

BigDataStack, a Holistic Stack for Big Data Operations and Applications

*Results, Adoption,
Open Source and
Impact for
End-users*

After 3 years of research and innovation, the BigDataStack consortium proudly delivered its complete infrastructure management system, which bases the management and deployment decisions on data from current and past application and infrastructure deployments. This complete infrastructure management system is delivered as a full “stack” that facilitates the needs of operation data and application. BigDataStack for end-users is depicted in the infographic below.

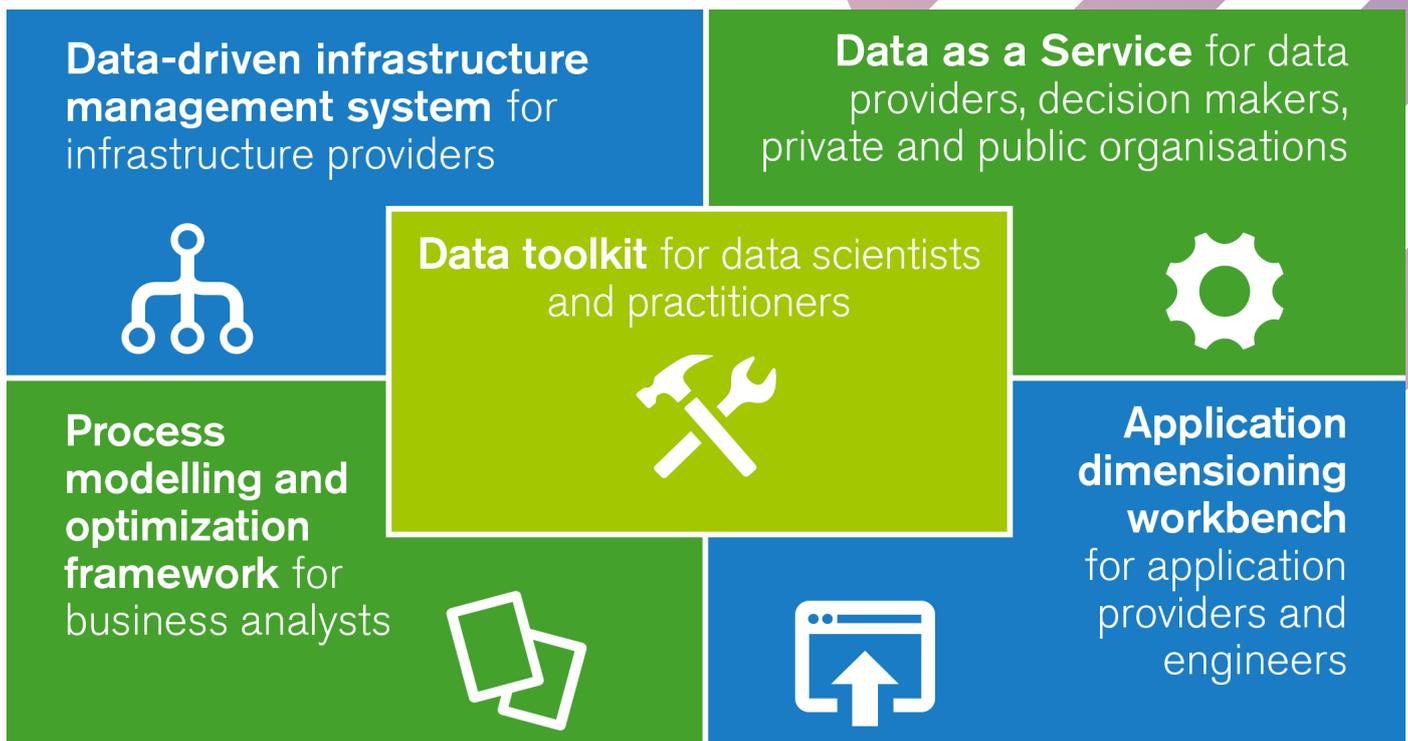


Figure 1: BigDataStack added Value for End-Users

“BigDataStack delivers a complete pioneering stack, based on a frontrunner infrastructure management system that drives decisions according to data aspects, thus being fully scalable, runtime adaptable and high-performant to address the emerging needs of big data operations and data-intensive applications.”

