YEARBOOK
2017
<table>
<thead>
<tr>
<th>AEGIS</th>
<th>BIG DATA OCEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIG DATA VALUE</td>
<td>DATABIO</td>
</tr>
<tr>
<td>DATAPITCH</td>
<td>E-SIDES</td>
</tr>
<tr>
<td>EUBUSINESSGRAPH</td>
<td>EW-SHOPP</td>
</tr>
<tr>
<td>EUBUSINESSGRAPH</td>
<td>EW-SHOPP</td>
</tr>
<tr>
<td>E-BUSINESSGRAPH</td>
<td>K-PLEX</td>
</tr>
<tr>
<td>EW-SHOPP</td>
<td>MY HEALTH MY DATA</td>
</tr>
<tr>
<td>EW-SHOPP</td>
<td>QROWN</td>
</tr>
<tr>
<td>K-PLEX</td>
<td>SLIPO</td>
</tr>
<tr>
<td>K-PLEX</td>
<td>SPECIAL</td>
</tr>
<tr>
<td>TRACK AND KNOW</td>
<td>TRANSFORMING TRANSPORT</td>
</tr>
<tr>
<td>TRACK AND KNOW</td>
<td>BIG DATA GRAPES</td>
</tr>
<tr>
<td>TRACK AND KNOW</td>
<td>BIG DATA STACK</td>
</tr>
<tr>
<td>TRACK AND KNOW</td>
<td>BIGMEDILY蒂CS</td>
</tr>
<tr>
<td>BODYPASS</td>
<td>BOOST 4.0</td>
</tr>
<tr>
<td>CLASS</td>
<td>CROSS-CPP</td>
</tr>
<tr>
<td>CROSS-CPP</td>
<td>DATABENCH</td>
</tr>
<tr>
<td>DATABENCH</td>
<td>E2DATA</td>
</tr>
<tr>
<td>E2DATA</td>
<td>EDI</td>
</tr>
<tr>
<td>EDI</td>
<td>FANDANGO</td>
</tr>
<tr>
<td>FANDANGO</td>
<td>FASHION BRAIN</td>
</tr>
<tr>
<td>FASHION BRAIN</td>
<td>I-BIDAAS</td>
</tr>
<tr>
<td>I-BIDAAS</td>
<td>ICARUS</td>
</tr>
<tr>
<td>ICARUS</td>
<td>LYNX</td>
</tr>
<tr>
<td>LYNX</td>
<td>SODA</td>
</tr>
<tr>
<td>SODA</td>
<td>THEYBUYFORYOU</td>
</tr>
<tr>
<td>THEYBUYFORYOU</td>
<td>TRACK AND KNOW</td>
</tr>
<tr>
<td>TRACK AND KNOW</td>
<td>TYPHON</td>
</tr>
</tbody>
</table>
Foreword

The European contractual Public Private Partnership on Big Data Value (BDV cPPP) was signed on 13 October 2014 marking the commitment of the European Commission, industry and partners from academia to build a data-driven economy across Europe, mastering the generation of value from Big Data and creating a significant competitive advantage for European industry, thus boosting economic growth and jobs. The Big Data Value Association (BDVA) is the private counterpart to the European Commission in implementing the BDV-cPPP programme.

The BDV cPPP activities address Big Data technology and application research and innovation development, new data-driven business models, data ecosystem support, data skills, regulatory and IPR environments and several social aspects.

The BDV cPPP started in 2015 and was operationalised with the launch of the Leadership in Enabling and Industrial Technologies (LEIT) work programme 2016/2017 of Horizon 2020 (H2020) with the first cPPP projects (Call 1) starting on January 2017.

2017 has been a year of major advancements for the Big Data Value cPPP. Funded projects started in 2017 ramping up from 0 to the 33 running projects at the beginning of 2018. During its first year of operations projects have reported over 40 innovations of exploitable value including technology, services, software, methods and tools, and including products and software components in the field of advanced privacy and security respecting solutions.

The PPP has already delivered services of high societal value and many are planned for 2018 onwards.

The BDV cPPP covers 4 major sectors with close to the market large-scale implementations (Bio Economy; Transport, Mobility and Logistics; Healthcare; Smart Manufacturing) and over 10 different sectors covered in total including Agriculture, Forestry and Fishing, Manufacturing, Wholesale, Retail and e-Commerce, Transport and Logistics, Information Service activities, Financial and Insurance activities, Public Administration, Health and Wellness, Safety and Security, Food, Tourism, Media and Energy. The BDV cPPP has developed 26 large Scale experiments (19 involving closed data) and over 150 use cases and experiments and made available 85,4 Petabytes of data for experimentation.

With an initial indicative budget from the Union of 534 M€ for the period 2016-2020, the BDV cPPP has mobilised 1,1 B€ private investments since the launch of the cPPP (484,6 M€ in 2017).

This Yearbook presents the different projects that have contributed to these numbers. One by one, each project is featured in terms of objectives, challenges and other related aspects. For those projects starting in 2018 a brief description is given and further information will be available in next editions of this document.

Please visit the BDV cPPP portal at www.big-data-value.eu for extended Information on our activities, projects, events and for access to our tools, our latest news and our Newsletter.

Enjoy your reading.
The Big Data Value Association (BDVA)

BDVA represents the private side of the BDV cPPP. The Big Data Value Association is an industry-driven and fully self-financed international non-for-profit organisation under Belgian law. BDVA has almost 200 members all over Europe with a well-balanced composition of large and small and medium-sized industries (over 30% of SMEs), as well as research and user organizations. BDVA members come together to collaborate in a common mission: developing the European Big Data Value Ecosystem that will enable the data-driven digital transformation in Europe delivering maximum economic and societal benefit, and, achieving and sustaining Europe’s leadership on Big Data Value creation and Artificial Intelligence.

