Big Data Analytics in a Healthcare Organization

John Wensink

MIS440 - Cloud Computing and Big Data

Colorado State University-Global Campus

Dr. Mazen Alkhatib

February 2, 2020

Big Data Analytics in a Healthcare Organization

A vast abundance of data is available to support healthcare professionals, but how can we as an organization make use of this data while still protecting patient’s privacy and maintaining compliance within governmental regulations. This organization stands to gain from the wealth of electronic data available such as digital health records, data from smartwatches and other wearables, genomic sequencing, multidimensional medical imaging, pharmaceutical research, as well as records from other healthcare organizations (New England Journal of Medicine, 2018). Healthcare information is a unique set of data with an incredibly high volume, the velocity of the data is fast-moving, and the variety of healthcare data is extraordinarily wide encompassing the full spectrum of the digital landscape. The value of this information is tremendous, and we need to be certain of the data’s authenticity and veracity. Traditional databases fall short of this data’s potential to revolutionize the healthcare industry and as such, our firm proposes a unique approach by leveraging cloud technology to facilitate acquiring, storing, analyzing, processing, and transmitting this valuable data, with emphasis placed on the patient’s right to own their personal digital information.

**Cloud Deployment Model**

The cloud deployment model which fits our mission best is that of the community cloud, which will be made available to employees within different areas of our organization, as well as a portal for the patients themselves (Mell & Grance, 2011). This community cloud will be managed by specialists within our own organization who have a deep understanding of the possibilities and requirements that managing such a system entails. Emphasis must be placed on acquiring and retaining individuals with a high degree of technical literacy as it pertains to cloud-based data management in the field of medical information. Third-party providers are to be avoided wherever possible. Although the use of local employees will come at a higher cost to the organization than reliance on third-party contractors, there are tangible benefits to this approach including accountability with information security, high standards of productivity are easier to enforce, and locally sourced employees will be more personally invested in providing a high-quality service to their own community which stands to benefit from the job creation this project will require (Albanese, 2018).

**Cloud Service Model**

The scope of this project calls for the use of the Platform as a Service (PaaS) cloud service model, as it emphasizes the creation of a new web application aimed at revolutionizing the healthcare industry by the use of big-data analysis. Although we are creating a new web application, we don’t need to re-invent the wheel and will want to utilize existing application development tools that have proven themselves to be secure, reliable, and efficient as they pertain to eliminating repetitive tasks that are commonly associated with web-app development. Google’s AppEngine (GAE) seems to be the obvious choice for our PaaS development goals as it contains tools for both developing as well as hosting our healthcare information system. GAE offers scalability as well as the efficiency with their on-demand fee structure, with the ability to commission and de-commission dynamic hardware resource requirements. Our organization will need to investigate the use of a Docker-based containerized infrastructure versus the use of hypervisor-based virtual machines to determine which approach will most closely fit our data isolation needs as a portable healthcare data information system (Schlosser, 2020). We understand that your organization’s resources are not infinite, and to that end, we recommend the utilization of open-source analysis tools including Apache’s Hadoop used in combination with Spark for gaining insights into the data collection which is sure to span vast data clusters (Apache, 2018). Using open-source tools is one way that our application will seek to balance the cost-benefit relationship which will carry a high cost burden from the avoidance of third-party contractors.

This project has great potential to revolutionize the way the healthcare industry uses big-data analytics to make intuitive tools that can provide our professionals with insights into treatment possibilities. It is of vital importance to remember our organization’s mission to improve the lives of the patients we treat while acting ethically and in compliance with existing and future governmental regulations. For this reason, we recommend that the search for high-quality data professionals begin without delay so as to gain insights from their experience as we begin to build this application which holds tremendous potential to improve the lives of the people we serve.

References

Albanese, J. (2018, July 30). Relying on the Business of Strangers: How to Manage Third-Party Vendors, Suppliers, and Contractors. Retrieved February 2, 2020, from <https://www.inc.com/jason-albanese/relying-on-business-of-strangers-how-to-manage-third-party-vendors-suppliers-contractors.html>

Apache. (2018). Apache Spark™ - Unified Analytics Engine for Big Data. Retrieved February 2, 2020, from <https://spark.apache.org/>

Mell, P., & Grance, T. (2011, September). The NIST Definition of Cloud Computing. Retrieved February 2, 2020, from <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>

New England Journal of Medicine. (2018, January 1). Healthcare Big Data and the Promise of Value-Based Care. Retrieved February 2, 2020, from <https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0290>

Schlosser, H. (2020, January 30). Docker vs. Virtual Machine: Where are the differences? Retrieved February 2, 2020, from <https://devopscon.io/blog/docker/docker-vs-virtual-machine-where-are-the-differences/>