Data Toolkit

BigDataStack Software Component developed by Ubitech

Description

The Data Toolkit is the component, which takes care to design an end-to-end Big Data application graph. Through the Data Toolkit is feasible to execute valid data analytics pipelines with auto-fill fields and services descriptions minimising the manual configuration required by the data scientist. It supports Spark and Spark Machine Learning (ML) library, but it may be also extended to support other machine learning frameworks (namely indicatively Tensorflow29, Scikit-learn30, PyTorch31, etc.).

Features

The main features of the component are, as follows:
- Automation;
- Online graph composition and application constraints enforcement;
- Auto-fill capabilities with services descriptions from the cloud infrastructure.

Areas of Application

Machine Learning and Application Setup Toolkit

Data Toolkit

Market trends & opportunities

End-to-end Big Data orchestration applications and machine learning pipelines set up and deployment.

Customer benefits

Easiness; Automation; Advanced User Experience / UI.
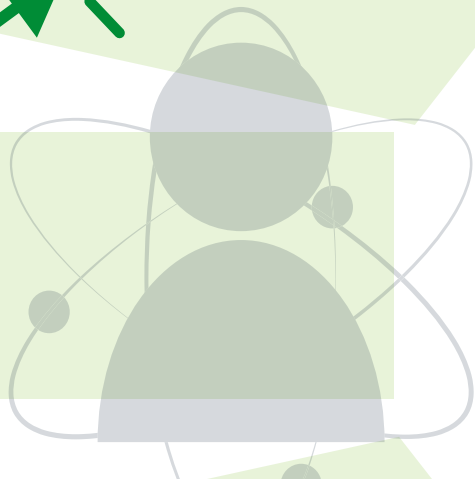
Technological novelty

N/A.

TRL level: 5

Find the Open Source code here:

- http://bigdatastack-tasks.ds.unipi.gr/Sidiropoulos/bigds-ui

bigdatastack.eu