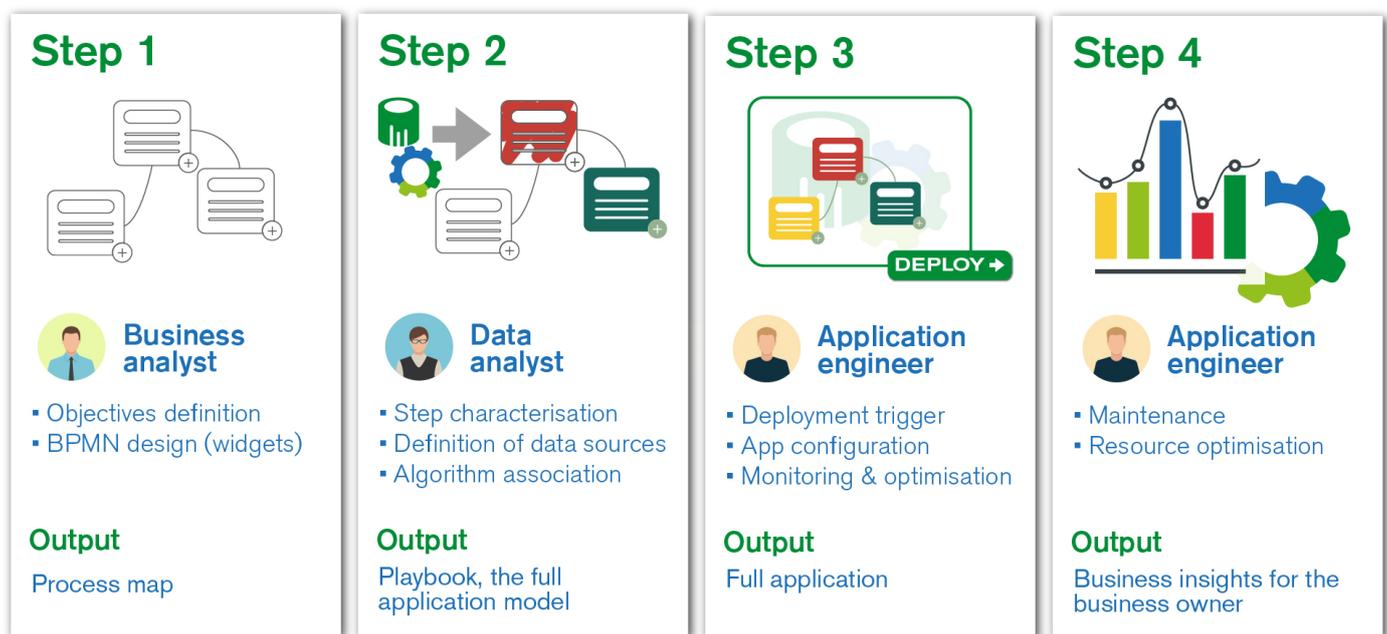
*Dimosthenis Kyriazis / Technical Coordinator
(University of Piraeus)*

Applicability and validation

To enable data operations and data-intensive applications to fully exploit the sustainability of BigDataStack and take full advantage of the developed technologies, the consortium onboarded three use cases that exhibited the applicability.

1. **REAL-TIME SHIP MANAGEMENT**, where the algorithms optimise and help cut costs on maintenance and spare parts inventory planning and dynamic routing.
2. **CONNECTED CONSUMER**: where the grocery retailer is provided with optimal insights into consumer preferences and improve the effectiveness of marketing strategies for improving consumer shopping experience.
3. **SMART INSURANCE**: where a multi-channel scenario facilitates data analytics-powered smart insurance, providing a 360-degree view of the customer and personalized services.

