Pricing Approaches for Data Markets

Alexander Muschalle[1], Florian Stahl[2], Alexander Löser[1], and Gottfried Vossen[2]

[1] TU Berlin, FG DIMA, Einsteinufer 17, 10587 Berlin, Germany
forename.surname@tu-berlin.de
[2] University of Münster, ERCIS, Leonardo-Campus 3,48149 Münster, Germany
forename.surname@ercis.de

Abstract. Currently, multiple data vendors utilize the cloud-computing paradigm for trading raw data, associated analytical services, and analytic results as a commodity good. We observe that these vendors often move the functionality of data warehouses to cloud-based platforms. On such platforms, vendors provide services for integrating and analyzing data from public and commercial data sources. We present insights from interviews with seven established vendors about their key challenges with regard to pricing strategies in different market situations and derive associated research problems for the business intelligence community.

1 Introduction

The analysis of freely available data, together with commercial and in-house data, is an increasing market segment. One example is market research focused analytics of Web data, with the aid of natural language processing technologies and statistical methods. In order to analyze data sets of such size large IT infrastructures need to be built. However, this is potentially costly as such systems usually have high implementation costs, as well as significant further costs for updating and analyzing data. Even though, we can reduce the problem to a few hundred GB that will fit into main memory, these companies still need capable staff in their IT department who can maintain and program complex in-memory multi-core infrastructures [22]. In particular, for small and medium enterprises (SME) the associated risks with such infrastructures are a strong barrier for innovation.

Arguably, cloud computing could lower this barrier as far as operating the hardware is concerned. It has to be taken into account that, though cloud computing infrastructure might be operated at lower cost, it is still likely to be to expensive for a single SME to rent the hardware necessary to crawl and analyze significant portions of the Web. This is why SMEs are not yet in a position to benefit from the latest research in cloud computing and Web mining. Nevertheless, SMEs often hold unique assets for transferring data into business relevant information, e.g., domain knowledge of a particular niche or relationships to potential customers that are willing to pay for the information.

Only recently, vendors of data, providers of data warehouse solutions, and algorithm providers have started to offer their products as platform-, software-,

M. Castellanos, U.Dayal, and E.A. Rundensteiner (Eds.): BIRTE 2012, LNBIP 154, pp. 129–144, 2013.

and data-as-a-service on so called data market places. These data markets supply analysts, business applications, and developers with meaningful data and an eased data access. Moreover, for data producers these marketplaces act as single integration platform. As a result, data marketplaces enable completely new business models with information and analysis tools as electronically tradable goods. The collective storage, analysis, and utilization of data from the Web on a central platform offers many cost savings and high innovation capabilities. Thus, especially for SMEs significant market entry barriers and impediments to innovation are eliminated.

In this paper we present an empirical study with seven data market owners and producers of data associated products and services. In this preliminary study with early adaptors, we collected answers from our interview partners for the following important questions:

1. What are common queries and demands of participants on a data market?
2. Which pricing models utilize beneficiaries for data associated products?
3. Which research challenges for the business intelligence community may arise from the combination of data and data associated services in data markets?

The reminder of this paper is structured as follows: In Section 2 we report on two query categories and seven types of beneficiaries that are common across all interview partners. In Section 3 we discuss the market situation of our interview partners, reveal current pricing strategies and derive major research challenges in Section 4. In Section 5 we discuss related work and conclude in Section 6. Finally in Appendix A we give a brief introduction to our research methodology.

2 What Is a Data Market?

In this section we review common elements of business models and business demands of our partners. We start with queries that a data market should be able to answer and then report our observations about seven beneficiaries and their demands.

2.1 Common Query Demands

During our interviews we asked our interview partners for common demands of their customers. We observed a heterogeneous set of queries from which we could derive the following two major scenarios:

Estimate the Value of a 'Thing', Compare the Value of 'Things'. In the first scenario, customers abstract the data market as a data warehouse that resides on an open set of data sources. These customers utilize the data market for collecting and measuring factual signals from public data sources, such as the Web or 'Open Data' from the UNO, and non-public sources, such as commercial data from Reuters and in-house private data, such as data from an ERP or CRM system. Given the set of available data sources on the data market and