Developing a cross-institutional open-source cyberinfrastructure to explore, analyze, and communicate greenhouse gas data and information

Speaker
Siddharth Chaudhary
NASA-IMPACT/UAH
siddharth.chaudhary@nasa.gov

Team: Siddharth Chaudhary, Deborah K. Smith, Jeanné le Roux, Slesa Adhikari, Vishal Gaur, Shanna Combley, Manil Maskey, Argyro Kavvada

IMPACT Interagency Implementation and Advanced Concepts Team
THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

What is the U.S. Greenhouse Gas (GHG) Center?

Vision

Inspire and accelerate the use of Earth science data and information for GHG monitoring, measurement, reporting and verification for decision support.

Mission

Extend accessible and integrated GHG data and modeling capabilities from U.S. Government and non-public sources for scalable impact.

Purpose

- Develop and enhance GHG data products that **meet user needs**;
- Foster c**ollaboration** with networks of interagency, international, intergovernmental and private sector partners to co-develop and increase adoption of **impactful applications;**
- Promote **scientific innovation** and transparency by leveraging **advanced data systems** capabilities and **open source science** principles; and,
- Establish **bidirectional knowledge transfer** and engagement with federal, state, local and tribal governments, researchers, and the general public.

Launch of U.S. Greenhouse Gas (GHG) Center

NASA, Partners Launch US Greenhouse Gas Center to Share Climate Data

NOVEMBER 29, 2023

Interagency Working Group Releases National Strategy to Enhance the Nation's Greenhouse Gas Measurement and Monitoring Capabilities

Climate Change

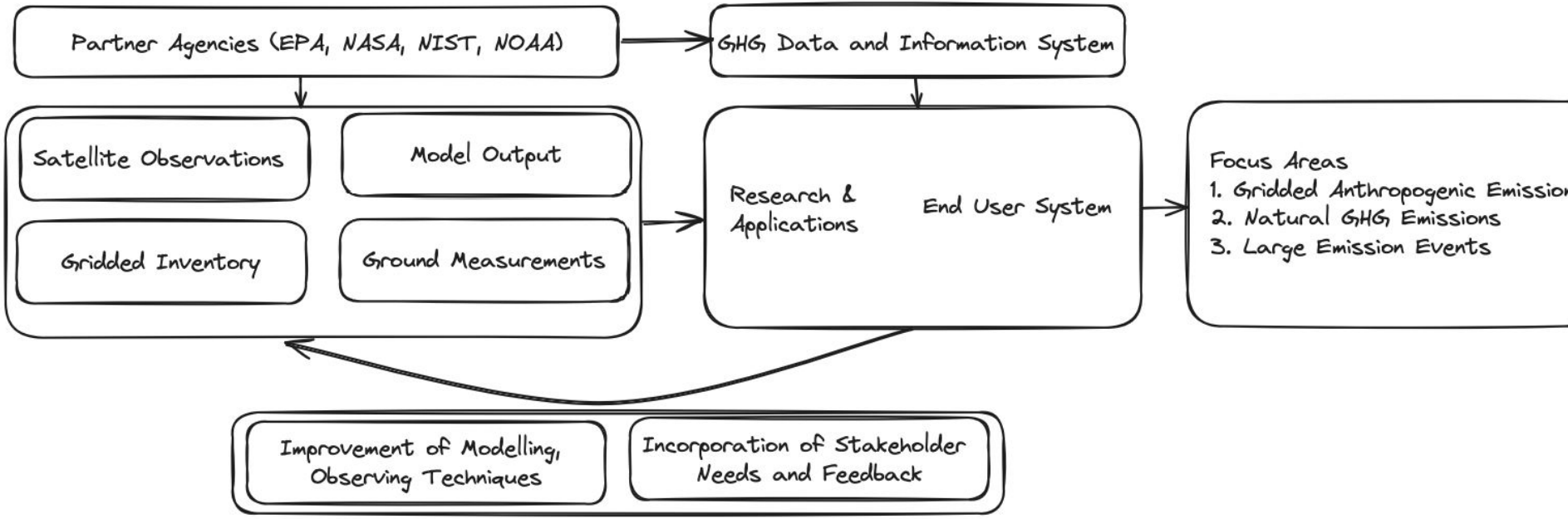
NASA's new 'Greenhouse Gas Center' track humanity's contribution to climate change