To achieve this mission BDVA members develop the following activities:

- **Develop Data Innovation Recommendations**: Providing guidelines and recommendations on data innovation to the industry, researchers, markets and policy makers.
- **Develop Ecosystem**: Developing and strengthening the European Big Data Value Ecosystem.
- **Guiding Standards**: Driving Big Data standardisation and interoperability priorities and influencing standardisation bodies and industrial alliances.
- **Know-How and Skills**: Improving the adoption of Big Data through the exchange of knowledge, skills and best practices.

The cross-technological nature of the data value chains, flowing across different technologies (IoT, Cloud, 5G, Cybersecurity, infrastructures, HPC, ...) has triggered and accelerated during 2017 the development of stronger collaborations between the BDV cPPP/BDVA and other technological (cross-sectorial) sectorial communities and in particular with other PPPs. This has resulted in the identification of common challenges, alignment of roadmaps, and development of activities of common value. This work has been reflected in the latest multi-annual version of the BDV Roadmap as part of the SRIA v4.0 and materialised in the LEIT WP 2018-2020 with some calls of common interest focused on technology integration and the signature of official declarations of collaboration with other PPPs during 2018.

During 2017 the Public-Private collaboration consolidated its foundations, not only through contractual and governance structures but joining efforts in major dissemination and engagement activities. During 2017 the European Commission and the Association merged their main European Data Community events in a single new branded European large event: The European Big Data Value Forum (EBDVF) whose first edition took place in November 2017 in Versailles. The second edition will happen in November 2018 in Vienna, supported by the Austrian Presidency of the European Council.

Please visit www.bdva.eu for a full picture on the Association, its members and activities.
AEGIS is a 30-month project, co-funded by the European Commission that aims to bring stakeholders, data and technologies together in a big data platform, utilizing the latest advancements in the linked and big data landscape.

AEGIS aims towards a data-driven innovation by taking into consideration heterogeneous, structured, unstructured, streaming and batch datasets, in order to rejuvenate the existing big data sharing and analysis models and facilitate companies and organisations to provide better and personalized services to their users.

During the first year of the project, the focus was on the definition and development of the AEGIS platform, based on end-user requirements and driven by the AEGIS demonstrator needs. For the second year, the main objectives are the evaluation of first version of the AEGIS platform through our demonstrators on three diversified sectors: Smart Home and Assisted Living, Automotive and Insurance, and the development and role-out of the second version of the platform. Project results have been detailed in the publicly available deliverables but also in a series of research papers and presentations, which can be found through the AEGIS website.

Q. What opportunities do you envisage by the use of Big Data technologies?
A. Modern technological breakthroughs in the areas of the data-driven economy bring together data owners from various sectors and enable their interrelation and interoperation. Big Data infrastructures offer the opportunity to roll out completely new value chains of interrelated data streams coming from diverse sectors and languages.

Q. What makes Big Data platforms attractive for businesses in the safety and security domain?
A. Currently, there is a lack of data and knowledge sharing mechanisms, that in the case of safety issues, if properly exploited, will enable the creation of value chains towards more accurate risk models and proactive thinking. The abundance of data that are currently siloed, calls for smart collaborations to be formed essentially maximising the value offered to the end user.

Q. What will the AEGIS project offer to the Big Data market?
A. The AEGIS project, through its Big Data platform, aims to offer:
• Big data processing, enrichment, storage, analysis and sharing
• Cross-domain batch and streaming data integration and harmonization
• DCAT-AP conformed metadata
• Data anonymization and semantic enrichment procedures
Project description

The main objective of BigDataOcean is to enable maritime big data scenarios for EU-based companies, organisations and scientists, through a multi-segment platform that will combine data of different velocity, variety and volume under an inter-linked, trusted, multilingual engine to produce a big data repository of value and veracity back to the participants and local communities.

BigDataOcean aims to capitalise on existing modern technological breakthroughs in the areas of the big data driven economy and roll out a completely new value chain of interrelated data streams coming from diverse sectors and languages and residing on cross technology innovations being delivered in different formats.

In 2017 BigDataOcean delivered the first release of the platform, designed and implemented based on the user requirements extracted from workshops with maritime stakeholders dedicated to identifying end users’ needs. BigDataOcean features four pilots showcasing significant scientific and technological breakthroughs reported in publications at prestigious conferences. BigDataOcean partners will further disseminate the project’s outcomes in 2018 and organise workshops and hackathons that will bring ideas and insights to the project. The project’s goal is to revolutionise the way maritime-related industries work, showcasing a huge and realistic economic, societal and environmental impact that is being achieved by introducing an economy of knowledge into a traditional sector which does not operate in an orchestrated manner and is rather fragmented.

Scope of activity

- Call: H2020-ICT-14-2016-2017
- Big Data Value Chain Scope:
  - Data Acquisition
  - Data Pre-processing
  - Data Curation
  - Data Storage
  - Data Analysis & Visualization

Big Data infrastructure and data analytics will increase the efficiency and accuracy of the business.
Project description

Effectively combining in a consortium Large Enterprises, SMEs and Academia the Big Data Value eCosystem Project (BDVe) provides coordination and support for the current and future H2020 projects within the Big Data Value Public-Private Partnership.

BDVe directly interfaces with numerous stakeholders, both from inside and outside the PPP to foster a truly vibrant community around Big Data in Europe, including the alignment and collaboration with the two entities with representing the private and public sides of the PPP: The Big Data Value Association and the European Commission.

BDVe has established the common bodies for the discovering and exploitation of synergies at project management level, the generation of a complete Big Data Value Reference Model at a technical level, and to multiply impact into the different targets at communication level.

During 2017 BDVe has also taken the first steps paving the path for current and upcoming projects by developing the different assets and setting up the communication platforms needed to multiply the PPP reach.

BDVe is putting its marketing capabilities at the service of the whole Community to co-produce the European Big Data Value Forum and the BDV PPP Meet-up as a series of workshops on different topics facilitating the networking between the PPP projects and the BDVA Taskforces.

