Chapter 1

Sifting Through the Sets: The Significance and Availability of Open Data

Claudia McGivney
Stony Brook University, USA

ABSTRACT

Data-driven decision making has become a driving factor across many industries and reliance on verifiable data is crucial in this environment. Equitable access to reliable data is a key component in establishing a just society. This chapter provides a brief overview and background to open data resources, including current initiatives and policies, along with an exploration of challenges to accessing this information. The researcher reviewed over 2,000 open data websites and repositories in order to provide a selection of these as exemplars. All resources have been annotated, with specific attention paid to repositories offering large datasets and government resources.

INTRODUCTION

The growing volume of data being generated has significant implications for everyone, from students and academic researchers to citizens and world leaders. Curating, distributing, and analyzing this data is complex and unwieldy. At the same time, we are becoming more dependent on the communication and utilization of big data for both large-scale strategy and our day-to-day work. Increased demand for open data is similar in nature to the current movements for open access, open government, and open educational and open source resources and its impact is wide-ranging since the
access to this information has the potential for change across all sectors of daily life. Open data has been defined as research data that has been made available to copy, download, research, or re-use, without barriers - financial, legal, or technological - save the ability to access the internet (SPARC, 2016). With information increasing at unprecedented rates, the ability to access and make use of such information has raised many questions for librarians, scholars, and researchers.

BACKGROUND

The sheer amount of data being produced by open government initiatives such as Data.gov, Data.gov.uk, and Open.canada.ca have attempted to make data transparent, easily accessible and available, but this data must remain protected under the label of “open data.” At the 2013 G8 summit, the US, UK, France, Canada, Germany, Russia, Italy, and Japan committed to the Open Data Charter (Sinai & Martin, 2013) and as of 2015 this mandates that data be: data should be: open by default, timely and comprehensive, accessible and usable, comparable and Interoperable, for improved governance and citizen engagement, and for inclusive development and innovation (Open Data Charter, 2015). In order to ensure that data remain available for everyone, resources must be allocated to support the data remaining accessible, and there must be a to assessing the use of this information, as well as a commitment by those providing the data to long-term support and continued growth.

The commitment to access, long-term availability, and growth has found support in the growing open access movement and legislation has begun to reflect this philosophy when it concerns research and information that have been government funded. The EU Horizon 2020 project has maintained the view that publicly funded information should be available for perpetual access by the public with no additional costs imposed in order to fully benefit citizens, researchers, and companies alike (Teixeira da Silva, J., & Dobránszki, J. (2015). The US has adopted similar policies, specifically The Office of Science and Technology Policy, as part of a larger plan for open government. These initiatives are key legislation in ensuring that publicly funded, open data remains available to the very people that supported the research for myriad uses for their direct benefit.

With many policies still relatively new and developing, the availability of data remains a key factor in determining strategic plans for open data since persistent access to digital information tends to decrease with age (Teixeira da Silva, J., & Dobránszki, J., 2015). Having caught the interest of businesses, government, and policy makers open data’s spread has necessitated public sector organizations to actively develop policies concerning the availability of data (Runardotter, M. & Ståhlbröst, A., 2016). These initiatives and policies are significant to the open data
How Continuous Improvement Can Support Logistics: A Reflection of Best Practices
www.igi-global.com/article/how-continuous-improvement-can-support-logistics/196601?camid=4v1a

Psycho-Aspects and Self-Determination on a Uniform Commonwealth Pedagogical Platform as a Scientific Category of Research Competence
www.igi-global.com/chapter/psycho-aspects-and-self-determination-on-a-uniform-commonwealth-pedagogical-platform-as-a-scientific-category-of-research-competence/196476?camid=4v1a