News By Meredith Garofalo published December 11, 2023

NASA, EPA share plans for greenhouse gas initiative at COP28

Natalie Harms Dec 7, 2023, 2:24 pm

U.S. Greenhouse Gas (GHG) Center



U.S. Greenhouse Gas (GHG) Center

It's providing **new scientific capabilities and insight**

- Science teams are working to provide the latest GHG data; collaborating with partners to bring together GHG data from diverse sources & programs

It's a **data and information system**

- Building a system that supports the GHG Center mission (web presence, data catalog, visualizations, data interactions, access, analysis capabilities, disclaimers)

It's building **new connections**

- The Stakeholder Engagement team is working to foster relationships & promote the GHG Center

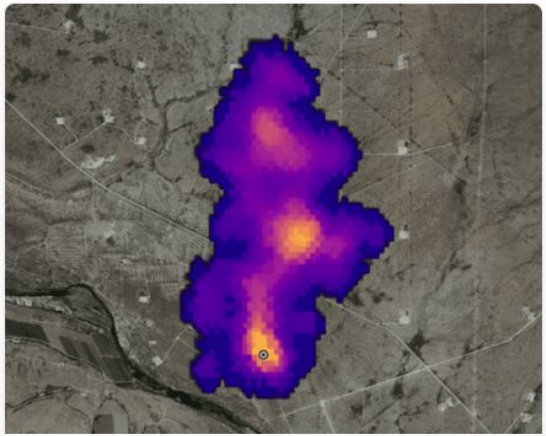
Use Cases/Demonstration Areas

Gridded Anthropogenic Greenhouse Gas Emissions

Emission estimates from human activities including the energy, agriculture, waste, and industry sectors