Scope of activity

- Call: H2020-ICT-17-2016-2017
- Big Data Value Chain Scope:
  - Data Acquisition
  - Data Pre-processing
  - Data Curation
  - Data Storage
  - Data Analysis & Visualization

Q. What is the challenge in coordinating such a large and diverse set of projects?
A. In fact, the challenge is the diversity and the size of the topic and of the PPP itself. More than 50 projects along 4 years should be supported, each one with its own objectives and areas of interest. The Steering Committee, the Technical Committee and the Communications team gathers respectively the project coordinators, technical leaders and communication experts to find common actions and directions that help to position Europe in a leading position in the Big Data Value global scenario.

Q. In terms of tools, what does BDVe provides for the benefit of the whole PPP?
A. We are working at three different levels: for the Companies, for their products and for the people. For the Companies the Big Data Landscape will include all the actors in the Ecosystem; for the products, the Big Data Marketplace will showcase the platforms and solutions coming from the different projects; and for the people we are building an ambitious skills platform, including the academic offering, a recognition programme, an internship offer and demand platform and a Mobility Programme for experts around Big Data in Europe.

Q. And about Communication and Dissemination, which are the major achievements so far?
A. The first year of the project has been devoted to the setup of quality platforms and procedures that guarantee a continuous service to a growing ecosystem. The website, intended as an always-on booth, is at the centre of all communications and always referred by other channels such as the Social Networks or the BDV Newsletter. Similarly, in terms of dissemination the European Big Data Value Forum, co-organized by BDVe, aims to be the Big Data flagship event in Europe the coming years, and will be complemented by the joint presence of the PPP in other major events in the coming years.
Project description

Data Pitch is a EU-funded open innovation programme that supports and develops an ecosystem of organizations building new services with shared data, from startups to large corporates.

Scope of activity

• Call: H2020-ICT-14-2016-2017
• Big Data Value Chain Scope:

Data Pitch is delivered by the University of Southampton, the Open Data Institute, both from the UK, Portuguese accelerator Beta-i and French company Dawex.

The Data Pitch innovation programme aims to develop a European data ecosystem, by connecting large-scale organization with data, with agile startups able to innovate with it.

Data Pitch launched in April 2017, since then the team have:
• Recruited six companies to share their data, including Deutche Bahn
• Launched the first open call with 12 data challenges, securing 142 applications
• Recruited the first cohort of 18 startups

The Data Pitch team has already reached businesses across Europe by:
• Building a Twitter community and setting up targeted advertising on Facebook.
• Attending events, including Pixels Camp Lisbon, Tech BBQ Copenhagen
• Developing news stories and securing media coverage including an Techcrunch, La Tribune
• Creating an influential stakeholder group
• Creating a Data Pitch newsletter

In 2018 we will support our 18 startups to excel and launch our second open call. At a more strategic level, we will be testing our open innovation model and feeding our learnings back into the programme.

Q. What has changed recently, around the idea of why and how we share data? Would you say that some people or organisations are more inclined now to share their own data and create Big Data pools?
A: With the implementation of the GDPR, data needs to be used in a controlled way that has value for and balances the rights of the person who is sharing it. Certain data needs to become more transparent, so that we are fully aware of and comfortable with sharing data, enabling companies to deliver a better, more personalised customer experience. This can create value for the data ecosystem, and ensure data hygiene allows organisations to make better, more informed decisions.

Q. What is distinctive between the Data Pitch model and other previous efforts with similar focus?
A. Data Pitch offers startups and SMEs the opportunity to solve a current and real problem within industry – helping them identify a market need and address it. Industries are dominated by large corporates and public sector organisations, but they are less inclined to be open to innovation, so by building collaborations between large corporates and startups, we are able to help create innovative solutions and solve real market issues through the process of sharing data.

Q. Which would be the main challenges and opportunities that the Big Data revolution is presenting, both for companies, public institutions and people in Europe?
A. Over the past two decades, since the adoption and evolution of the Internet, we have seen a large growth of big data. However, this growth has its limitations. The technology, tools and processes for using, storing and collecting data need to be improved, as much of the data that is collected is underused. When organisations adopt an open mentality to sharing data, startups can help them improve tools and understand the value within their data through innovation. This enables businesses to make better informed decisions, based on the quality and volume of the data collected and processed and in turn enable job creation through the need to create innovative solutions. By creating jobs, and advocating data transparency, we are able to help boost the EU data economy.
## Project description

DataBio is a H2020 Lighthouse project focusing on utilizing Big Data technologies (software components and datasets) to sustainably and responsibly improve the productivity in the main bioeconomy sectors. It deploys over 90 state-of-the-art Big Data, Earth Observation and ICT technologies, and partners' existing infrastructure and solutions, linked together through the DataBio Platform. The project is driven by the co-design, implementation and evaluation of 26 demonstrative piloting trials in agriculture (13), forestry (7) and fishery (6). DataBio runs supporting dissemination and training effort to boost technological take-up, and emergence of new solutions and business opportunities.

During the DataBio’s lifecycle, big data (multi-terabyte or larger) will be collected in a wide range of data types from numerous conventional and new sources. The project provides a big data toolset to support bioeconomy pilots and the forming of reusable and deployable pipelines of interoperable components (mostly provided by partners) thus extending the impact of DataBio to new bioeconomy projects as well as to other business areas.

### Scope of activity

- **Call:** H2020-ICT-15-2016-2017
- **Big Data Value Chain Scope:**

- **Data Acquisition**
- **Data Pre-processing**
- **Data Curation**
- **Data Storage**
- **Data Analysis & Visualization**

Q. What are the main challenges and opportunities that the use of Big Data brings to agriculture, forestry and fishery sectors?

A. Big data and big data technologies will benefit entire key social economic sectors including the primary, secondary, and the tertiary. It is acknowledged big share will go to the big data's analytic area but, big data beneficiaries include also data sets, business value, and sales and marketing activities. Big data's volume, velocity and diversity will lead to broader data sourcing for analytics, and hence improved data staging and warehousing. On the other hand, improved datasets are going to enhance harnessing the information and growing new data-driven applications. Several challenges face big data as a new technology paradigm. The most critical challenges include inadequate staffing or skills, weak business support, inadequate data management infrastructure, and immaturity with big data properties.

