Open Data: Opportunities And Challenges

Matthew N. O. Sadiku
Roy G. Perry College of Engineering
Prairie View A&M University, TX 77446
sadiku@ieee.org

Mahamadou Tembely
Roy G. Perry College of Engineering
Prairie View A&M University, TX 77446
mtembely@student.pvmau.edu

Sarhan M. Musa
Roy G. Perry College of Engineering
Prairie View A&M University, TX 77446
smmusa@pvamu.edu

Abstract—**Open data requires government and other sources to make some data freely available for users to use as they deem fit without restrictions from copyright or other control mechanisms. It is growing rapidly in the United States, Canada, China, India, Germany, Ghana, and other places. It is transforming government across the globe as they provide services to citizens and remain relevant. This paper presents the new concept of open government data and addresses its benefits and challenges.**

Keywords—Open data, open data policies, open government data

I. INTRODUCTION

Data is open if people can use and distribute it freely. Open data is a recent phenomenon. The concept of an open data ecosystem is still at the early stage of development. Open data is the material which is available for anyone to freely use for any purpose, without restrictions from copyright or other means of control. The three most important characteristics of open data are: access and availability, reuse and distribution, and universal participation. The structure and semantics are also important as they allow the open data to be distributed in reusable formats.

Governments (which are recognized as the largest data producers) worldwide can see the potential advantages of opening up their data for public use. Significant amount of government open data, covering topics such as infrastructure and population profiles, are provided online for use by citizens, other arms of government, and private sector. These open data are opened to the public at no cost and with few restrictions on usage or distribution. This is in contrast to intellectual property rights and also to the classic open science in which only the final results are shared.

II. OPEN GOVERNMENT DATA

Why would governments share their information like weather and environmental data, census reports and public projects? Primary motivations to publish government data include [1]: (1) increase democratic control and political participation, (2) foster service and product innovation, (3) strengthen law enforcement.

Since 1813, Federal Depository Libraries have been allowing free public access to select US government documents. Federal, state, and local governments are now releasing a new kind of information, which is known as open government data (OGD). In 2009, President Barack Obama required government data to be made open to the public. The government established data.gov as the portal for US federal agencies to publish open data. Obama's OGD initiative is being adopted by state and local governments and even by governments around the world. In order for publicly available government data to be qualified as open, it must be complete, primary, timely, accessible, machine-processable, nondiscriminatory, nonproprietary, and license-free [2].

Government open data is usually available and collected at the websites of the various departments and agencies. These data sets are made available online. Government publishes a variety of data online to provide transparency, foster applications, and satisfy legal obligations. These portals allow one to browse by topics such as business, education, health, and public safety. Governments and local authorities now see open data as a tool for economic development.

Open data can be described in three ways. First, it must be machine-readable. It must to be made available online to increase accessibility. Third, it must be published under an open license [3]. Governmental data is regarded as open if they comply with 10 principles: complete, primary, timely, accessible, machine processable, non-discrimination, non-proprietary, permanent, license-free, and preferably free of charge [4]. It should be born in mind that open data are not limited to government data. They may also come from private sector, schools, and universities.
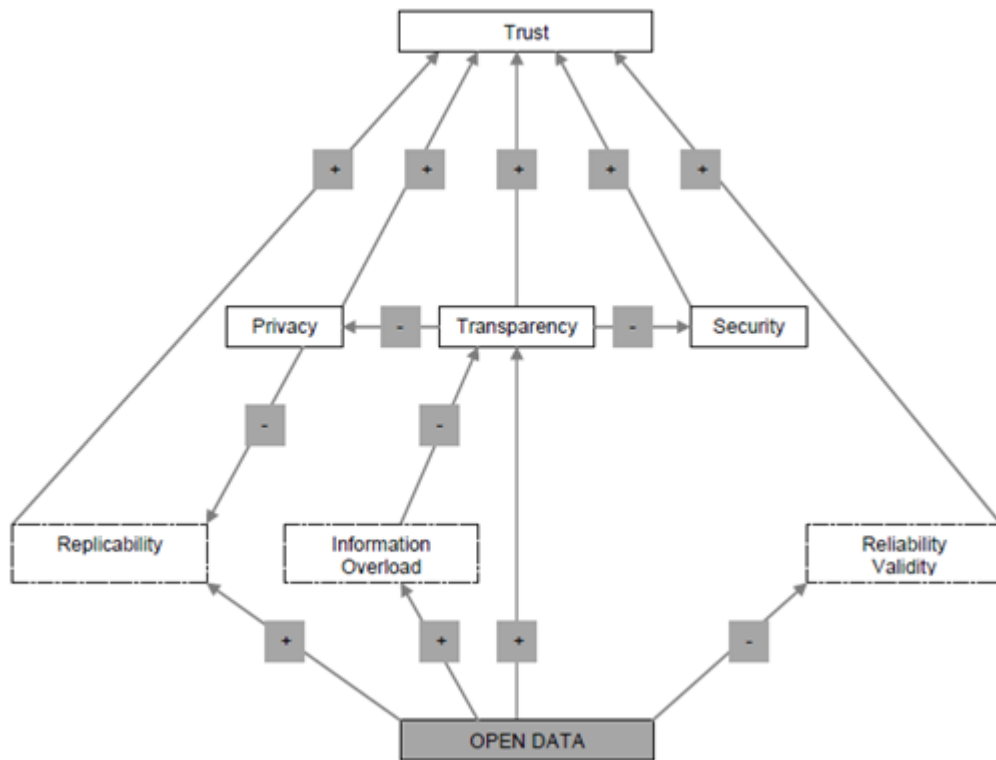
.