Natural Greenhouse Gas Sources and Sinks

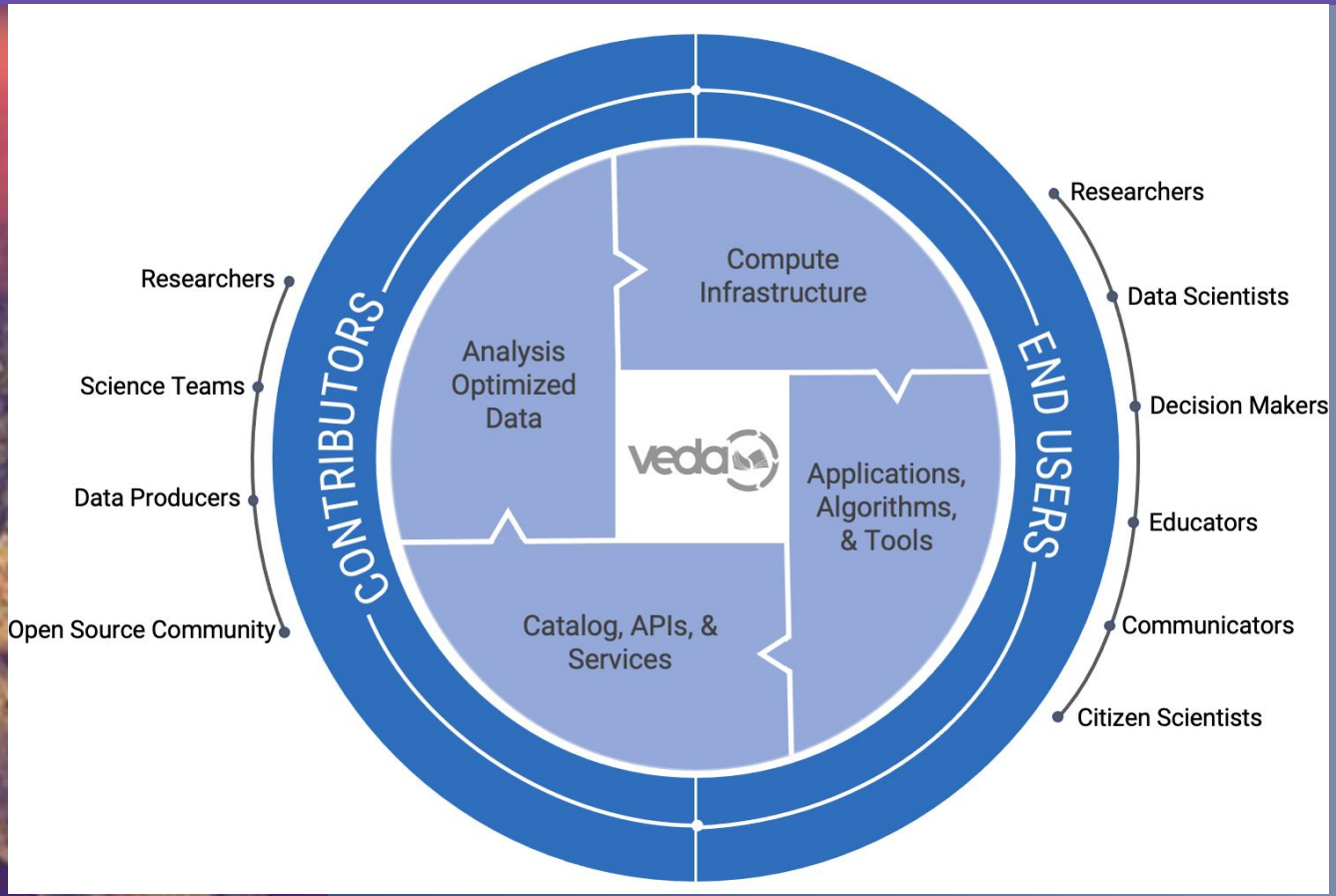
Naturally-occurring greenhouse gas fluxes from land, ocean, and atmosphere



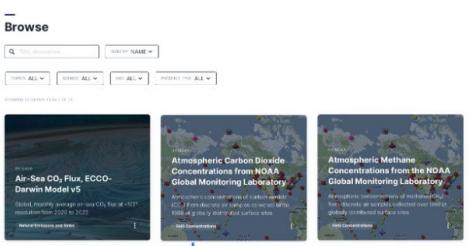
New Observations for Tracking Large Emission Events

Identify and quantify large methane leak events leveraging aircraft and space-based data