Q. How is DataBio contributing to the Big Data Technology?

A. Big data technologies are being deployed on big datasets that are being extensively produced in pilots run in the three bioeconomy sectors: agriculture, forestry, and fishery. The project deploys over 90 state-of-the-art Big Data, Earth Observation and ICT technologies as well as partners' existing infrastructure and solutions, linked together through the DataBio Platform. On the other hand, DataBio is driven by the design, development, implementation and evaluation of 26 demonstrative piloting trials covering agriculture (13), forestry (7) and fishery (6). The main objective of the project is a significant contribution to increased productivity in the three European economic sectors of interest.

Q. Which are the methods and technologies used on this project?

A. Big information is being derived from a variety of sources including proximal and satellite sensors, and from production operations and several databases. The aim is to provide a streamlined, Big Data Infrastructure for data discovery, retrieval and processing related to Agriculture, Forestry and Fishery, engaging global communities of data providers and users, promoting pilot activities in agriculture, forestry and fishery towards innovative Big Data powered applications using DataBio infrastructure.
Project description

Although there are potentially enormous economic and social benefits, data-driven innovation also brings new challenges for individual and collective privacy, security, as well as democracy and participation. The main objective of e-SIDES is to complement the research on privacy-preserving big data technologies, by analyzing the challenges emerging from the adoption of big data technologies, conforming to the principles of Responsible Research and Innovation. e-SIDES is investigating stakeholders' concerns and framing the results in a conceptual framework showing the potential trade-offs between conflicting needs and providing a basis to validate privacy-preserving technologies.

Scope of activity

• Call: H2020-ICT-18-2016
• Big Data Value Chain Scope:

  Data Acquisition  Data Pre-processing  Data Curation  Data Storage  Data Analysis & Visualization

This may raise privacy issues, but also (other) ethical and societal issues, such as discrimination, human dignity, justice, fairness and trust. In view of such considerations, businesses and governments are often unsure about how to deal with the data collected through their operations while developers are unsure how to avoid introducing unfair or unethical snares in their code. e-SIDES will help companies, researchers, policy makers and civil society identify these pitfalls as well as strategies and recommendations to develop and promote technologies that are inherently privacy-preserving.

Q. Can technology guarantee the anonymisation of personal data without losing the value added of analytics?
A. Traditional anonymisation techniques fail in the context of big data applications because there are hundreds of data points for every single individual. A full de-identification cannot be achieved, but a re-identification can be made costly. There are different approaches that may be used when anonymising data such as k-anonymity, which seeks to prevent record re-identification by hiding each original record within a group of k indistinguishable anonymised records, and differential privacy, which seeks to limit the knowledge that users obtain from query responses. However extremely large datasets may reduce the privacy of individuals, and some approaches may reduce the accuracy of analytics. This may lead to losses in value added.

Q. What is the best approach to design privacy-aware solutions and services without falling into ethical traps, such as fostering discrimination and unfairness?
A. As there are far reaching, often unforeseen implications of big data applications, it is important to conduct a careful analyses of applications as well as the big data technologies that are being developed or used in order to evaluate which ethical values may be affected. Ideally, these considerations are taken into account not only in advance of using a privacy by design approach to development, but also throughout the whole development and use process. Responses need to be proactive to address threats to ethical values when they arise. Ethical considerations must be revisited and re-evaluated periodically, especially when new technologies are added or used in new ways, or new or more complete datasets are utilised.

Q. Can we move from technology as the problem (violating privacy) to technology as the solution?
A. Big data analytics can have negative implications, which are not limited to privacy issues. The implications may also include, for instance, discrimination or unfairness. Additionally, massive data breaches are occurring with alarming regularity. Measures taken by both policy makers as well as business leaders dealing with large datasets do not appear to be effective. It is therefore not surprising that big data technology itself is sometimes seen as the root of many problems. However, it is the way society uses data and technology that determines whether the overall impact of big data applications is positive or negative. The ability to reliably mitigate the negative implications through a combination of technological and non-technological measures may lead to a change of perspective.
Project description

euBusinessGraph aims to create a knowledge graph of companies, and a set of innovative business products and services build in connection to the knowledge graph.

Scope of activity

• Call: H2020-ICT-14-2016-2017
• Big Data Value Chain Scope:
  - Data Acquisition
  - Data Pre-processing
  - Data Curation
  - Data Storage
  - Data Analysis & Visualization

EuBusinessGraph is a 2.5 years lasting project that is highly product-oriented and cross-border-thinking. Apart from project leader SINTEF (Norway), the consortium consists of nine partners from different countries and backgrounds: research and public institutions as well as businesses.

The main object of euBusinessGraph is to create a cross-border knowledge graph of companies, and a set of innovative business products and services build in connection to the graph. The graph aims to enable the creation of data-driven products and services starting with a set of six corresponding business cases. The business cases are developing well and the overall graph is shaping up on its way to deliver our vision: world-class company datasets and analytics across borders and languages.

EuBusinessGraph collaborates with other projects, such as EW-Shopp. The consortium attended several (networking) events and will continue so. In Year 1 we focussed on establishing visibility in corresponding communities; for the remaining project period we aim to strengthen collaborations and finding synergies with other projects. On a dissemination level we will focus on promoting the progress of our business cases and the knowledge graph.

Q. Which role do Private Public Partnerships have in the European Big Data community?
A. PPPs are essential, especially when it comes to community building, networking and funding.

Q. In which direction is the commercialisation of business data heading?
A. Overall the market is growing and the entrance barriers are getting lower. We observe that a transparent and cost-effective marketplace for company-related data is essential.

Q. From a technical aspect, where do you see promising innovations within the scope of the project?
A. We see a huge potential in the euBusinessGraph data marketplace, where we focus on company-related data. The marketplace is supported by a self-service platform for data onboarding for data providers.

