This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

SAND2022-9391C

Timeline

FOR PRESENTATION DEVELOPMENT

Please note that these dates are HARD deadlines. A failure to meet any of the above deadlines may result in incomplete or incorrect presentations available at the final event.

June 6 Dry run #1 with initial presentation in template

June 15 Dry Run #2 with presentation in template

June 21 Presentation is sent to designer

VERSION CONTROL: *no edits may be made by the presenter during this time without designer communication*

July 5 Designer sends back presentation

July 7 Dry Run #3

July 12 Final edits due to designer for compilation

R&A: *please send your presentation through R&A at this time as designer will not be altering any content, and provide the designer with your SAND number when it is approved.*

July 18 Final edits returned to presenters

July 19 Compile final presentations

July 21 **SANDIA TALKS**

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

SANDIA TALKS

Sharing is caring:

The open data initiative

at Sandia Labs

opendata.sandia.gov (Sandia Open Data Access)

Tony Alberti, Org. 6376

What is open data?

Open data is data that can be freely used, re-used and redistributed by anyone*

* https://opendatahandbook.org/guide/en/what-is-open-data/

What is Sandia open data?

Open data is data that can be freely used,
re-used and redistributed by anyone on the SRN

Shared Pain Points

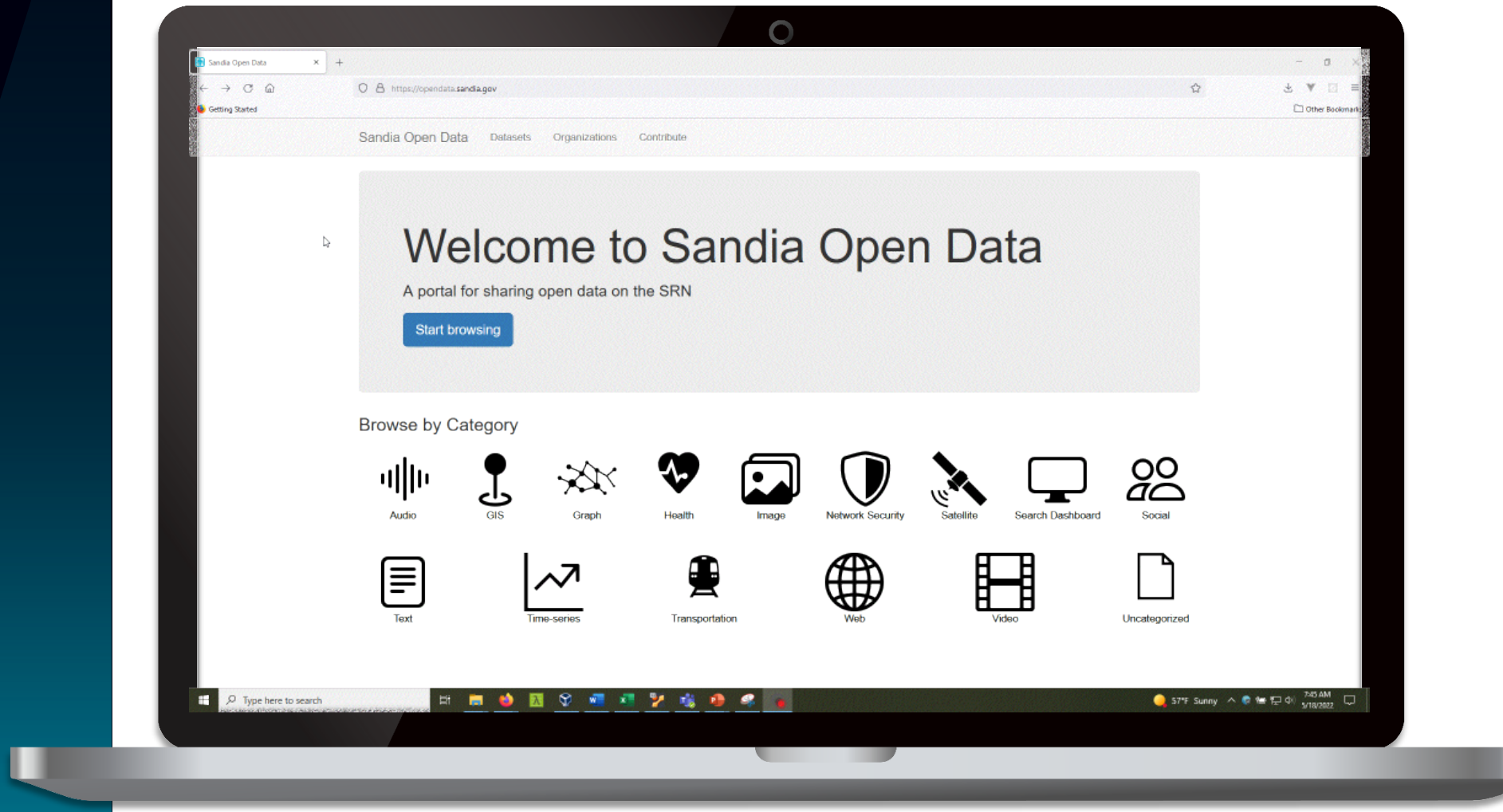
Image courtesy Taylor Vick

Shared Pain Points

Image courtesy Taylor Vick

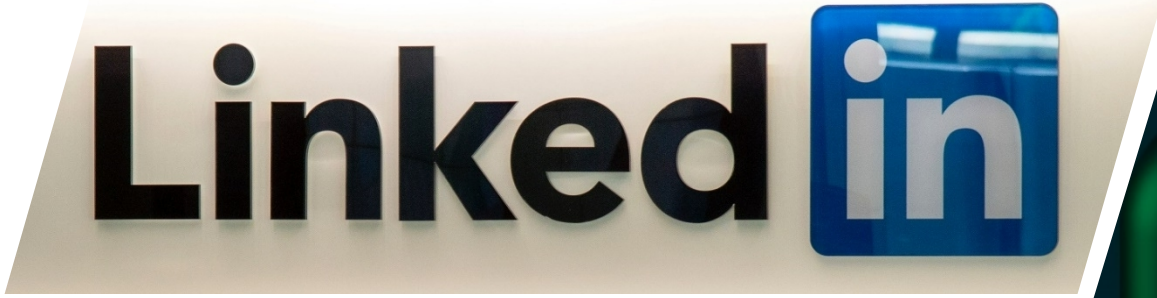
opendata.