Leveraging Visualization, Exploration, and Data Analysis Platform

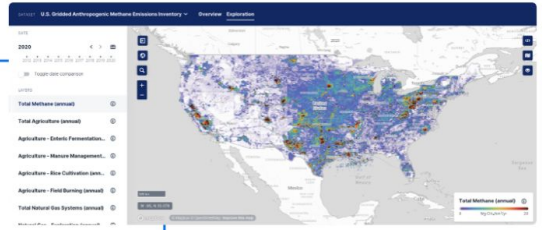


VEDA - Reusable in part or in whole



EXPLORE

- Map Exploration
- Data Access



RESEARCH

- Analysis on the fly
- Analysis Hub
- Notebooks
- QGIS

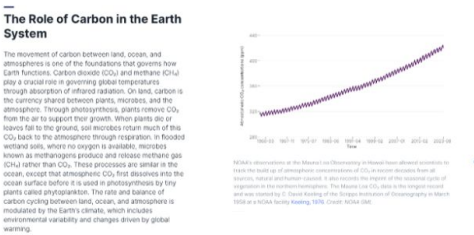
Open VEDA Platform

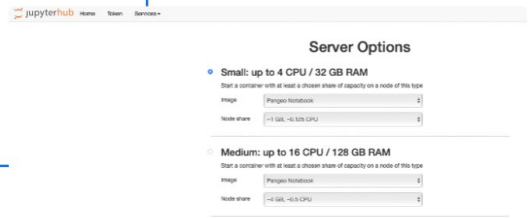
PUBLISH

- Data Catalog

COMMUNICATE

- Data Insights
- Data Driven Storytelling





U.S. GHG Center Data and Information Management - Beta

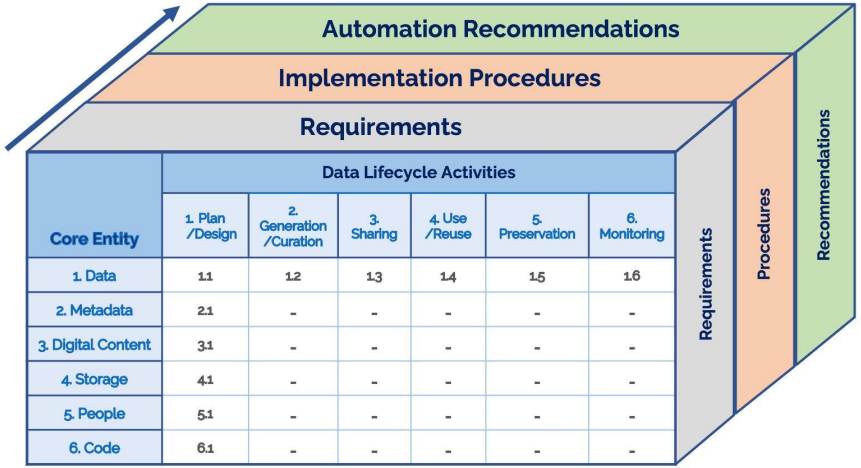
1. Data Governance Framework
2. GHG Center System Data Flow
3. Gathering Dataset Information
4. Verifying Data Ingestion
5. Data Incorporation

U.S. GHG Center Modern Data Governance Framework (mDGF)

Developed a new modern Data Governance Framework (mDGF), designed to organize and simplify data stewardship for projects

All aspects of data handling have been planned and are in the process of being implemented

