

Workshop C2

Big Data Interoperability for Enterprises

Workshop C2 Report

MARTEN VAN SINDEREN and MARIA-EUGENIA IACOB
University of Twente

Technological advances and their wide adoption have led to the current era of ‘big data’, i.e. the availability of large volumes of data to enable better decision making and process optimization. Sources of such big data include:

- Low-cost networked sensors embedded in physical objects and environments, including mobile devices, household equipment, homes, production tools, offices, roads and outdoor environments, which allow for real-time monitoring of many ‘things’ of interest.
- Corporate websites publishing information on weather, traffic, public transport and fares, assets, infrastructure, public records, etc., which provide important auxiliary input to decision processes.
- IT-based social networking services such as Facebook, Twitter, LinkedIn, Google+ and Tumblr, used by people to inform others about their experiences, observations and activities, and which collectively are good indicators for existing sentiments and emerging trends.

Although these sources produce a wealth of data, only by extraction of ‘value’ through advanced analytics the potential of big data can be exploited. Which value an enterprise is looking for, and thus which criteria for extracting and evaluating data should be applied, is determined by the business goals that are to be supported.

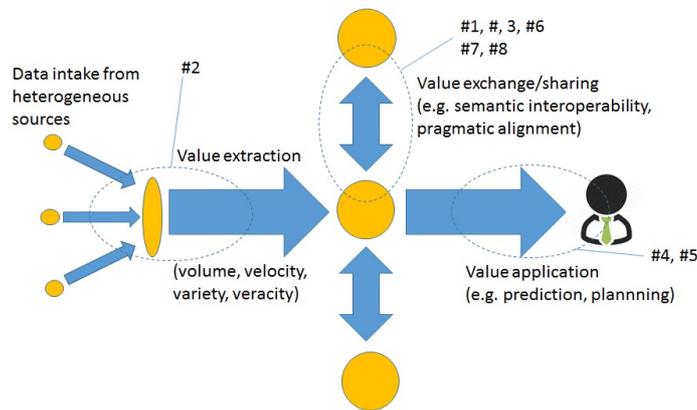
More recently, data-driven enterprises manifest an increasing interest in big data and its (business) potential. They want to know the state of their processes, resources, products, services, markets, customers, etc., and use this knowledge to make decisions that will help them achieve their business goals. Enterprises can use big data to become more ‘attentive’ and ‘proactive’, in order to prepare for and adapt to changing circumstances and upcoming opportunities. Big data provides the underpinning of context-awareness and business intelligence, allowing ‘smart’ decision-making based on patterns discovered in historical data, current situations and future trends. Cost-efficient use of big data can create significant competitive

advantages in the firm-level value chain, but also in industry-level and global value chains, across enterprise boundaries.

However, cost-efficient use of big data by enterprises is challenging, as illustrated by the figure below. Data from multiple heterogeneous sources are typically combined. This data needs to be handled by various interacting components in different systems for automated transformation, filtering, processing and analysis, as well as for representation according to adopted industry standards and interpretation according to prevailing industry models. The derived information may be shared with many other systems, and applied for many purposes. Hence, big data interoperability and integration is a major concern that must be addressed at different levels and along the (extended) value chain.

The BD4E workshop featured 8 paper presentations:

- FAIR data points supporting big data interoperability;
- a big data wrangling approach for machine and sensor data in manufacturing;
- product-service design structure matrix – a method to synchronize developing processes;
- a benchmark for predicting turnaround time for trucks at a container terminal;
- the role of domain analysis in prediction instrument development;
- achieving big data interoperability in disease surveillance;
- big data interoperability challenges in logistics;
- benefits of multi-level modeling in ecosystem interoperability.



These presentations addressed different aspects of the big data interoperability challenge, as indicated in the figure by the dashed ovals, and in different application areas, ranging from life sciences and health (1 and 6) to manufacturing (2 and 3), transportation and storage (4, 5 and 7), and oil and gas (8).