sandia.gov



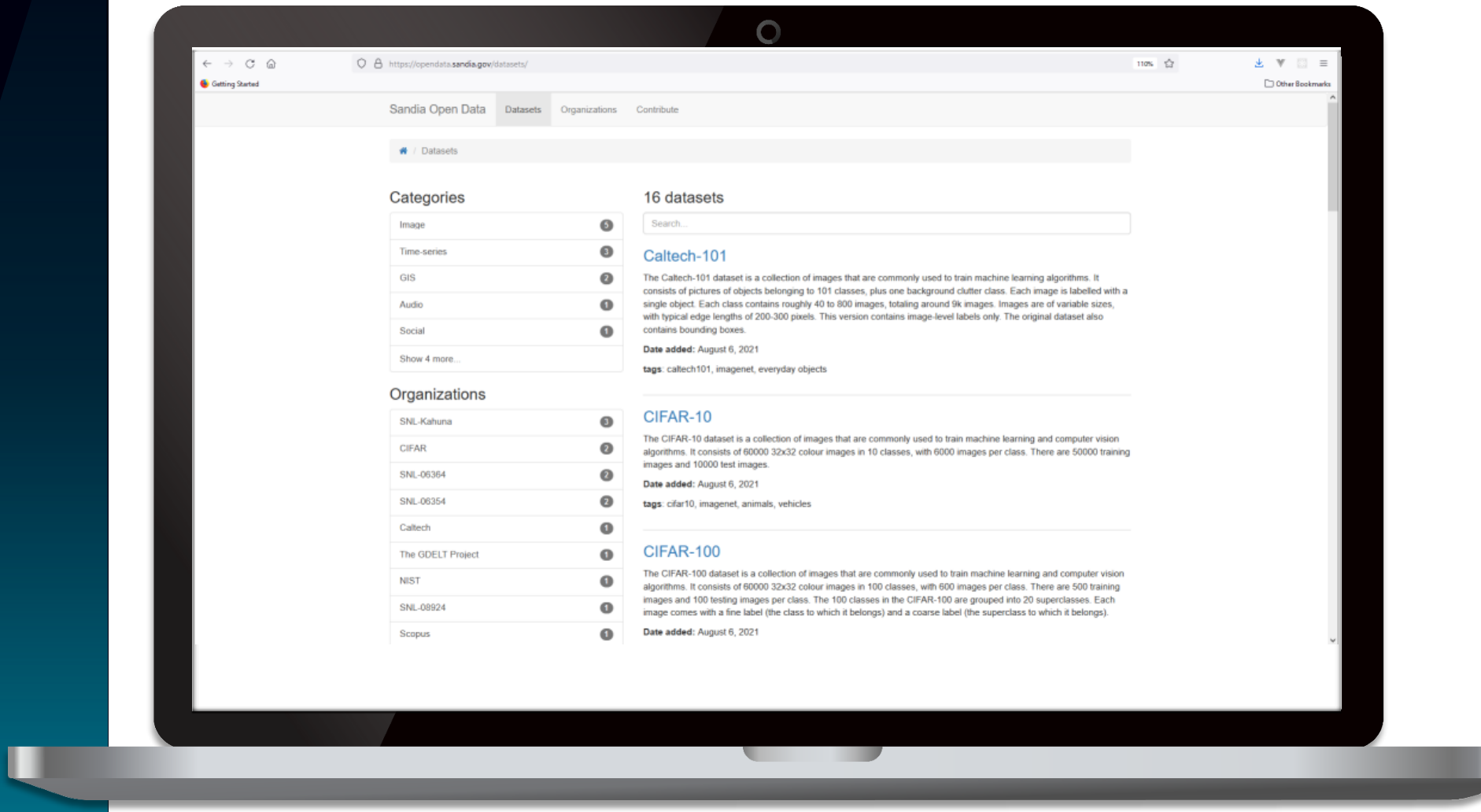
How did big tech address this?

- DataHub (LinkedIn)
- Amundsen (Lyft)
- Marquez (WeWork)
- Databook (Uber)
- Lexikon (Spotify)



Images courtesy of Thought Catalog, Greg Bulla, P.L., Alexander Shatov

Low Maintenance & Simplicity



Bettering the Platform

CONTINUED OUTREACH

Reaching out to potential users, **investigating** what users want in the platform, and **making informed decisions** on what we believe should be incorporated from those interviews

