Assignment-9

Sai Hemanth Nirujogi
G01065588

The main purpose of this paper is to study the practices, strategies, and failures of Big Data management in the industry. The author discusses about the main data strategies of an enterprise as data governance, sharing, standards, analysis and adds the fourth factor as “We need to devote at least as much time to planning and managing the people and business processes that make data analysis possible as we devote to the analysis process itself and the technologies that support it.” [1] The author also made some suggestions for planning and managing data-intensive projects. The top 5 best practices for Big Data management and plan are:

1. Don’t let the tools drive your policy
2. Always think about the boundaries
3. Understand where the data comes from and where it’s going
4. Keep track of where you are and where you’re going
5. Know your stakeholders

Don’t let the tools drive your policy

There are various kinds of data cleaning tools present in the market which are been used by various companies for the purpose of removing duplicate data, fixing the incorrect and incomplete values from the databases. These data cleaning tools save money by omitting bad records and deals with inaccuracies within seconds saving a lot of time and effort of the organization. According to studies organizations end up wasting upto 6 million dollars only because of no proper data cleaning tools. The data cleaning software help to correct incorrect data or duplicate data such as the emails, phone numbers of the customers resulting in contacting them. One of the most preferred example in this scenario is due to duplicate data the marketing organizations may send multiple marketing advertisements to the same customer due to duplicate data resulting in losing the customer. These type of errors can be controlled by the organization within no time and save valuable resources of the organization as it would take a much longer time to delete the records manually.[2] These tools are incredibly powerful and available in the market for big organizations for a high price. But relying completely on automated analytic tools for analysis can be consequential. We have been using traditional tools for more than 30 years now for the data analysis, so an organization should make sure the tools they are using for the process should be appropriate.

Always think about the boundaries

Many big organizations like Microsoft, Google, Adobe, SAP etc. has started the open-data initiative, which means that the data can be used by the public via portals, visualizations and other analysis. This process is user-centric in the approach and is meant for the brands and small organizations. The users of these platforms can also transfer their data from one platform to another platform. This allows the organizations what the consumer needs and achieve a better business

understanding. This also helps the organizations to create and adopt new powerful applications that natively analyze the data and the relationships from multiple services. [3] The author poses a question about the standards that will need to be implemented with different data sources and they need to be considered while creating the project plan.

Understand where the data comes from and where it's going

A data source has no beginning and no end. It is either the data format may change forcing us to change the process of analytics. Data validation is a crucial and ongoing challenge which is highly complex for Organizations because they have to come up with a new technology and policy changes. It is very essential for an organization to ensure the data integrity and immutability because if the data gets modified will be of no value at all. When an erroneous information enters an organization's complex system even a minute error can lead to revenue and efficiency loss, and failure to comply with industry regulations. [4] The initial capture of the data at the beginning of the project should be the responsibility of the project manager. The manager needs to make sure with all the technical team that all of them has the understanding of where the data goes after the analysis. The analyzed trends will have a great value only if the data is kept to a better use.

Keep track of where you are and where you are going

With a team in place for the management plan, understand the skills and wisdom of each individual to assign better tasks. This helps in tracking who is responsible for which part of the project and can easily get updates from them. The team under the management should perform a needs assessment, which shares all the information and characteristics of the project requirement gathering for the future use. The project manager must take care of how the data and the working group are performing in the progression of the management plan to deliver useful functionality. The strategy used should be able to maintain the systems and business processes for each functional area and division of the plan. [5]

Know your stakeholders

No one in the organization has complete awareness into the requirements of the business. The development, operations, marketing, and finance teams have a lot of requirements in common and unique ones as well. The CDO has to take care of the business transformation that relies on the data and intelligence and should represent the interests of the organizations. The data analysts should fix the data issues, curate datasets, and fix the hoc requests. The first step in the project management plan always should be to know your stakeholders and assemble a trusted working group of people. This group will play a major role in the project to find and understand the requirements of ownership across the organization. [5]

References

1. McDonald, D. D., Ph.D. (2015). Planning and Managing Big data Projects. Planning and Managing Big Data Projects. Retrieved December 2, 2018, from https://static1.squarespace.com/static/52fbe871e4b060243dd758d6/t/578e649337c581c4be12cb4d/1468949654944/.
2. B, Susan. “How Businesses Can Benefit from Data Cleansing Software.” The Data Cleansing Blog, The Data Cleansing Blog, 4 Oct. 2018, accessed on 2 Dec. 2018 www.winpure.com/blog/how-businesses-can-benefit-from-data-cleansing-software/.
3. Nalawade, Shalaka. “Is the Open Data Initiative as Good as It Sounds? MTA Looks at This Data Alliance.” MarTech Advisor, accessed on 2 Dec. 2018 www.martechadvisor.com/articles/data-management/data-alliance-open-data-initiative/.
4. Gao, Jerry, et al. “Big Data Validation and Quality Assurance -- Issuses, Challenges, and Needs.” 2016 IEEE Symposium on Service-Oriented System Engineering (SOSE), 2016, accessed on 2 Dec. 2018, doi:10.1109/sose.2016.63.
5. Author: “Four Steps To A Data Management Strategy In Light Of Big Data.” The Hassle Free Technology Blog, accessed on 2 Dec. 2018, www.zlhent.com/blog/2017/09/four-steps-to-a-data-management-strategy-in-light-of-big-data/.