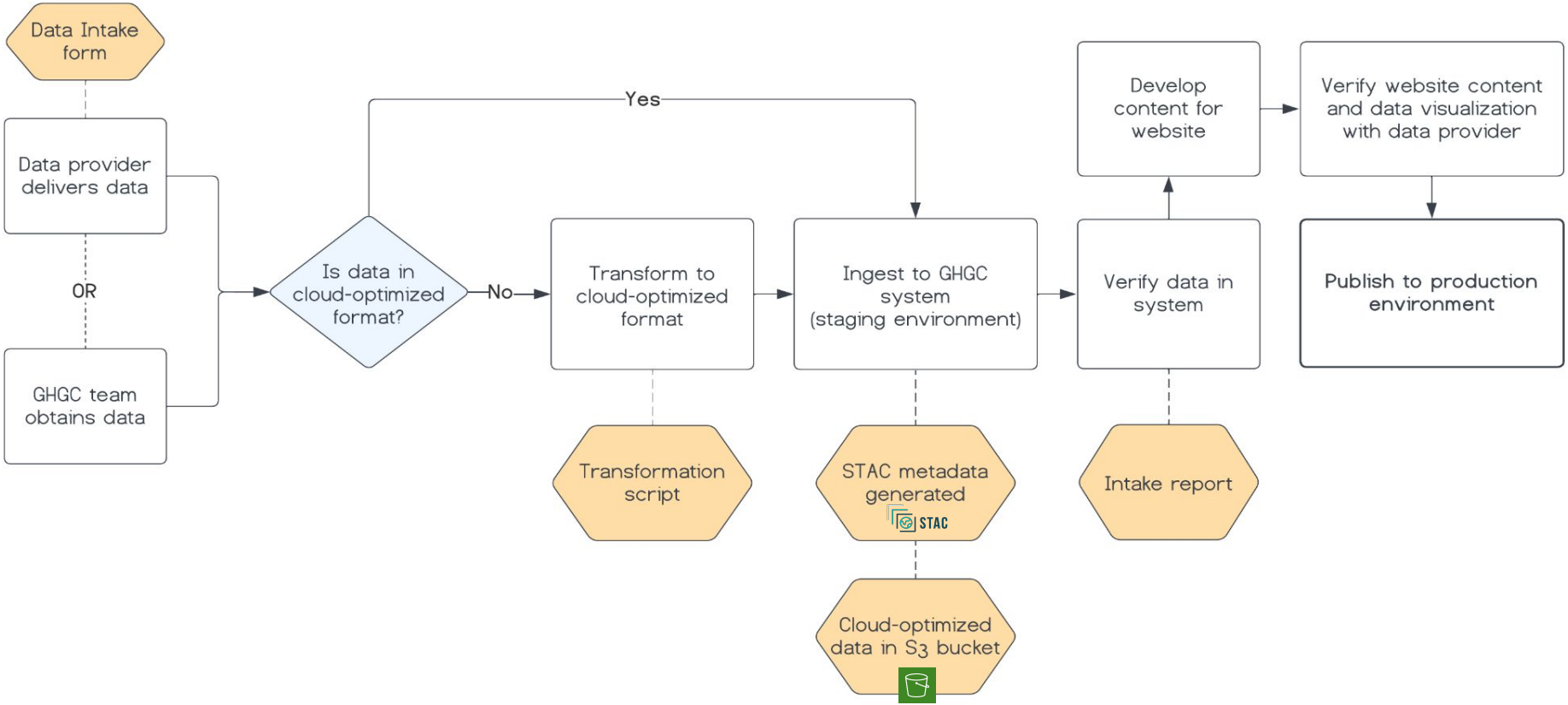
The mDGF is a data management plan, a resource management plan, a software management plan all in one



Core Entity	Data Lifecycle Activities					
	1. Plan /Design	2. Generation /Curation	3. Sharing	4. Use /Reuse	5. Preservation	6. Monitoring
1. Data	1.1	1.2	1.3	1.4	1.5	1.6
2. Metadata	2.1	-	-	-	-	-
3. Digital Content	3.1	-	-	-	-	-
4. Storage	4.1	-	-	-	-	-
5. People	5.1	-	-	-	-	-
6. Code	6.1	-	-	-	-	-

Entities and Lifecycle Activities relevant to a given project can be selected to customize the framework

Data Flow Diagram



Data Intake Form

Greenhouse Gas Center (GHG Center) Data Intake Information Form

The purpose of this document is to collect important information about datasets to be incorporated into the GHG Center from the data provider. The *gray italicized text* provides instructions for how to fill out each item.

Dataset Name
The full/formal title of the dataset.

Short Name/Unique Dataset ID
Shortened or abbreviated dataset title or other dataset id. Note: the original short name may be updated to reflect processing applied for GHG Center (e.g. MOD4A > MOD4A-cog). If the dataset is cataloged in the Common Metadata Repository (CMR) and discoverable in Earthdata Search, also provide the unique collection identifier for the dataset. E.g. C1214470488-ASF

Dataset POC
Name of the primary point of contact for the dataset. The GHG Center may reach out to this POC for questions or additional dataset information.

POC Email
Dataset POC email address.

POC Affiliation
Dataset POC organization/affiliation.

Access URL
The link(s) for accessing the dataset (e.g. https or s3).

Access Instructions/Credentials
Provide any additional instructions needed to access and read the data (e.g. directory structure, access credentials, read routines).

