EU Innovation Radar Identifies I-BiDaaS Innovations and Key Innovators

The Innovation Radar is a European Commission initiative to identify innovations with high potential, key innovators in EU-funded research and innovation projects and market readiness of these innovations. The Innovation Radar platform builds on the information and data gathered by independent experts involved in reviewing ongoing research and innovation projects funded by the European Commission. These experts also provide an independent view regarding the innovations in the projects and their market potential. The aim is to make information about EU-funded innovations from high-quality projects visible and accessible to the public via the Innovation Radar platform.

Five innovations of the I-BiDaaS project were evaluated as excellent innovations by the EU Innovation Radar. One innovation was deemed market ready[1], one business ready and three exploring. The institutions that were main developers of the five innovations were identified as key innovators and the innovations have been published in EU innovation radar platform.

Moreover, having as a goal to broaden the audience, as well as to help carry out matchmaking for potential future users of the I-BiDaaS platform and its components, these five Innovations detected by EU Innovation Radar, were submitted in the BDVe Innovation Marketplace, an excellent dissemination and exploitation opportunity for I-BiDaaS to promote all the innovation results developed within the course of the project. The Innovations have been published published after the review process.

[1] https://www.innoradar.eu/methodology/#maturity-info

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.

Multidimensional Storage with Efficient Sampling (MuSES)

Deep Tech Innovation

Market Maturity: Market Ready

Key Innovator: Barcelona Supercomputing Center (BSC)

Brief description: Our technology allows organizing the data according to their multidimensional attributes while building stratified samples at high-speed. Such a design has three main advantages: we can use a single system to manage transactional and analytical workloads; we enable interactive data exploration, both via visual interfaces and comparative analysis; third, Machine Learning algorithm can exploit a new type of data parallelism achieving faster results with a fraction of the computation cost.

Main features:

- Faster analysis of data using fewer computing resources
- Enable analyzing data that was previously difficult to assess due to its size
- Computation flexibility through cloud resources
- Fast and scalable Machine Learning algorithms
- 3D Visual explorative analysis with space reduction techniques
- Integrated transactional and analytical platform
- Timely results: Choose between precision and execution time

Technological Novelty:

- Qsience: A laboratory for data science, equipped with ready to use tools to analyze Big Data on our platform.
- Qdev: The gateway to Qbeast's engine. The place for Data Engineers to build and design data processing pipelines.
- Qinsights: The BI tool for the data-driven business users. Empower your whole organization to use Big Data insight.
- Qadmin: Access and Security Management. Console to manage multiple user access. Administer security groups, create advanced access control rules, review data accesses and more.

Find out more…

EU Innovation Radar

BDVA Innovation Marketplace

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.

AEGIS Advanced Visualization Toolkit (AVT) supporting scalable data visualisation

Deep Tech Innovation

Market Maturity: Business Ready

Key Innovator: AEGIS IT Research (AEGIS)

Brief description: AEGIS has designed and implemented the AEGIS Advanced Visualization Toolkit that is an extensible software with a wide application scope, ranging from Digital Forensic Analysis to Big Data analytics.

The AEGIS Advanced Visualisation Toolkit provides an extensible platform for data visualization and transformation. It is a versatile solution that can use different visualization techniques to cover the needs of IT experts as well as non-IT literate users. Via its intuitive monitoring dashboards, configurable visual representations and collaboration features, the toolkit serves as an advanced system that can support easy exploration and insight gaining from big volumes of multidimensional data.

Main features: AEGIS Advanced Visualization Toolkit (AVT) is a set of specialized tools that can turn input data into meaningful insights through creative visualizations. Key features include:

- Visualizations of both batch data analytics and streaming analytics
- Processing of large and complex data sets
- Addressing needs of experts and non-IT experts
- Multiple Granularity in data representations
- Advanced visualisations that include dynamic real-time graphs, spatiotemporal representations (interactive maps, etc) apart from traditional charts (pie, bar, etc)
- Extensibility and customization for specific needs and requirements
- Powerful User & Access Management

Technological Novelty: AEGIS AVT is designed with modularity and scalability in mind, offering multidisciplinary, intuitive and effective visualization and interaction capabilities for the end-users. Several types of programming languages and advanced visualization tools have been used to develop a platform easy-to-use for experts and non-expert users. The agnostic-domain nature of AVT allows the user to grasp information and insights from o large amount of data with ease. AVT also helps decision-makers to identify focus areas between multi-dimensional data-sets and offers effective features and tools to interpret the data.

Find out more…

EU Innovation Radar

BDVA Innovation Marketplace

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.

ADMM Machine Learning Algorithms

Deep Tech Innovation

Market Maturity: Exploring

Key Innovators: University of Novi Sad Faculty of Sciences (UNSPMF), Barcelona Supercomputing Center (BSC)

Brief description: This is an open-source implementation of the Alternating Direction Method of Multipliers (ADMM)[2] optimization algorithm that relies on CVXPY[3], a Python-based toolbox for convex optimization, and implemented with the COMPSs programming model[4]. The solution has been developed in the context of the I-BiDaaS project[5], and it allows to train several Machine Learning models, including Least Absolute Shrinkage and Selection Operator (LASSO)-based and elastic net regression, logistic loss-based classification, and clustering, as well as several other models with minimal additional coding effort. A number of related implementations are available at the I-BiDaaS knowledge repository[6], and the ADMM implementation of LASSO has been included in the dislib library[7], an open source COMPSs based library oriented to Machine Learning. This innovation has been has been applied to several industrial real data sets in the context of the I-BiDaaS project.

