Data Democratization

M. N.O. Sadiku[1], S. R. Nelatury[2], S.M. Musa[1]

[1]College of Engineering, Prairie View A&M University, Prairie View, TX 77446
[2]School of Engineering and Engineering Technology, Pennsylvania State University, Erie, PA 16563-1701

Abstract Data democratization is allowing digital information to be accessible to an average user. Traditionally, access to data is granted to executives, data scientists and information technology staff. The goal of data democratization is to allow non-specialists to gather and analyze data. It empowers a broader-than-ever audience to extract value from data and use it to effect positive social change. This paper provides a brief introduction to data democratization.

Keywords data democratization, democratization of information

Introduction

Data is generated by multiple sources every day and mobile data is experiencing exponential growth. Governments and organizations are generating open data these days for the sake of transparency. Some local and state government believes that information is power. The technological obstacles that hindered data accessibility have now disappeared. However, almost all of the data remains unused or unanalyzed because there are not enough data scientists to analyze and derive meaningful result from the data.

Data is everywhere, but being able to understand complex huge volume of data has been the preserve of a handful of highly trained, paid data scientists and analysts. However, organizations are increasingly realizing the need for everyone being able to access and understand data for effective decision making. The concepts of enabling everyone have access and understand data is known as data democratization [1]. Proponents of the concept believe it is a game-changer and will narrow the gap between big brands and smaller businesses. It empowers people everywhere with rich information. It is a source of decisive competitive advantage and even survival for most organizations. Though data democratization may appear like a tall order, it will be a key benefactor for organizations embracing it.

Components of Data Democratization

Data democratization efforts focus on four pillars [2]: data, tools, training, and people.

- *Data*: Open data is moving towards a big data scenario, in which the data is often too hard to tame and some of it typically remains in silos spread across Microsoft SQL Server. The quality data can always help to inform strategic decision-making and tactical planning.
- *Tools*: A single analytical tool cannot process all kinds of data. Some user-friendly data mining techniques can be applied to discover some knowledge from data. Data democratization requires such data analysis techniques can be used by everyone.

- *Training*: To ensure that data democratization does not come with data misinterpretation, people should be trained on how to use data. Graduate programs with emphasis on GIS and urban policy could be encouraged. For example, training on how to use GPS now takes days rather than months.
- *People:* For people to cope with this new age, they need to become data literate. People should be involved and motivated to think, ask questions, and play with data. Democratization requires that we share information with people in a format that they can easily read and understand.

Data is usually repackaged and made available in the forms of tables, maps, graphs, or charts. Advances in virtualization make data democratization much easier to adopt. Visualization of library data highlights purchasing decisions, future library needs and goals. Librarians, as de facto experts of data visualization, can assist students, faculty and researchers visualize their data [3]. As organizing tools, maps and visual representations of data can be used to mobilize people and gain their commitment on important issues. Using maps and data can lead to increased credibility of a company [4].
.

Issues with Data Democratization
A company that embraces democratization can face some issues and challenges. Several obstacles lie between merely making data available and the reality to use the data for decision-making. Professional data analysts are good at turning data into information, information into knowledge, and knowledge into sound decision [5]. A remarkable gap in the demand and supply of data scientists is impeding the data analysis potential of corporate and governmental organizations. This is debate on whether data democratization really benefits companies that adopt it. There are concerns about the security risks involved in allowing everyone access to sensitive data. This may require that the company provide adequate training to those using the data.

Conclusion
The promise of easy access to data is achieved today through the idea of data democratization. Democratization of data is the phenomenon of making data available to people who need it, when and where they need it, *i.e.* to all the layers of the organization while transcending departmental boundaries. It is important these days that non-expert users can exploit the vast amount of data in order to make well-informed decisions [6].

References
[1]. "Beginners Guide To: Data Democratization," https://channels.theinnovationenterprise.com/articles/beginners-guide-to-data-democratization

[2]. V. Vasileiou, "Data Democratization. What, why, and how?" http://blog.zopa.com/2016/12/02/data-democratization/

[3]. T. J. Brigham, "Feast for the eyes: an introduction to data visualization," *Medical Reference Services Quarterly*, vol. 35, no. 2, 2016, pp. 215-223.

[4]. S. Treuhaft, "The democratization of data: how the Internet is shaping the work of data intermediaries," Working Paper Series, UC Berkeley, 2006.

[5]. D. S. Sawicki and W. J. Craig, "The democratization of data," *Journal of the American Planning Association*, vol. 62, no. 4, Autumn 1996, pp. 512–523.

[6]. R. Espinosa et al., "Linked open data mining for democratization of big data," *Proceedings of IEEE International Conference on Big Data*, 2014, pp. 17-19.

About the authors
Matthew N.O. Sadiku (sadiku@iee.org) is a professor at Prairie View A&M University, Texas. He is the author of several books and papers. He is an IEEE fellow. His research interests include computational electromagnetics and computer networks.

Sudarshan R. Nelatury (srn3@psu.edu) is an associate professor at Penn State University, The Behrend College, Erie, Pennsylvania. His teaching and research interests lie in electromagnetics and signal processing.

Sarhan M. Musa (smmusa@pvamu.edu) is a professor in the Department of Engineering Technology at Prairie View A&M University, Texas. He has been the director of Prairie View Networking Academy, Texas, since 2004. He is an LTD Spring and Boeing Welliver Fellow.