Q. Your project has been directly collaborating from the start with other ones within the PPP. Do you felt that collaboration will increase following the creation of an European Big Data Ecosystem?
A. We are convinced that collaboration with other related projects helps us in both promoting the idea of European Business Graph and shaping the profile of what services we are going to offer. An European Big Data Ecosystem would even more enable opportunities of mutual exchange.
Project description

EW-Shopp aims at helping companies, operating in the eCommerce, Retail and Marketing industries, embed events and weather in their daily data analyses so as to gain deep customers’ insights and optimize the services offered to their customers.

**Scope of activity**

- Call: H2020-ICT-14-2016-2017
- Big Data Value Chain Scope:

EW-Shopp aims at supporting companies to gain customer insights and build innovative data-driven services for their customers. Such insights come from the increased capacity of factoring events and weather into analysis of companies’ data. The project will deliver a platform to facilitate the enrichment of companies’ data, and their visual and predictive analysis.

The consortium includes companies (GfK Italy and Engineering Ingegneria Informatica from Italy; Ceneje, Big Bang from Slovenia; Measurence from Ireland; JOT from the UK; and research organizations (University of Milano-Bicocca, Jožef Stefan Institute, and SINTEF).

In 2017, we cleared out ethics-related issues, and collected business requirements. We addressed interoperability by specifying formats of and providing access to weather, event, spatial, temporal, and product data. Finally we designed and released a preliminary version of the EW-Shopp platform.

In 2018, we plan to release six pilot services in four business cases, and finalize the key technical features of the platform.

We have been sharing insights about the project on our Data Blog, online industry magazines, key dissemination events, and social media, in order to engage with target users and build up our project’s brand awareness.

Preliminary data analyses confirmed that Digital services should be optimized using event and weather-based analytics because context does matter for customers. We opened up a discussion group on LinkedIn (Weather & Event-based Analytics in Business) to allow private and public bodies share challenges or solutions.

**Supporting event and weather-based data analytics and marketing along the shopper journey**

Q. You are targeting a specific business domain, linking the beginning and end of its value chain and connecting it to the weather information. How is Big Data at the center of all that?

A. We rather say that we target the eCommerce, Retail, IoT and Digital Marketing business domains. A more precise formulation could be that we target an ecosystem that touches and interconnects several business domains. We then connect this ecosystem to information about weather and events to generate advance and value-added information to improve business KPIs. Big Data is at the center of our attempt because of three main reasons. First, we target the Variety dimension. Second, in one of our domains, Digital Marketing, we target the Volume dimension of Big Data, preparing and enriching historical data. Third, technologies used in the project to allow users to integrate, enrich and analyze a vast amount of heterogeneous data are based on the best practices of Big Data architecture and software integration.

Q. Which would be the main opportunities that the use of Big Data technologies is offering to the retail and marketing sectors as they are joining the global digitalization processes?

A. The main opportunities offered by Big Data technologies to these sectors consists in enabling companies in these domains to gain in-depth insights on customer behavior, which can then be used to develop targeted services and to optimize sales, interactions and margins. In addition, since EW-Shopp vocation is all about knowledge and technology transfer, we can say that business partners have also enriched their knowledge about these technologies.

Q. Which would be the ethics-related questions that should be addressed by any project willing to promote the use of Big Data?

A. We believe that an important ethic-related question in our domains is how to find a good trade-off between the capability of acquire actionable customer insights that companies can benefit from, and the respect of users’ privacy. In our project, for example, we focus on insights that can be gained using aggregate data enriched with non-privacy-sensitive third-party data, as opposed to extreme personalisation strategies that may require deep and invasive user models. All the actions executed in this project are in fact compliant with the actual GDPR regulation.
**Project description**

Humanities-led and ICT-funded, the KPLEX project is investigating the delimiting effect digital mediation and datafication may have on rich, complex cultural data in a comparative, multidisciplinary and multisectoral fashion.

The KPLEX partners have been funded to undertake a 15-month investigation of the ways in which a focus on ‘big data’ in ICT research elides important issues about the information environment in which we live. The project will be delivered by the following 4 organisations:

- Trinity College Dublin (TCD)
- Data Archiving and Networked Services (DANS)
- Freie Universität Berlin (DE)
- Tilde (LV)

The team is committed to sharing results widely to inform both industry regulation and philosophical debates in the technical sphere.

**Scope of activity**

- Call: H2020-ICT-35-2016
- Big Data Value Chain Scope:
  - Data Acquisition
  - Data Pre-processing
  - Data Curation
  - Data Storage
  - Data Analysis & Visualization

In addition to consistently working towards achieving the individual objectives as set forth in the project Grant Agreement, the project partners have actively pursued an ambitious dissemination programme. Particularly noteworthy examples are Dr. Georgina Nugent-Folan’s feature piece in ERCIM News 111 and Dr. Jennifer Edmond’s feature piece on the RTÉ Brainstorm website which has attracted significant attention from within the private ICT sector.

The project has completed its work on 31st of March and the main activities and goals of 2018 will focus on ensuring the timely and thorough completion of the project deliverables.

Q. How would you succinctly describe the “delimiting effect” that datafication has in an area as vast as the one you are targeting?

A. The digital record can suffer from impoverishment due to what can be captured explicitly and effectively. As one researcher in the KPLEX interviews described it, “all this documentation stuff functions as a kind of exogram or external memory storage... the sensual qualities of field notes, photographs or objects from the field have the capacity to trigger implicit memories or the hidden, embodied knowledge.” If we are concerned about the development of pan-European identities, for example, and of the strength of cultural ties able to create resilient societies, then we should be very concerned about how the digital record, for all of its global reach and coverage, represents cultures and languages unequally.

Q. Your project may create new links between the academic community and the European Data-centered business ecosystem. How do you plan to achieve that?

A. The datafication of research requires us to rethink how we create the problem-solving toolkits with which we equip people at every level of society. Life in the 21st century requires a range of problem solving approaches. While it is important, perhaps, that experts in culture learn to code, and engineers and computer scientists develop their ethical sensibilities, narrative imaginations, cultural competencies, and sensitivities to the communications skills. Understanding the limits and potential of ICT must become a foundational skill, regardless of the context in which it will be applied. Universities, active researchers, and professional societies would all have a role to play in this transformation, as would the publishers and funders that manage the incentive systems at the top of the research ecosystem.