Main features: The implementation can be used to train various Machine Learning models, including regression, classification, clustering over a computer cluster over which COMPSs has been installed. Users can run a pre-defined Machine Learning model, or they can encode a new Machine Learning model by setting a different objective function for training. In this sense, the available implementation can be seen as a code template, where a new Machine Learning model can be obtained with a minor programming effort. Current implementation assumes that the input data is structured, organized into numerical-valued matrices, and split in multiple files, one per machine in the cluster. Parallel, scalable execution over the cluster is achieved by the inherent parallelization of ADMM and the underlying COMPSs runtime system

Technological Novelty: This solution is novel essentially due to the reusability of the developed code, since with easy and small code changes new models can be supported. It is the first ADMM method implementation in COMPSs and corresponds to a novel

[2] Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, Foundations and Trends in Machine Learning, 3(1):1–122, 2011.
[3] CVXPY: A Python-Embedded Modeling Language for Convex Optimization, S.
Diamond and S. Boyd, Journal of Machine Learning Research, 17(83):1-5, 2016.
[4] ServiceSs: an interoperable programming framework for the Cloud, Journal of Grid Computing, March 2014, Volume 12, Issue 1, pp 67–91, Lordan, F., E. Tejedor, J. Ejarque, R. Rafanell, J. Álvarez, F. Marozzo, D. Lezzi, R. Sirvent, D. Talia, and R. M. Badia, DOI: 10.1007/s10723-013-9272-5
[5] http://ibidaas.eu/
[6]
https://github.com/ibidaas/knowledge_repository/tree/master/tools_technologies/sources/batch_processing/unspmf
[7] J. Álvarez Cid-Fuentes, S. Solà, P. Álvarez, A. Castro-Ginard, and R. M. Badia, "dislib: Large Scale High Performance Machine Learning in Python", in Proceedings of the 15th International Conference on eScience, 2019, pp. 96-105.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.

addition to dislib. The technological novelty also lies in the integration of COMPSs and CVXPY through the ADMM framework to develop efficient methods for Machine Learning training. In this way, we exploit for the first time the benefits of parallel execution due to COMPSs, and the efficient convex optimization problems solutions due to CVXPY.

Find out more…

EU Innovation Radar

BDVA Innovation Marketplace

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.

Parallelization of Constraint Satisfaction Problems (CSP) solver

Deep Tech Innovation

Market Maturity: Exploring

Key Innovator: IBM Israel – Science and Technology LTD (IBM)

Brief description: The CSP solver is the heart of the IBM's Data Fabrication Platform technology. The generated data is the solution of a constraint satisfaction problem created by the Platform engine out of the user defined data model. The complexity of a CSP depends on the number of variables and the complexity of the defined constraints. To fabricate significant amount of synthetic data many CSPs should be solved. The aim of this innovation is to solve many such CSPs concurrently. Cross-CSP data dependencies are handled as well by the new parallel solver. The new implementation enables order-of-magnitude improvement of the Platform performance and fabrication of big data.

Main features:

1. Solving many CSPs concurrently
2. Support for cross-CSP data dependencies
3. Order-of-magnitude improvement of the Platform performance and support for fabrication of big data

Find out more…

EU Innovation Radar

BDVA Innovation Marketplace

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.

Specification of an end-to-end Big Data as-a-self-service platform

Deep Tech Innovation

Market Maturity: Exploring

Key Innovators: University of Novi Sad Faculty of Sciences (UNSPMF), The University of Manchester (UNIMAN), ATOS Spain SA (ATOS)

Brief description: This innovation corresponds to specification of an end-to-end platform for Big Data as-a-self-service, developed within the I-BiDaaS project. The platform specification is mainly based on the consolidated system requirements derived from the I-BiDaaS project end user requirements analysis (functional requirements), and non-functional requirements which stem from software engineering good practices. The specification takes into consideration existing research in big data requirements and complies with existing Big Data Reference Architectures. The platform involves a number of open source and proprietary technologies provided by the I-BiDaaS partners.

Main features: The platform specification ensures multiple functionalities offered to end users, including fabrication of realistic synthetic data for experimentation and testing, batch and streaming analytics, and simple, intuitive, and effective visualization and interaction capabilities. The specification assumes three different modes of the platform, offered to different user types.

1. The **Self-service mode** allows a user that has the required domain knowledge and a degree of (non-expert) knowledge about data analysis to construct Big Data pipelines in a user-friendly way, by selecting a pre-defined data analytics algorithm from an available list.
2. The **Expert mode** allows experts (Big Data developers) to upload their own data analytics code based on the available reusable templates.
3. The **Co-develop mode** corresponds to an end-to-end solution for a given industry project that is developed with the use of expertise of the personnel that builds and maintains the platform.

Technological Novelty: The technological novelties are manyfold. For example, the results produced by the batch processing module are fed back to the data fabrication tool; these results are then used for training and to help building rules that will be used for future data generation purposes. Furthermore, the platform specification allows for stream processing that the parts of the streaming analytics that can be parallelized are offloaded to the GPU-accelerated streaming analytics module. By carefully performing part of the queries at the lowest level (especially for filtering), only the required data is forwarded to the stream-processing engine for a more sophisticated analysis, while the remainder is ignored at the earliest possible.

Find out more…

EU Innovation Radar

BDVA Innovation Marketplace

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.

Find & Follow us

Website | Twitter | LinkedIn

Zenodo | OpenAIRE | GitHub

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780787.