Modeling of Value Chain for Big Data

Sangwon Lee\textsuperscript{a}, Sungbum Park\textsuperscript{b}, Jumin Lee\textsuperscript{c}, Hyunsup Ahn\textsuperscript{d}, Yong Goo Choi\textsuperscript{e}

\textsuperscript{a}Division of Information and Electronic Commerce (Institute of Convergence and Creativity), Wonkwang University
\textsuperscript{b}Big Data Strategy Center, National Information Society Agency
\textsuperscript{c}Department of E-Business, Kyung Hee Cyber University
\textsuperscript{d}Department of Wirtschaftsinformatik, Technische Universität Braunschweig
\textsuperscript{e}Department of Computer Information, Dong Seoul University

The volume sub-challenge requires novel approaches, often referred to as Big Data technologies and methodologies. Data is generated constantly in an ever growing number of places and by an ever growing number of actors while a large proportion of potentially re-usable data resides within silos within institutions or companies. These are needed when conventional database technologies cannot be applied to storage and computing issues. The issue of big data has been referred to as the next frontier in computing. In this paper, we research on factors to design an organizational value chain for Big Data.

\textbf{Key Words:} Big Data; Value Chain; Modeling

I. Introduction

A better use of data, building on progress in data-analytics and processing, has the potential to transform service industries and significantly increase their efficiency. The data value chain is at the centre of the future knowledge economy, bringing the opportunities of the digital developments to the more traditional sectors. This unlocked value of data will enhance a wide range of new innovative information products and services thereby increasing the productivity of many companies, including SMEs. It is also a key area for web-entrepreneurship, given that there are no capital-intensive barriers preventing market-entry, in particular, in the light of developments.

II. Concept of Data Value Chain

For some companies the production and sale of data is a business goal in itself, but more and more, valuable data is produced as a by-product of activities that serve a different primary purpose. The focus of the data value chain strategy will be on the value generation at the centre of the process (data creation, data analysis and usage, data storage and long-term preservation). Several layers of value can be added to the original raw data. Ideally data is generated by public sector bodies or private companies in a way that facilitates re-use and value generation based on the data over and over again. Special attention is to be paid to a particular aspect of the data value chain, namely to citizens’ use of their own personal data. This specific
aspect also appears in the graph as it seems to be crucial to facilitate individuals' empowerment to better use their personal data for private or economic purposes. The data value chain includes the followings: (1) aggregated services with aggregated value: Thomson Reuters financial information service integrating a wide range of data-services. (2) data services and products: Decision Support System (DSS) designed for the oil sector; apps, based on traffic information, maps and meteorological information. (3) analysis, visualization, reporting: activities of companies like Augify, Quadrigram and InstantAtlas. (4) aggregation, integration, brokering and pre-processing: activities of companies like Mendeley, DataMarket and Talend. (5) data validation: activities of companies like Data Publica, Timetric and Duedil.

III. Value Extraction from Data

Problems caused by these differences have been addressed by policy initiatives and legislation in different areas. Different policy practices and legislation have a large impact on what can be done with data, and lead to fragmentation within the digital internal market. The data value chain strategy will focus on the aspects that are directly affected by or have a direct relevance for the way the data market will develop in future.

The lack of interoperability between datasets is relevant across the board and represents a major problem when combining data from different sources and turning them into new data products and services or using them for business intelligence purposes. It makes it also more difficult for machines to analyse, combine and extract value out of large amounts of greatly varied data. Better standardization and the creation of data in harmonized machine-processable format can help to overcome this barrier. In fact interoperability issues need to be taken into account at the stage of the production of the data. Data-harmonisation will contribute to a 'Semantic Web' allowing data to be shared and reused across borders, languages and sectors.

A pool of data that can be easily found and used by everyone is an important base for innovation. A portal for open data as one of the digital service infrastructures will make it easier to find and use data from different sources for innovative services and products. At the same time it will drive the harmonisation of data-formats and licensing conditions across the country, and will offer services such as an apps store.

IV. Conclusions

Given the fact that data economy is fragmented to a great extent along sectors, languages and borders, it is essential that significant efforts are made towards a coherent data ecosystem that will have a positive effect on the whole of the economy. The strategy on the data value chain aims at extracting the maximum value from data by building on the intelligent use of data sources. Therefore a strategy on the value chain strategy is necessary to establish the pre-conditions for the cross-border, cross-sector and cross-language flow of large amounts of public and private data throughout the world in order to create value again and again. This value will then be translated into growth and jobs which in the end bear benefits for the whole economy and all individuals.

References