As a ‘sister’ project intended to undertake research linked to other Horizon 2020-funded big data research areas, KPLEX itself did not have either the time or resources to fully develop the many potential interventions. We can hope, however, that future research and policy development will encourage the big data research community as a whole to take these opportunities to rethink fundamental assumptions and foster a more symbiotic relationship between technological and social progress. This will be a necessary development for Europe if the recognised risks of big data research are to be countered at the macro level.

Q. What opportunities do you envisage using Big Data technologies for the European population?

A. Big data research should be a means, but not an end. While computer or data scientists may be able to extract a certain kind of knowledge from large data sets, by their very nature the original sources contain more complexity than those results necessarily represent. Decision-making in big data research should not be driven by perceived technical imperatives to meet an algorithmic challenge or commercial imperatives to serve a market niche, but must also contain a natural braking function to ensure that the technical and the commercial don't outstrip the human and the social.
MyHealthMyData aims at facilitating sensitive data sharing. MHMD is poised to be the first blockchain-based biomedical information network centred on the connection between healthcare providers, biomedical industries and individuals.

Scope of activity

MHMD is composed of SMEs, research centers, universities, a big industry, a law firm, and four hospitals, distributed over 9 EU countries.

MyHealthMyData leverages the blockchain technology allowing hospitals to share privacy protected data, while prompting citizens to regain control over their health data.

During 2017, MHMD designed the blockchain backbone, the smart contracts-enabled consent management tool, the data catalogue, and the user interfaces.

MHMD attended a large number of dissemination events. MHMD will also have a booth at the forthcoming EU Digital Week in Sofia (June 25th-26th 2018).

During 2018, MHMD will mainly focus on: privacy-preserving and security tools for data sharing and analytics, smart contracts for data access and exchange, and mobile/web applications. MHMD is intended to become a multi-sided health data marketplace, based on new mechanisms of trust and direct, value-based relationships between EU citizens, hospitals, research centers and businesses.

Q. What are the main opportunities that the use Big Data is bringing to the health sector?
A: Big data is allowing new diagnostic and treatment models to be designed in healthcare systems worldwide. The availability of diverse, multilevel and high volume data set enables personalization of treatments, predictive care pathways which can improve clinical outcomes at scale, making the entire care process tailored to the specific needs of the patient.

Q. Privacy of data have always been at the center of any digitalization trend around health and bio-medical information. Which would be the point the security and sharing meet?
A: Several strategies have been proposed to increase the sharing of data available for clinical and research purposes, while preserving privacy and security. MHMD is targeting advanced technologies for making that possible and for leveraging the value data in a variety of operational contexts. On one hand, the application of secure multi-party computation and homomorphic encryption can greatly facilitate the sharing of data by guaranteeing security and privacy. On the other hand, the approach of bringing the “algorithms to the data”, thus minimizing data mobilization is increasingly becoming the solution of choice for secure data sharing.

Q. Which kind of new data is getting incorporated to health data pools? Who and how may profit from it?
A: The new phenomenon of the Medical Internet of Things (MIoT) is bringing to healthcare an incredible amount of citizens/patients-generated health data, by wearable devices and mobile applications (it’s forecasted that 161 million healthcare IoT devices will be shipped worldwide in 2020). Healthcare providers will benefit from this new data flow for improved management of chronic diseases and use of resources through teledmedicine services (and subsequent reduction of – e.g. – visit at the GP or at the emergency rooms); pharma companies are becoming able to better understand adherence to therapy, effectiveness of drugs in specific sub-sets of patients, also improving their post-market adverse effect control and drugs usage monitoring. Patients and citizens will benefit from this new data flow thanks to improved and more personalized services.

Data Acquisition
Data Pre-processing
Data Curation
Data Storage
Data Analysis & Visualization

MHMD

- Call: H2020-ICT-18-2016
- Big Data Value Chain Scope:
Project description

QROWD offers local governments and transportation businesses innovative solutions to improve mobility, reduce traffic congestion and make navigation safer and more efficient. We involve the people and capitalize on hybrid Big Data integration.

Scope of activity

• Call: H2020-ICT-14-2016-2017
• Big Data Value Chain Scope:

Data Acquisition → Data Pre-processing → Data Curation → Data Storage → Data Analysis & Visualization

The goal of the project is ultimately to make transport in its various forms smarter.

QROWD project will advance in the development of the QROWD platform, which will allow various data analysis processes. Three core use-cases have been identified as valuable for citizens, municipalities and industry.

1. Completing mobility infrastructure
   This aim of this is to gain deeper knowledge about existing infrastructure (type and position) based on crowdsourced data from citizens (app-based sensors and manual feedback).

2. Better data for computing and estimating the modal split of the use cases to inform transport policies.
   Most service providers that are working in this space, including the likes of Google maps, are not very accurate in doing the modal split. Understanding what transportation services citizens are using in their daily routines will save transport cost and driving time on a daily basis (together with fuel, personnel stress...) and will get improved customers flows due to less road congestion issues. On the other hand, data driven services will benefit from increasing demand for info to support mobility related decision making coming from both public and private sector customers.

3. Decision support through the creation of the dashboard to allows citizens and municipalities to be more informed about transport in the city.
   This is about analyzing different aspects of transport and presenting them visually so that they are useful for individuals and communities: to be of real value to cities citizens should easily and quickly see the value in it.
SLIPO develops technologies for the scalable and quality-assured integration of Points of Interest (POI) Big Data assets using Linked Data technologies. We address the challenges of integrating, enriching, and sharing POIs, and deliver POIs of greater coverage, accuracy, and value.

Scope of activity

SLIPO aims to address the limitations, gaps, and challenges of current POI data integration. Within its first year, the development of the SLIPO Toolkit, a cloud-based software-as-a-service for large-scale data integration, has exceeded our expectations, increasing scalability and performance by one order of magnitude. SLIPO simplifies and speeds up the data integration process, reducing the associated costs and complexities, and delivering POIs with increased coverage, quality, timeliness, and completeness.

