Terramycelium: A Reference Architecture for Adaptive Big Data Systems

Pouya Ataei

School of Engineering Computer and Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand

Email: pouya.ataei@aut.ac.nz

Abstract—Big Data is a nascent term that emerged to describe a large amount of data that comes in different forms from various channels. In the modern world, users are the ceaseless generators of structured, semi-structured, and unstructured data that, if gleaned and crunched precisely, will reveal game-changing patterns. While opportunities exist with big data, the unprecedented amount of data has brought traditional approaches to a bottleneck, and the growth of data is outpacing technological and scientific advances in data analytics. It is estimated that approximately 75% of the big data projects have failed within the last decade, according to multiple sources. Among the challenges, system development and data architecture are prominent. This paper aims to address these challenges by introducing a novel big data reference architecture. This reference architecture is developed absorbing the principles of complex adaptive systems, domain-driven design, distributed systems and event-driven systems. The refence architecture is developed following the guidelines of creating empirically grounded reference architectures and is evaluated using two distinct methods: case-mechanism experiment and expert opinion. The results of the case-mechanism experiment demonstrates the artifact's capability in meeting the requirements of big data systems, while being highly maintainble and scalable. Moreover, the feedback collected from the experts presents that Terramycelium offers advantages for organisations committed to big data, yet its alignment with organisational cultures or structures varies.

Index Terms—Big Data, Big Data Reference Architectures, Big Data Architectures, Big Data for business, Data analytics, Data engineering, Data-intensive applications, Reference Architectures