Bettering the Platform

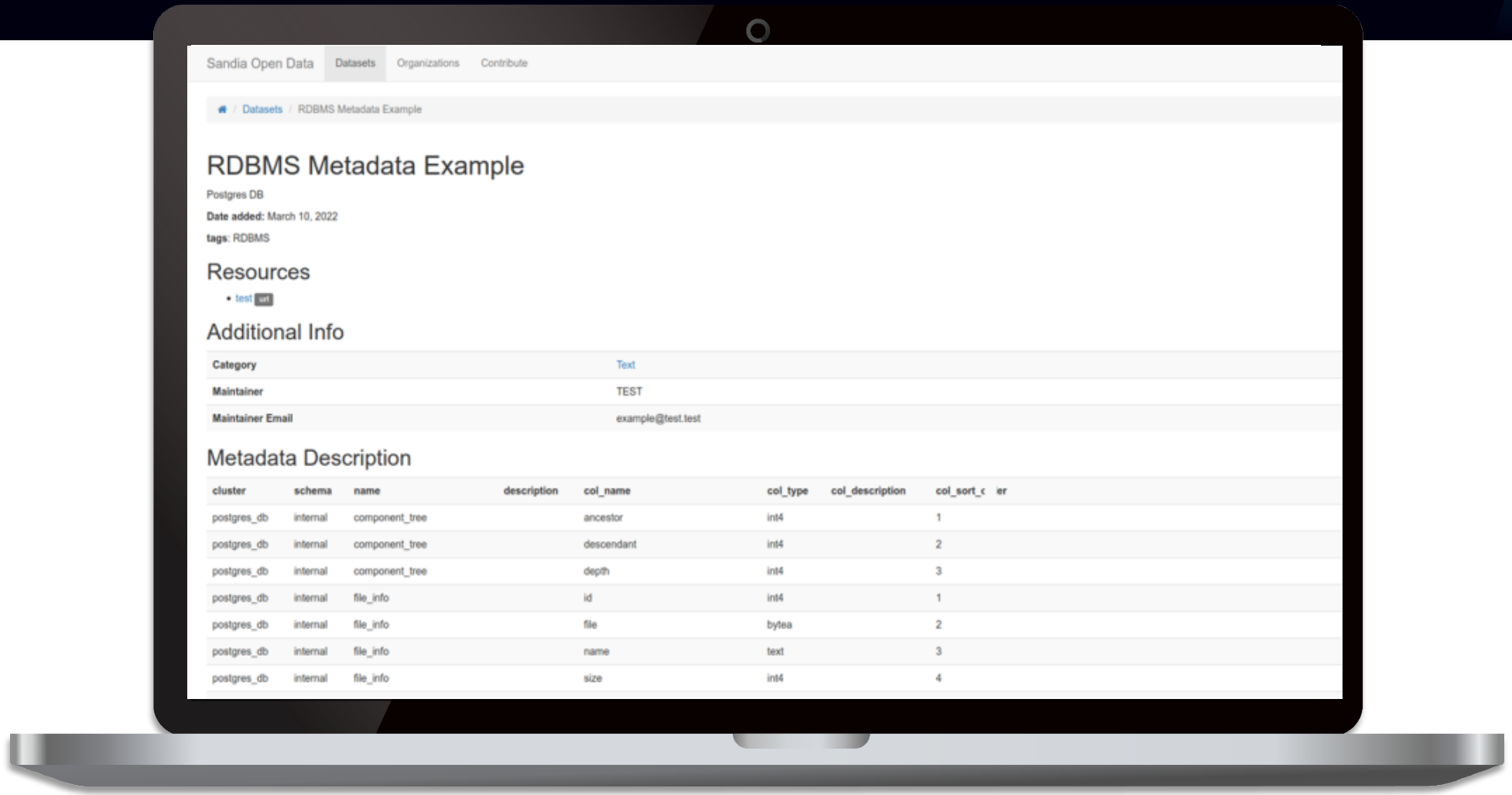
CORPORATE STORAGE

Corporate provided petabyte storage available.

Could be used as an option if people don’t have a place to host their data.

