# 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 by absorbing the principles of complex adaptive systems, domain-driven design, distributed systems, and event-driven systems. It follows the guidelines for creating empirically grounded reference architectures and is evaluated using two distinct methods: a case-mechanism experiment and expert opinion. The results of the case-mechanism experiment demonstrate the artefact's capability to meet the requirements of big data systems while being highly maintainable and scalable.