Figure 1.   Relationship between open data and privacy, transparency, and security [9].

III.   BENEFITS

Many governments, such as U.S., U.K., Canada, Germany, and India, and institutions, such as the United Nations, World Bank, and the African Development Bank, have established open data portals for public access. These portals enable the citizens to transform data into innovative solutions to societal problems and improve quality of life [5].

The main benefits of open government data are more transparency, participation of citizens, greater efficiency and effective government, working processes, societal and economic benefits, and filling a need [6]. The data may be used to answer questions on politics and government. Several governments (national, state or local) have started to release their data to gain these advantages. Some have launched open data initiatives such as Data.gov, Data.gov.uk, Data.gov.be, and Data.gv.at.

Advocates of open data argue as follows [7,8]:

- Public money was used to fund the work and so the work should be made available to the public.
- Equal access to data and creation of trust in government.
- Public data aids the creation of new products and services.
- Open data can facilitate the creation of new and personalized services.
- Open data can strengthen accountability, promote economic growth, stimulate innovation, build trust, enhance transparency, and improve citizen satisfaction in the nation.
- In research, the rate of discovery is accelerated by having free access to data.

Arguments against open data include [7,9]:

- Government must be accountable for proper use of taxpayer's money.
- If the data will commercially benefit the user, then the user should pay the government for the data.
- There is lack of standardization of information.
- What would motivate researchers to engage in data and resource sharing.
- Data should be made available for a charge.
- There are privacy, security, and legal implications that affect the release of some sensitive data.
- It is questionable whether citizen participation will inevitably follow from open data.

IV.   CHALLENGES

To open up government data, some policies are needed, including legislative directives and strategies. Some nations still lack a clear, well-defined open data policy. Right now, there is no guidance on what data could be open and how to track the usage of open data. Open data is still subject to licenses which determine how the data can be used.

The task of providing structured data is daunting but worthwhile. Combining open data from different nations has heterogeneity problems and integration challenges. While large amounts of data are publicly available, they do not conform to any well known schema. Making the published data understandable is crucial for adoption.

While open data is seen as a driver for innovation and transparency, it might also lead to privacy breaches and security violations [10]. Security is indispensable in order to prevent accidental or malicious disclosure of data sets. Figure 1 illustrates the relationship between open data and privacy, transparency, and security.

Standardization efforts and development of robust metadata can help data access but are not enough. Although open data portals provide a platform for a new culture of openness and creativity, there is no guarantee that all who need the data have access to it or those who have access understand it.

V. CONCLUSIONS

Open data refers to the idea that data should be available freely to everyone to use without restrictions. In return, governments obtain benefits such as long-term transparency, citizen participation, and innovation. Open data should lead to an open government and play an important role in good e-government.

Open data is still in its early stage of development. Publishing open data is gaining momentum worldwide. Open data allows citizens and businesses to have free access to public information. Several governments around the world have started including open data issues in their government strategies. They maintain websites for publishing open data for public use. In addition to the government, the private sector recognizes the potential advantages of making the data available as open data but they have not opened their data storages at the same speed as state owned organizations.

REFERENCES

[1] N. Huijboom and T. V. den Broek, "Open data: an international comparison of strategies," *European Journal of ePractice*, no. 12, March/April 2011, pp. 1-13.

[2] K. Okamoto, "Introducing open government data," *The Reference Librarian*, vol. 10, 2016, pp. 1-13.

[3] K. O'Hara, 'Government open data and transparency: Oakeshott, civil association and the general will," *Global Discourse*, vol. 5, no. 1, 2015, pp. 135-152.

[4] F. W. Donker and B. V. Loenen, "How to access the success of the open data ecosystem?" *International Journal of Digital Earth*, 2016, pp.1-23.

[5] L. M. Amugongo, S. H. Nggada, and J. Sieck, "Open data portal, a technical enabler to drive innovation in Namibia," *Proceedings of the 2nd International Conference on Open and Big Data*, 2016, pp. 80-86.

[6] A. Zuiderwijk and M. Janseen, "A comparison of open data policies and their implementation in two Dutch ministries," *The Proceedings of the 13th Annual International Conference on Digital Government Research*, 2012, pp. 84-89.

[7] "Open data," Wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/Open_data

[8] M. Janssen, Y. Charalabidis, and A. Zuiderwijk, "Benefits, adoption barriers and myths of open data and open government," *Information Systems Management*, vol. 29, no. 4, 2012, pp. 258-268.

[9] M. Wiener et., "Enabling an open data ecosystem for the neurosciences," *NeuroView*, vol. 92, Nov. 2, 2016, pp. 617-621.

[10] R. Meijer, P. Conradie, and S. Choenni, "Reconciling contradictions of open data regarding transparency, privacy, security, and trust," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 9, no. 3, Sept. 2014, pp. 32-44.

About the authors

Matthew N.O. Sadiku (sadiku@ieee.org) is a professor at Prairie View A&M University, Texas. He is the author of several books and papers. He is a fellow of IEEE.

Mahamadou Tembely (mtembely@student.pvamu.edu) is a Ph.D student at Prairie View A&M University, Texas. He received the 2014 Outstanding MS Graduated Student award for the department of electrical and computer engineering. He is the author of several papers.

Sarhan M. Musa (smmusa@pvamu.edu) is an associate 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.