Bettering the Platform

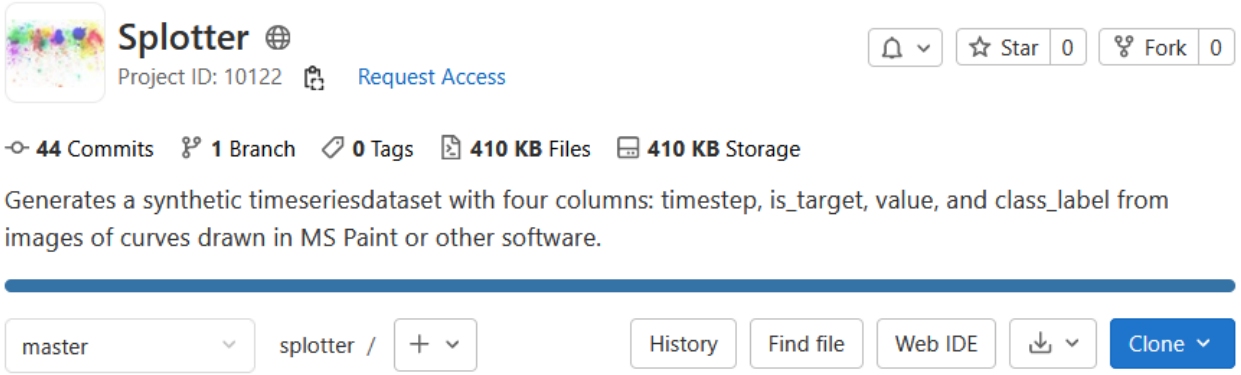
RDBMS METADATA



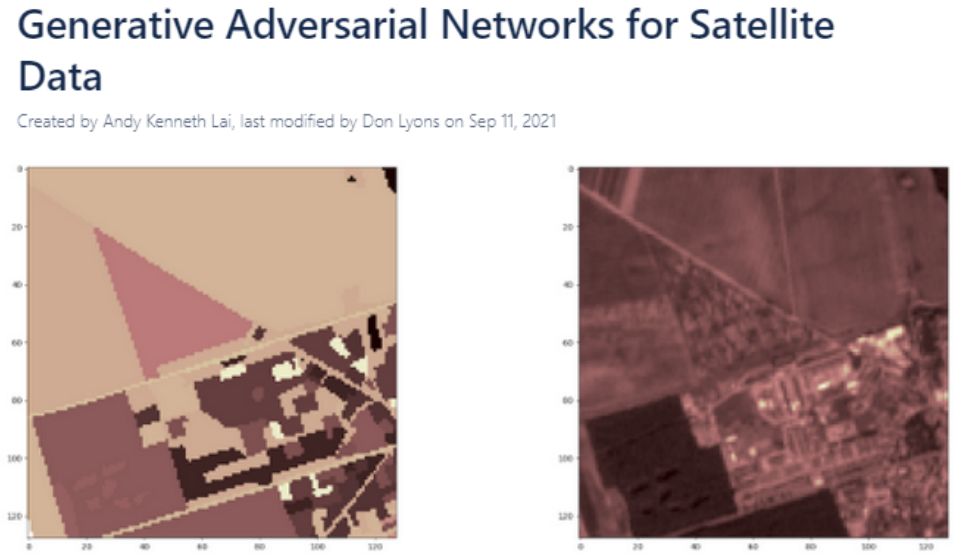
Bettering the Platform

SYNTHETIC DATA GENERATION

Created by Sarah Ackerman



Created by Andy Lai and Dalton Cole



Takeaways

Open data solutions **reduce friction in day-to-day work**, and a good way to implement one is to have a single entry point for finding all open data.

Takeaways

CONTINUED

Large tech companies have invested resources and formed entire teams coming up with solutions of how to best tackle this problem.

Bottom line: These pain points are shared from all organizations that depend on data to do their jobs.

Takeaways

CONTINUED

Having a **lightweight architecture** with **low barrier to entry** where users can **share open or even slightly access-controlled datasets** is how *opendata.sandia.gov* alleviates these pain points

Plug

We are **continuously searching for relevant data to add** to the platform. If you have:

- Unstructured data
- Processed/cleaned/labeled/derivative data from popular datasets
- A relational database you'd like to share metadata from
- Synthetic data generation tools
- Anything else that matches the decision tree from the *Contribute* page

Let us know!

https://opendata.sandia.gov/contribute/

SANDIA TALKS

THANK YOU

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories

SANDIA TALKS

NEXT TALK WILL BEGIN AT