Service ID S00177

Location Italy, Remote

Design of testing protocols for digital testing

Provider service

Politecnico di Milano (POLIMI), Università degli Studi di Milano (UMIL)

Link to content

https://agrifoodtef.eu/catalogue-of-services/design-testing-protocols-digital-testing

Type of Sector

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

Accepted type of products

Data, Design / Documentation, Software or AI model

Type of service

Test design

Description

Any test activity involves three main components, i.e., environment (where the tests take place), protocol (defining what activities are executed and how), and evaluation metrics (used to assess the results of the tests). This service concerns the second element, i.e., the design of the testing procedure for digital systems such as (for instance) AI models or computer vision software. The digital environment and the evaluation metrics can be designed – if required – via services S00176 and S00178.In the context of testing customers' solutions within digital environments, this service is targeted at designing a suitable protocol for digital testing based on the use cases specified by the customer. The components of the testing protocol can include:Selecting the datasets to be used for testingSelecting reference AI models to be used for testing (if needed)Choosing data formats and metadata standards Defining data pre-processing and preparation steps Defining values and ranges of test parametersDefining the different phases of the protocolOutlining the operations to be executed in each phaseThoroughly describing the protocol specifics to ensure reproducibilityDatasets to be used for testing can be provided by agrifoodTEF and/or by the customer; if nothing suitable is available, other agrifoodTEF services can be leveraged to collect and/or generate tailored data.The technical team executing this service comprises expert engineers but can also involve agronomists when this is necessary to ensure the relevance of the tests for the use case, e.g., to determine the distribution of test repetitions across the variation ranges.

How can the service help you

Building a digital system (e.g., a software module) that solves a problem and performing thorough testing of such a system are separate and different activities and involve different competencies. Testing is important not only to check that the system performs satisfactorily but also to identify ways to further improve it.

This service supports the customer in defining the optimal experimental protocol to validate the system with data relevant for the customer's use case, to enable quantitative performance evaluation, and to demonstrate its performance to potential users.At the end of the service, customers are provided with a fully documented digital testing protocol which they can immediately use to set up their own experimental activity.

If required, agrifoodTEF can support the customer in designing also the computational environment needed for testing and the evaluation metrics (via services S00176 and S00178), in the setup of the experimental activities (via services S00180 and S00181), and in the execution of the tests (service S00182) and associated data collection (service S00183). Finally, agrifoodTEF can also support the customer in performance evaluation (service S00184), thus offering the full set of activities composing a digital testing pipeline.

How the service will be delivered

This service description is intentionally generic. Every instance of this service is in fact customised to adapt it to the needs and requirement of the specific customer. The following is an example of service instance.

Example service: In the context of testing the capability of a Computer Vision model to discriminate weeds from crops, a list of candidate methods is defined that are already largely-applied on the market, to consider as a benchmark and compare against the customer's existing solution. Incremental variations of these models are also identified (e.g., values and ranges of model parameters, isolation of individual sub-modules during training and fine-tuning, different optimization functions, etc.). Based on the model size and performance on different datasets, estimates are made of the minimum number of examples required per weed and crop class for training the models.

A list of comparable datasets available within the consortium and/or publicly is also defined to reduce the model training cost and reuse existing datasets wherever possible. In the defined experimental protocol, methods are tested 10 times in a row, to collect average and standard deviation values for all evaluation metrics.

Service customisation

The duration of this service is on average 1-3 weeks.The first phase involves one or more interviews, in person or remote, where the customer provides information about the features of the system(s) to be tested and the performance elements that the protocol to be designed needs to investigate. Additional documentation (e.g., details about the internals of software components) may be provided (under NDA if confidential) in order to precisely match the protocol to the necessities of the customer's system.

After this preliminary phase, protocol design takes place. Additional interaction with the customer will occur to ensure compliance with their necessities and to fine-tune the design. At the end of the service, the customer receives a comprehensive design that they can immediately employ to set up and perform the tests.