Q. What new opportunities and challenges do you envisage by the use of Big Data technologies for innovation experts and industrial stakeholders?
A: Generating value from Big Data is a reality for most aspects of our Data Economy, leading to increases in competitiveness and growth. We are past the point of pondering the potential impact of Big Data; it has already arrived! New systems, algorithms, and concepts successfully address the ever-growing volumes of data, and gradually find their way across most domains of science and economy.

SLIPO contributes in this goal by making it simpler, faster, cheaper, and more efficient for industrial stakeholders to extract value for POI Big Data assets.

Q. And how that process will be benefiting the people in Europe?
A: The emerging convergence of science and economy domains towards data-driven innovation, will safeguard our financial growth, but most significantly, ensure our security and quality of life. Big Data cannot address all challenges, and is only part of the solution, but Big Data has a transformative effect for EU society, which will inclusively benefit all EU citizens.

Q. Who are the prime users of the SLIPO Toolkit and what kind of feedback have you received from them?
A: SLIPO targets SMEs that produce or consume POIs regardless of their specific domain and Big Data expertise. We help them increase the value of their data, lower production costs, shorten update times, increase quality, and expand to new markets by ensuring data availability. Early user feedback is extremely positive, with users highlighting how SLIPO makes POI data integration simple, fast, and effective.

POI data integration is technically challenging and labor-intensive, resulting to large fixed update cycles for commercial data and lost revenue opportunities. SLIPO simplifies and speeds up the data integration process, reducing the associated costs and complexities, and delivering POIs with increased coverage, quality, timeliness, and completeness.

SLIPO COMPONENTS

- Transform
- Interlink
- Enrich
- Fuse

TOOLKIT

SLIPO WORKSPACE

SLIPO API

SLIPO ANALYTICS
Project description

SPECIAL is a H2020-ICT-2016-1 Big Data PPP funded R&D project focused on developing technologies supporting compliance with the European data protection framework. Our scientific results are accessible here: www.specialprivacy.eu/publications/public-deliverables

Scope of activity

SPECIAL focuses on solutions for Big Data industries which need to comply with the European General Data Protection Regulation (GDPR). We develop technology that enables user consent management, policy enforcement, transparency, control features, and auditability. These technologies will be tested with a number of use cases reflecting real-world business scenarios in the context of Big Data environments. In 2017, requirements from different relevant domains were identified, such as legal, ethical, economical and technical requirements.
**Project description**

Big Data will have a profound economic and societal impact in the mobility and logistics sector, which is one of the most-used industries in the world contributing to approximately 15% of GDP. With freight transport activities projected to increase by 40% in 2030, transforming the current mobility and logistics processes to become significantly more efficient, will have a profound impact. However, big data is still under-used in the domain. TransformingTransport aims at improving efficiency in the mobility and logistics sectors, which may lead to significant cost savings, through the increase in investing in big data utilization, demonstrating its results in all transport modes.

**Scope of activity**

• Call: H2020-ICT-15-2016-2017  
• Big Data Value Chain Scope:

The TransformingTransport (TT) project demonstrates, in a realistic, measurable, and replicable way the transformations that Big Data will bring to the mobility and logistics market. To this end, the project validates the technical and economic viability of Big Data to reshape transport processes and services to significantly increase operational efficiency, deliver improved customer experience, and foster new business models.

TT addresses seven pilot domains of major importance for the mobility and logistics sector in Europe: 1. Smart Highways, 2. Sustainable Vehicle Fleets, 3. Proactive Rail Infrastructures, 4. Ports as Intelligent Logistics Hubs, 5. Efficient Air Transport, 6. Multi-modal Urban Mobility, 7. Dynamic Supply Chains. Initial results from those pilot domains have been disseminated in various events within 2017 (e.g. BDVF, etc.), whereas 2018 started with a very successful project presence at TRA 2018 and is planned to continue with further presentations in events, as well as partners publications in high quality journals and magazines.

The TT consortium combines knowledge and solutions of major European ICT and Big Data technology providers together with the competence and experience of key European industry players in the mobility and logistics domain.
**BIG DATA GRAPES**  
**GA Number:** 780751  
**Call:** H2020-ICT-16-2017

BigDataGrapes aims to help European companies in the wine and natural cosmetics industries become more competitive in the international markets. It specifically tries to help companies across the grapevine-powered value chain ride the big data wave, supporting business decisions with real time and cross-stream analysis of very large, diverse and multimodal data sources.

[www.bigdatagrapes.eu](http://www.bigdatagrapes.eu)

**BIG DATA STACK**  
**GA Number:** 779747  
**Call:** H2020-ICT-16-2017

The world is moving to Big Data. As much as Cloud Computing was a breakthrough technology, cloud management and provisioning has mainly be aimed at optimizing computation, with less of a consideration for efficient handling of Big Data applications. BigDataStack is a data-centric view of the Cloud and will provide a complete infrastructure management system that will base the management and deployment decisions on data aspects thus being fully scalable, runtime adaptable and high-performing for big data operations and data-intensive applications in the Cloud.

[www.bigdatastack.eu](http://www.bigdatastack.eu)

**BODYPASS**  
**GA Number:** 779780  
**Call:** H2020-ICT-14-2016-2017

BodyPass ambition is to create a platform to exchange 3D data from different sectors as healthcare and consumer goods sectors. BODYPASS will offer tools that open new opportunities for solving business and social challenges. It will help the European Consumer Goods’ Industries to improve product design, to enable mass customization and to increase online sales; to the stakeholders of the Image Diagnostics Market for accessing 3D data to enable new diagnosis services and to improve the patient’s journey in the health sector.

[www.bodypass.eu](http://www.bodypass.eu)

**BOOST 4.0**  
**GA Number:** 780732  
**Call:** H2020-ICT-15-2016-2017

The biggest European initiative in Big Data for Industry 4.0, will lead the construction of the European Industrial Data Space to improve the competitiveness of Industry 4.0 and will guide the European manufacturing industry in the introduction of Big Data in the factory, providing the industrial sector with the necessary tools to obtain the maximum benefit of Big Data.

