Dealing with software engineering challenges during the I-BiDaaS baseline phase

Dr. Evangelia Kavakli - April 30, 2019

I-BiDaaS aims to develop a unified Big Data as-a-service solution, which will be easily configured and adopted and will empower Big Data users to easily utilize and interact with Big Data technologies. I-BiDaaS end-users include non-technical business users, who only consume the analytics results, data scientists that configure analytic services and data flows, as well as data integrators, system administrators and analytics application developers.

The project is driven by a number of real use case scenarios within the three industrial sectors where the project's partners are involved, namely: telecommunications, finance and manufacturing (reported in D1.1). Although company-specific, these scenarios cover generic objectives and deploy Big Data analytics in order to address problems that cross sectorial borders (summarized in Table 1).

Table 1. I-BiDaaS use cases

Industrial Sectors	Business Objectives	Big Data Deployment
– **Telecommunications** – **Finance** – **Manufacturing**	– improve and optimise current operations – improving decision making – improve efficiency of Big Data solutions	– customer analysis – data driven services – fraud detection – operational efficiency – predictive maintenance

Given the broad scope of business processes, technologies and types of expertise involved, the software engineering challenges faced by the I-BiDaaS consortium during the baseline phase of the project were: the requirements engineering challenge: "which are the business, user and technical requirements underpinning all use cases?"; the system architecture challenge: "which are the system functional components that collaboratively satisfy these requirements?"; the quality challenge: "what are the

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

quality indicators, benchmarking tools and testing processes for evaluating the performance of the I-BiDaaS solution and its applicability to the different I-BiDaaS use cases?".

To address the first challenge, a Big Data requirements typology was defined providing a uniform representation of all key requirement concepts, thus facilitating the identification of cross-sectorial industrial challenges. Using this typology a systematic analysis and consolidation of end-user requirements were carried out, leading to a generalized set of requirements that are industry-neutral and technology-agnostic (reported in D1.3).

Towards the second challenge, the defined set of requirements were used as the key input of the I-BiDaaS software architecture design, which forms the reference model of the integrated I-BiDaaS solution. In order to deal with the system complexity a viewpoint-oriented approach was used (reported in D1.2), whereby the system is seen as: a set of implementation units (logical view); a set of runtime elements interacting to carry out the systems work (informational view); and a set of elements existing in and relating to external structures in its environment (physical view).

Finally, to address the quality challenge the industrial requirements and architecture specification described above, guided the definition of the I-BiDaaS experimental protocol (reported in D1.3), consisting of two phases. The first focuses on the evaluation of the I-BiDaaS architecture and its individual components using applicable Big Data benchmarks, while the second phase focuses on the evaluation of the application of the I-BiDaaS solution in the different use cases in the business context and against the identified high-level business requirements.

The methodological approach followed during the I-BiDaaS baseline phase, aims to ensure that the project will deliver a technical solution whose components are aligned with the business goals of the industrial partners. It builds upon and advances current research in the context of requirements elicitation, architecture specification and experimental validation of industrial Big Data solutions. Furthermore, it resulted in the systematic specification of an end-to-end Big Data as-a-self-service conceptual architecture which complies with and extends existing Big Data Reference Architectures.

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