are encouraged to communicate any specific needs or constraints in advance.

## **Service customisation**

The service is delivered remotely. The process involves multiple iterations to refine data quality, applying techniques like normalisation, dimensionality reduction, noise addition, and labelling. The duration of the service depends on the dataset's size and complexity, generally taking a few days to weeks. Customers receive a detailed report and the augmented dataset ready for use in their AI or analytics workflows. Customers need to provide the raw datasets and specify their goals for augmentation.