Big Data Analytics – Use Case Identification Framework
03/05/2019
Sameer S Paradkar

Abstract

This white paper provides a methodology for identifying big data use cases and shares the best practices and standards that should be leveraged for big data use case identification. The later section of the document is a catalogue of big data use cases providing a consolidated view for various industry verticals.

Big Data User Cases:

Application of big data is proliferating rapidly across industries. Big data if analysed properly has the power to uncover hidden patterns and insights, which helps organization make informed business decisions. Besides, organizations can draw valuable insights and conclusions on overall business processes. For proper analysis of big data, there is a need for organizations to leverage advanced technologies like big data analytics.

But before that, organizations need to carry out big data use case identification. Identifying the right area to harness the data will maximize organizational growth and increase revenue. But, finding the right use case is itself a challenging process.

Methodology identifying - Big Data Use Cases

Use cases are gaining more and more importance when it comes to the exploration of ideas, requirements analysis, and prioritization of tasks in analytics and Big Data projects. To identify, structure, and prioritize the most promising use cases, we recommend combining two common approaches: ideation (top-down) and data-driven (bottom-up).

The ideation approach leverages innovation and creativity to come up with new ideas. Methods like design thinking help subject matter experts to identify and formulate potentially beneficial use cases. Thinking outside the box supported by creativity techniques has proven helpful in coming up with ideas. The biggest weakness is the lack of validation and substantiation of ideas with data and facts that could lead “theoretically designed” innovations, which are difficult to implement in actual organization environments.
The second approach is a data-driven approach to finding use cases. It starts with the data at hand, potentially adding additional external or previously not accessible internal data. Applying techniques like pattern recognition help understand what is in the data and which use cases can be derived from it. This pragmatic approach works very well for initial analysis of data, but higher-value and successful use cases can be identified easily if a data-driven approach is combined with an ideation approach.

With digitalization, companies are increasingly setting up “systems of innovation” and supporting organizations for explorative data analysis. But neither the exact requirements nor the real business value of these new analytic applications and their models can be defined upfront. Instead, many use cases usually must be examined during the POC/Project to see if there is really added value.

To reduce costs and the failure rate it is therefore helpful to identify and prioritize the most promising use cases early on. To select use cases as the foundation of a data or analytics strategy or to steer projects that are initiated to support digitalization and create innovative applications.

The starting point for the methodology is typically a workshop to collect ideas from domain and data experts from different levels and areas of the business. The goal is to understand what ideas there are for an innovative use of data. At the same time, data sources and business processes are looked at to identify where more, different, or better data could bring value in steering processes, forecasting behaviour, or making strategic decisions. Methods like design thinking may help in this situation to extract and formulate ideas, solve conflicts, or clarify ambiguous aspects in terms of available data, tools, skills, or processes.

The next step is to filter and concretize all these ideas into use cases that could generate high business value. All use cases get systematically documented in a use case grid. It is important to identify use cases that not only create insights but also have an actual impact on the organization’s bottom line.

Information gathered contains, for example, the involved data, data sources, required functionality and technology as well as organizational aspects. Yet this may also reveal possible limitations that need to be considered in further activities. At the end of this process, an understanding of the potential benefit and limitations of all use cases is reached so they can be prioritized using a scoring model.

Highly prioritized use cases are subject to a proof of value. In this test, domain, analytics, and data experts team up to validate the assumptions and possible outcomes of a use case. It is especially important to achieve rough checks of data and first versions of models quickly. In a “fail-fast” culture, time is of more importance than thoroughness since testing and failing is an integral part of explorative process. Once the goal of an analysis is defined and the data prepared, a data scientist can generate initial results that help to decide whether it is worth spending more time exploring a certain use case.
Summary

The primary purpose of big data analytics is to identify hidden patterns and correlations, to extract value from the routine data. Big data analytics help companies create new opportunities and improve their existing business processes. In turn, this helps companies to conduct their business better, thereby meeting their goals with high efficiency and accuracy.

Business analysts and leaders should understand the business objectives before identifying the business use cases. List all the business use case categories by predicting the respective business outcomes. Keep a regular check for new business use cases from the data that pours in real-time.

The bridge that connects digital data and advanced technologies is a business use case. Hence, with a team of highly-skilled professionals and analysts, one should identify the right use cases by following the earlier approach, thereby breaking the data silos for enhanced business outcomes.
About the Author

Sameer Paradkar is an Enterprise Architect with 15+ years of extensive experience in the ICT industry which spans across Consulting, Product Development and Systems Integration. He is an Open Group TOGAF, Oracle Master Java EA, TMForum NGOSS, IBM SOA Solutions, IBM Cloud Solutions, IBM MobileFirst, ITIL Foundation V3, COBIT 5 and AWS certified enterprise architect. He serves as an advisory architect on Enterprise Architecture programs and continues to work as a Subject Matter Expert. He has worked on multiple architecture transformation and modernization engagements in the USA, UK, Europe, Asia Pacific and Middle East Regions that presented a phased roadmap to transformation that maximized the business value, while minimizing costs and risks.

Sameer is part of Architecture Group in AtoS. Prior to AtoS he has worked in organizations like EY - IT Advisory, IBM GBS, Wipro Consulting Services, TechMahindra and Infosys Technologies and specializes in IT Strategies and Enterprise transformation engagements.

Disclaimer

The views expressed in this article are the author’s views and AtoS does not subscribe to the substance, veracity or truthfulness of the said opinion.