Each of the use cases can be explored further through a short [explanatory video](#) and an [end-to-end demonstration](#), available at the project website.



A Journey into BigDataStack



Figure 2: BigDataStack User Journey

Further adoption of BigDataStack

Data is amongst the most valuable products nowadays, its exploitation for industry will twist the market over the next few years. BigDataStack defined its [adoption roadmap](#), for organisations interested in implementing the [full BigDataStack](#) or its individual components, as developed in its [exploitation plan](#). BigDataStack offers a perfect solution for Industry, handling real-time analytics on big data in flight and at rest, across multiple data stores and from several sources. Here we provide a summary of the main results of BigDataStack, after three year of research,

adoption and team building among the [fourteen public and private partners of the consortium](#).

In 2020, the third and final year of BigDataStack, the consortium organised the adoption and exploitation focused events for end-users [A BigDataStack for Industry - a focus on retail, shipping and insurance: Is Big Data the real future of emerging business?](#) and [European Big Data Research for Industry – 3 projects, 7 sectors, 9 applications, 41 software components. Now what?](#)

In addition to the three BigDataStack use cases Smart Insurance, Real-Time Ship Management and Connected Consumer, another [four exploitation stories were documented](#).

“We focus specifically on monitoring a vessel in real-time, the prediction of possible malfunctions in the main engine, the timely spare part order and the optimisation of the requisition process.”

*Dr. Takis Varelas
from Danaos, on
the Real-Time Ship
Management Use
Case*

The BigDataStack Catalogue of Software Components

The BigDataStack project published all software components, Open Source and proprietary, in its [online catalogue](#). The catalogue provides you with downloadable factsheets, for each of the components, with information on its license and access to the code on GitHub for the Open Source components. As the consortium is so very proud of each of its components and wants to make sure you don't miss out on them for your own work on Big Data analytics. Partners have prepared a series of videos on the BigDataStack components. You can watch them all [here](#).



Figure 3: BigDataStack Architecture

European Open Source Initiative

“The importance of OS is increasing and can be seen from the users and suppliers sides. From the supplier side, we see big companies such as Microsoft, Apple, Facebook and IBM, actively participating in OS communities and even contributing to their OS solutions.”

Céline Xu, Big Data expert at the BigDataStack final event
Case

BigDataStack Upstream Contributions to Open Source Communities

- *Spark*
- *Kuryr-kubernetes (CRDs, Network Policies, E/W distributed loadbalancer)*
- *Kuryr-Tempest-plugin (testing of Kuryr-kubernetes)*
- *Octavia (integration of ovn-octavia)*
- *Neutron (integration of ovn-octavia plus bug fixes)*
- *Kubernetes (testing side)*
- *Gophercloud (integration of OpenShift on OpenStack)*
- *Terraform (integration of OpenShift on OpenStack)*
- *OpenShift Installer (integration of OpenShift on OpenStack)*
- *Cluster Network Operator (Kuryr integration)*
- *Cluster API and Cluster API provider OpenStack (integration of OpenShift on OpenStack)*
- *Machine Config Operator (integration of OpenShift on OpenStack)*

In October 2020, the European Commission approved its new [Open Source Software Strategy 2020-2023](#), a part of the overarching [Digital Strategy Of The Commission](#) and contributing to the Digital Europe programme. The internal strategy is called “Think Open”, and sets out a vision for encouraging and leveraging the transformative, innovative, and collaborative power of open source. It promotes the sharing and reuse of software solutions, knowledge and expertise, in order to deliver not only better European services, but services that benefit society and have lower economic cost. The aim of the strategy is to increase the use of Open Source in both practical and strategic sectors, while recognising the importance of collaboration between governments, researchers, industry, and the public at large for building new, innovative digital solutions that work across borders and towards technological sovereignty.

