



BIG DATA EUROPE

Empowering Communities
with Data Technologies

6 September 2016

PRESS RELEASE

“Big Data Europe” addresses societal challenges with data technologies.

European project makes big data technologies easy to use.

Across society, from health to agriculture and transport, from energy to climate change and security, practitioners in every discipline recognise the potential of the enormous amounts of data being created every day. The challenge is to capture, manage and process that information to derive meaningful results and make a difference to people’s lives. The Big Data Europe project has just released the first public version of its open source platform designed to do just that. In 7 pilot studies, it is helping to solve societal challenges by putting cutting edge technology in the hands of experts in fields other than IT.

Although many crucial big data technologies are freely available as open source software, they are often difficult for non-experts to integrate and deploy. Big Data Europe solves that problem by providing a package that can readily be installed locally or at any scale in a cloud infrastructure by a systems administrator, and configured via a simple user interface. Tools like [Apache Hadoop](#), [Apache Spark](#), [Apache Flink](#) and many others can be instantiated easily.

As Maritina Stavrakaki of Agroknow explained, *“the availability of masses of ground based and satellite sensor data is changing my field of agriculture but it’s hard to combine it using traditional tools like spreadsheets and simple databases. The BDE platform allows me to process large amounts of data and process it in real time. Pattern recognition, accessing data tables from within publications and research across a wider range of datasets is now much easier than in the past.”*

The tools included in the platform were selected after a process of requirements-gathering across the seven societal challenges identified by the European Commission (Health, Food, Energy, Transport, Climate, Social Sciences and Security). Tasks like message passing are handled using [Kafka](#) and [Flume](#), storage by [Hive](#) and [Cassandra](#), or publishing through [geotriples](#). The platform uses the Docker system to make it easy to add new tools and, again, for them to operate at a scale limited only by the computing infrastructure.

“The platform makes it possible to carry out much smarter scheduling for a trip”, says transport expert Maxime Flament. “It’s not just published timetables that are important, but road maintenance schedules and data generated by vehicles in real time”.

Project Coordinator, Prof. Dr. Sören Auer of Fraunhofer IAIS, explained that in the second half of the project, the platform will be refined and expanded collaboratively with the users and the societal challenge partners (use case). The next step will be to improve data connectivity using graph technologies to introduce the concept of a semantic data lake, a semantic analytics stack engine and a semantic platform for logging and integration for cluster resilience.

The project has published a simple set of [instructions](#) for installing the platform, as well as a [video](#) on how to get started. Users can then pull ready-to-use images from the [BDE repository](#) or create a [dataflow pipeline](#) to realise a full data value chain using the components available.

One of the architects of the platform, Vrije University of Amsterdam’s Ronald Siebes concluded: *“We believe that the platform lowers the barrier to entry for new big data users and scientists from different domains, enabling them to experiment with a variety of big data tools in a plug and play fashion. It’s fun – try it!”*

About Big Data Europe

Big Data Europe is a project within the EU’s “Horizon 2020” framework programme that aims to build a knowledge- and innovation-based society so as to strengthen the competitiveness of Europe’s economy. Together with Fraunhofer IAIS, the following institutions and companies also participate in the project: AgroKnow (GR), CESSDA (NO), National Center for Scientific Research DEMOKRITOS (GR), ERCIM (FR), ERTIO European Union Satellite Centre (ES), Food and Agriculture Organization of the United Nations FAO (IT), Institute for Applied Informatics (InfAI) e.V. (DE), Kentro Ananeosimon Pigon Ke Exikononismis Energeias (Center for Renewable Energy Sources, (GR), National and Kapodistrian University of Athens (GR), The Open PHACTS Foundation (GB), Tenforce (BE), Semantic Web Company (AT), VU University (NL) and W3C (FR).

For more information, please visit: www.big-data-europe.eu

The platform can be downloaded from <https://github.com/big-data-europe>

See also the [installation instructions](#) and [video](#).

Press contacts:

Alexandra Garatzogianni
Fraunhofer IAIS

Phil Archer
W3C/ERCIM

Phone +49 2241 14-2974
Alexandra.Garatzogianni@iais.fraunhofer.de

Phone +44 7887 767755
phila@w3.org