[www.boost40.eu](http://www.boost40.eu)

**BIGMEDILYTICS**  
**GA Number:** 780495  
**Call:** H2020-ICT-15-2016-2017

BigMedilytics will transform Europe's Healthcare sector by using state-of-the-art Big Data technologies to achieve breakthrough productivity in the sector by reducing cost, improving patient outcomes and delivering better access to healthcare facilities simultaneously.

[www.bigmedilytics.eu](http://www.bigmedilytics.eu)

**CLASS**  
**GA Number:** 780622  
**Call:** H2020-ICT-16-2017

CLASS aims to develop a novel software architecture framework to help big data developers to efficiently distributing data analytics workloads along the compute continuum (from edge to cloud) in a complete and transparent way, while providing sound real-time guarantees. This ability opens the door to the use of big data into critical real-time systems, providing to them superior data analytics capabilities to implement more intelligent and autonomous control applications.

[www.class-project.eu](http://www.class-project.eu)
Cross-CPP
GA Number: 780167
Call: H2020-ICT-14-2016-2017
The objective of Cross-CPP is to establish an IT environment offering data streams coming from various industrial sectors (vehicle, smart home devices, etc.). Providing access to these Cross Industrial Data streams will enable to build new and innovative business ideas for many stakeholders.

www.cross-cpp.eu

DATABENCH
GA Number: 780966
Call: H2020-ICT-17-2016-2017
Organisations rely on evidence from the Benchmarking domain to provide answers to how their processes are performing. There is extensive information on how and why to perform technical benchmarks for the specific management and analytics processes, but there is a lack of objective, evidence-based methods to measure the correlation between Big Data Technology (BDT) benchmarks and an organisation's business benchmarks and demonstrate return on investment.

www.databench.eu

E2DATA
GA Number: 780245
Call: H2020-ICT-16-2017
E2Data proposes an end-to-end solution for Big Data deployments that will fully exploit and advance the state-of-the-art in infrastructure services by delivering a performance increase of up to 10x while utilizing up to 50% fewer cloud resources.

www.e2data.eu

EDI
GA Number: 779790
Call: H2020-ICT-14-2016-2017
EDI provides an incubation programme and up to €100k equity free for EU Big Data startups sorting out data challenges from companies in a myriad of sectors. EDI will help Big Data entrepreneurs with the following: a free infrastructure with open source tools; training on the most known off-the-shelf solutions; support and business services to develop their business idea; and equity-free funding.

www.edincubator.eu

FANDANGO
GA Number: 780355
Call: H2020-ICT-14-2016-2017
The FANDANGO project aims at providing unified techniques and an integrated big data platform to support traditional media industries to face the new “data” economy with a better transparency to the citizens under a Responsible, Research and Innovation (RRI) prism. Hence, the goal of FANDANGO is to aggregate and verify different typologies of news data, media sources, social media, open data, so as to detect fake news and provide a more efficient and verified communication for all European citizens.

www.fandango-project.eu

FASHION BRAIN
GA Number: 732328
Call: H2020-ICT-14-2016-2017
Combining data from different sources to support different fashion industry players by predicting upcoming fashion trends from social media as well as by providing personalized recommendations and advanced fashion item search to customers.

www.fashionbrain-project.eu
I-BiDaaS
GA Number: 780787  Call: H2020-ICT-16-2017
I-BiDaaS is a self-service solution, aiming to empower users to easily utilize and interact with big data technologies by designing, building and demonstrating a unified framework that significantly increases the speed of data analysis while coping with the rate of data asset growth and facilitates cross-domain data-flow towards a thriving data-driven EU economy.

www.ibidaas.eu

ICARUS
GA Number: 780792  Call: H2020-ICT-14-2016-2017
ICARUS aims to deliver a novel framework and architecture that leverages data, primary or secondarily related to the aviation domain, coming from diverse sources to help companies and organisations whose operations are directly or indirectly linked to aviation. Thus, simultaneously enhance their data reach, as well as share and trade their existing data sources and intelligence, in order to gain better insights into airplanes’, airports’ and passengers’ quantified selves and contribute to improving their operations whether in real time or “offline” and increasing passengers’ safety and satisfaction.

www.icarus2020.aero

LYNX
GA Number: 780602  Call: H2020-ICT-14-2016-2017
Lynx will provide more effective ways of accessing huge amounts of digital regulatory compliance documents, including legislation, case law, standards, industry norms and best practices. In particular, this solution envisages an ecosystem of smart cloud services to better manage compliance documents, based on a Legal Knowledge Graph, which integrates and links heterogeneous compliance data sources.

www.lynx-project.eu

SODA
GA Number: 731583  Call: H2020-ICT-18-2016
Scalable Oblivious Data Analytics. Development of robust technologies, including a control dashboard to be applied in the acquisition of user consents and workflows of data, addressing the contradiction between Big Data Innovation and privacy-aware Data policies.

www.soda-project.eu

THEYBUYFORYOU
GA Number: 780247  Call: H2020-ICT-14-2016-2017
TheyBuyForYou will explore how procurement knowledge graphs, paired with data management, analytics and interaction design could be used to reform four key procurement areas: economic development, demand management, competitive markets and supplier intelligence.

www.theybuyforyou.eu

TRACK AND KNOW
GA Number: 780754  Call: H2020-ICT-16-2016
Track & Know will research, develop and exploit a new software framework that aims at increasing the efficiency of Big Data. This will be applied in the transport, mobility, motor insurance and health sectors.

www.trackandknowproject.eu

TYPHON
GA Number: 780251  Call: H2020-ICT-16-2017
The aim of TYPHON is to provide an industry-validated methodology and integrated technical offering for designing, developing, querying, evolving, analysing and monitoring scalable hybrid data persistence architectures that will meet the growing scalability and heterogeneity requirements of the European industry, which drives the project through four use cases from the domains of automotive, earth observation, banking, and motorway operation.

www.typhon-project.org