For the benefit of the BigDataStack project, the strategy is well aligned with its principles as the collaboration across the Commission, with Member States, companies and the public at large for building new, innovative digital solutions that work across borders and towards technological sovereignty is key. This openness and collaboration is very much in line with the [European Commission’s Goals For A Digital Europe](#) and, in this particular case, for Open Source Software.

BigDataStack introduced the [European open-source initiative](#) towards increased impact of project outcomes. The European Open Source Initiative is led by RedHat and open to the complete research community, welcoming participation from EU-funded projects, industrial and academic institutions, and individual contributors. Through the initiative the project has published 18 software components in Open Source, carried out 4 upstream contributions, of which one by IBM. As part of the European Open Source Initiative, RedHat started to organise the RedHat Research Day, a one-day event connecting research and Open Source developers.

BigDataStack for End-Users in Numbers

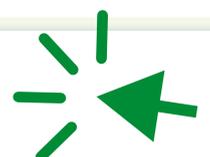
- 1** **BIGDATASTACK** | with 21 Software Components
- 2** **BUSINESS MODELS FOR EXPLOITATION** | Individual & Joint
- 1** **EUROPEAN OPEN SOURCE INITIATIVE** | Continued in the annual RedHat Research Day
- 3** **SOFTWARE COMPONENTS PUBLISHED** | with factsheets (proprietary)
- 1** **TECHNOLOGY TRANSFER** | through exploitation agreement for CEP component with LeanXcale
- 1** **SQL DATA SKIPPING BY IBM** | Published in the EC Innovation Radar
- 1** **BIGDATASTACK** | on the BDV Marketplace
- 3** **END-TO-END ONLINE DEMOS**
- 1** **BIGDATASTACK.EU** | Knowledge Platform
- 14** **PUBLIC PRIVATE PARTNERS**
- 1** **ADOPTION ROADMAP**
- 18** **SOFTWARE COMPONENTS PUBLISHED** | with factsheets and Open Source Codes
- 3** **PATENTES FILED** | by IBM, LXC and NEC, and one being prepared for filing
- 1** **COLLABORATION AGREEMENT** | with the EOSC DIH
- 4** **IBM SERVICES** | use SQL Data Skipping
- 12** **OPEN SOURCE UPSTREAM CONTRIBUTIONS**
- 4** **EXPLOITATION STORIES**
- 1** **BIGDATASTACK ZENODO COMMUNITY** | with 30+ resources

BigDataStack Impact in Numbers

- 3** USE CASES | In Insurance, Retail and Shipping
- 5** JOINT PARTICIPATIONS | in R&D projects to evolve our results
- 54** EVENTS ATTENDED BY BIGDATASTACK | Addressing Open Source, Industry, Research&Academia
- 1600+** VIEWS | of the BigDataStack software catalogue
- 4057** VIEWS | of BigDataStack OS components on GitLab
- 1** JOINT POLICY BRIEF
- 15** SCIENTIFIC PUBLICATIONS | and counting

- 2200+** SOCIAL MEDIA CONNECTIONS
- 1000+** PEOPLE | attended BigDataStack organised webinars
- 6** INFORMATION BOOTHS | for end-users in Open Source, Industry, Research & Academia
- 33K+** VIEWS | of the 3 BigDataStack use cases videos
- 784** VIEWS | of the BigDataStack OS components on GitHub
- 280+** TIMES | BigDataStack deliverables on Zenodo were consulted
- 2** PUBLICLY AVAILABLE EVENT REPORTS

Click the numbers to discover more



Interested in adopting the BigDataStack or its software components?

Contact us at info@bigdatastack.eu



Watch: BigDataStack in 3 words →

Check out Website & Social

- www.bigdatastack.eu
- [@BigDataStackEU](https://twitter.com/BigDataStackEU)
- [company/bigdatastack](https://www.linkedin.com/company/bigdatastack)
- [ZENODO Search: bigdatastack](https://zenodo.org/search?q=bigdatastack)



This project has received funding from the European Union's Horizon 2020 Research and Innovation program under grant agreement No 779747.