

IAF EARTH OBSERVATION SYMPOSIUM (B1)
Earth Observation Data Management Systems (4)

Author: Prof. Mariel Borowitz
Georgia Institute of Technology, United States, marielborowitz@gmail.com

WHAT HAPPENS WHEN OPEN DATA GETS BIG? OPPORTUNITIES AND RISKS FOR EARTH
OBSERVATION**Abstract**

Many space and meteorological agencies around the world have been leaders in implementing open data policies with respect to their Earth observation satellites. This makes it possible for the environmental satellite data to be used to conduct research and advance science, inform and improve policy-making, and develop new applications and data products. Space agencies have found that open data policies have led to dramatic increases in data access and use, and significant economic benefits for society as a whole.

Satellite-operating agencies now find themselves at the forefront of a new trend: production of big data. Landsat 8, launched in February 2013, collected more data in its first two years of operation than Landsat 4 and 5 collected over their collective 32-year lifespan. The NASA-ISRO Synthetic Aperture Radar Mission (NISAR), planned for launch in 2021, is expected double the size of NASA's Earth science data archive over its three-year mission. These developments provide exciting new opportunities, and scientists and others are eager to work with this new data.

However, the interaction of the open data and big data trends poses challenges to existing technical and economic models of data provision. The growth of open data initiatives was enabled by the wide availability of the internet, which made it possible to post data online which users could access and download at no cost to be analyzed on their own system. However, as data volume increases, neither data suppliers nor data users are able to quickly or easily upload, download, or analyze these very large datasets.

One potential solution to this technical challenge is the use of cloud computing, in which data is stored, accessed, and analyzed in a cloud environment, removing the need to actually transfer the data to the user. A transition to this new technology - whether government-built or commercially provided - thus raises a number of policy questions: should the government shoulder the costs of both storage and processing, or should some portion of these costs be paid by the user? If it becomes necessary for the user to pay a fee to access or use the data, can it still be considered "free and open" data? This paper examines these questions, comparing potential data provision models under the new big data paradigm, examining how they compare to existing open data provision models, and exploring the potential implications of these new models for the future of open data provision.