Chapter 16
Big Data Governance in Agile and Data-Driven Software Development:
A Market Entry Case in the Educational Game Industry

Lili Aunimo
Haaga-Helia University of Applied Sciences, Finland

Ari V. Alamäki
Haaga-Helia University of Applied Sciences, Finland

Harri Ketamo
Headai Ltd., Finland

ABSTRACT

Constructing a big data governance framework is important when a company performs data-driven software development. The most important aspects of big data governance are data privacy, security, availability, usability, and integrity. In this chapter, the authors present a business case where a framework for big data governance has been built. The business case is about the development and continuous improvement of a new mobile application that is targeted for consumers. In this context, big data is used in product development, in building predictive modes related to the users and for personalization of the product. The main finding of the study is a novel big data governance framework and that a proper framework for big data governance is useful when building and maintaining trustworthy and value adding big data-driven predictive models in an authentic business environment.

DOI: 10.4018/978-1-7998-8954-0.ch016

INTRODUCTION

A big data governance framework is critical when a company performs data-driven software development and business. The authors adhere to the common definitions of big data governance and of data governance, which were presented by Sarsfield (2009), Soares (2012), and DAMA (2017). To meet the requirements of an agile start-up company that needs to manage and govern big data based predictive models and the data related to them, the authors propose a framework for big data governance. This framework includes five key dimensions for big data governance: data privacy, security, availability, usability, and integrity. Each dimension is described in the big data governance framework section. These five key dimensions are important in building and managing a successful big data driven business.

In this chapter, the authors present the business case based on which the authors have derived the proposed big data governance framework. The business case concerns the development and continuous improvement of a new mobile educational application that is targeted at children and young people who wish to learn to play soccer. In this context, big data are used in three contexts: 1) product and business development, 2) building predictive modes about the learners' progress in general and 3) personalisation of the product to meet the needs of each individual learner.

Previous research and guidelines on how to govern big data governance (e.g., Soares, 2012; DAMA, 2017) have been published. However, there is a research gap concerning big data governance in agile big data driven start-up companies and especially on how to govern big data in the product development phase.

The main finding of the study is the proposed novel big data governance framework and the fact that a proper framework for data governance is necessary when developing trustworthy and value-adding big data driven software products in an authentic business environment. Without big data governance, the predictive models and other data-driven applications may not bring added value to the business because their trustworthiness is uncertain, they might violate the right to privacy of customers, or they might not be available for use when needed. In addition, without proper big data governance, they might not meet the needs of the business, or valuable data might be leaked to competitors because of the poor governance and weak security of the big data. If a company succeeds in big data governance, data can become its most valuable asset (Panian, 2010).

BACKGROUND

The field of big data governance emerged with the advent of big data. Big data are data that cannot be processed using traditional data processing software and infrastructure on a personal computer or on a dedicated server (e.g., Liebowitz, 2013). In addition, big data differ from traditional data in volume, variety, and/or velocity (Liebowitz, 2013). Typical examples of big data include web and social media data, machine-to-machine data, big transaction data, biometric data and human-generated data (Soares, 2012). The authors include Internet of Things (IoT) data and data generated in industrial processes in the broad set of transactional data.

Need for Big Data Governance in Companies

In technology companies, there is an increasing need to develop data analytics especially in launching new products and services. The diffusion of technology has increased rapidly (Downes & Nunes, 2013),