Secure Cloud Computing for SDG Data Challenges

Paulo Cunha
AWS South America Head
Worldwide Public Sector Team
cunhapa@amazon.com
Amazon.com...

???

... AWS: the global leader in utility (aka cloud) computing
16 regions worldwide growing to **19** soon

**74 Global Points of Presence**
At AWS, Security is Job Zero

Familiar security model

Validated by security experts
Collaboration on Enhancements

Every Customer Benefits

Physical Security

Network Security

Platform Security

People & Procedures
Role of **compliance** and 3\textsuperscript{rd} party auditors

- Vendor claims alone are not good enough!
- Testing, auditing and certification by multiple teams of 3\textsuperscript{rd}-party pros provides needed proof
- Far more rigorous process than any gov’t agency or corporation could reasonably sustain
“You should probably start engaging with the idea that the cloud can be considerably more secure than the private cloud or your own data centre, and start engaging with the risks that are building in the spaces where you haven't moved to the cloud yet.”

Customer control of data in the cloud

• Extra-territoriality of data
  – Growing realization that physical location is meaningless for most levels of data classification
  – Strong and ubiquitous encryption plus transparency via logging/auditing provide first-class isolation and control

• Privacy and cross-border data flows
  – CSPs provide virtual infrastructure, they are not “data processors” with direct access to data
  – Customers control all data (including encryption thereof) and data flows
“...data must be organized, well-documented, consistently formatted, and error free. Cleaning the data is often the most taxing part of data science, and is frequently **80% of the work**.”

— *Data Driven* by DJ Patil and Hilary Mason

**Cloud eliminates the “undifferentiated heavy lifting” behind data science**

Cloud democratizes access to data, *and* the computing power needed to process it

When (cleansed and curated) data is shared in the cloud, **anyone** can analyze **any volume of data** without needing to download or store it themselves.
Public Datasets on AWS

High-value datasets available for anyone to access at no cost on AWS. We use the AWS Public Datasets program to explore new ways to stage data for analysis in the cloud. Examples include:

- 3K Rice Genome
- Landsat on AWS
- NEXRAD on AWS

Info at https://aws.amazon.com/public-datasets
Thank you!