Version
Dataset version number.

Processing Level
The data processing level or processing level description. If possible, map to the equivalent EOSDIS data processing level:
https://www.earthdata.nasa.gov/engage/open-data-services-and-software/data-information-policy/data-levels

Data Maturity Level
Indicate the data maturity level (Beta, Provisional, Validated Stage 1-4) based on the ESDS data maturity level criteria.

Production Status (Static/Active)
Indicate whether the dataset is complete/static or active (e.g. new data routinely added to the dataset from an operational satellite or sensor).

Planned Version Update
Describe any plans for data reprocessing or version updates. Include a timeline, if known (approximate

Gather important information about a data product needed to appropriately portray it in the GHG Center

Data Processing and Confirmation Report (Intake Report)

Processing and Confirmation Report

Background Information

- Original Dataset Name: ODIAC Fossil Fuel Emission Dataset
- GHGC Dataset Title: ODIAC Fossil Fuel CO_2 Emissions
- Dataset Provider: NASA/NIES
- Date Obtained: 06/01/2023
- Location Obtained from: https://db.cger.nies.go.jp/dataset/ODIAC/DL_odiac2022.html
- Data location in GHGC: odiac-fossil-fuel-co2-emissions-2022
- Data POC(s): Dr. Tomohiro Oda, Dr. Lesley Ott
- Dataset File Type(s): GeoTIFF
- Projection (if different from WGS84): NA

Data Transfer Confirmation

An SHA-256 checksum is used to detect high-level errors within data transmissions

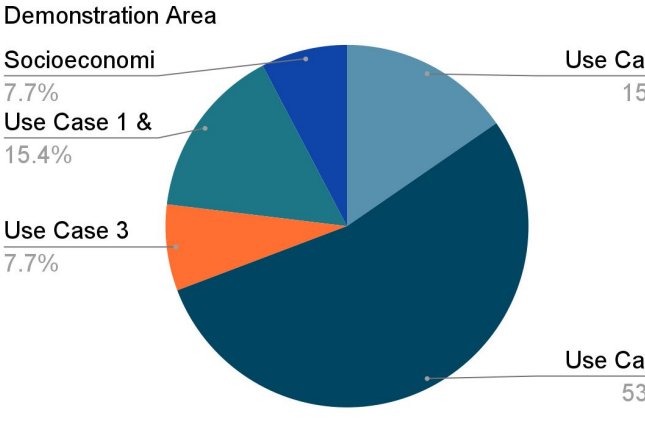
- Results from individual checksum file comparisons of pre-transfer and post-transfer
  - All files were transferred successfully

Filename	SHA 256 Original file	

- Report any individual file issues: NA

Report for data providers that documents the data ingestion process and data verification outcomes

Dataset Name	Source	Demonstration Area
Gridded Anthropogenic Methane Emissions Inventory	EPA	1
ODIAC Fossil Fuel CO_2 Emissions	NASA/NIES	1
Atmospheric Carbon Dioxide Concentrations from the NOAA Global Monitoring Laboratory	NOAA	1 & 2
Atmospheric Methane Concentrations from the NOAA Global Monitoring Laboratory	NOAA	1 & 2
Wetland Methane Emissions, LPJ-wsl Model	NASA	2
CASA-GFED3 Land Carbon Flux	NASA	2
OCO-2 MIP Top-down CO_2 Budget	NASA/NOAA	2
GOSAT-based Top-down Total and Natural Methane Emissions	NASA	2
OCO-2 GEOS Column CO_2 Concentrations	NASA	2
Air-Sea CO_2 Flux, ECCO-Darwin Model v5	NASA	2
TM5-4DVar Isotopic CH_4 Inverse Fluxes	NASA/NOAA	2
EMIT Methane Point Source Plume Complexes	NASA	3
SEDAC Gridded World Population Density	NASA	Socioeconomic

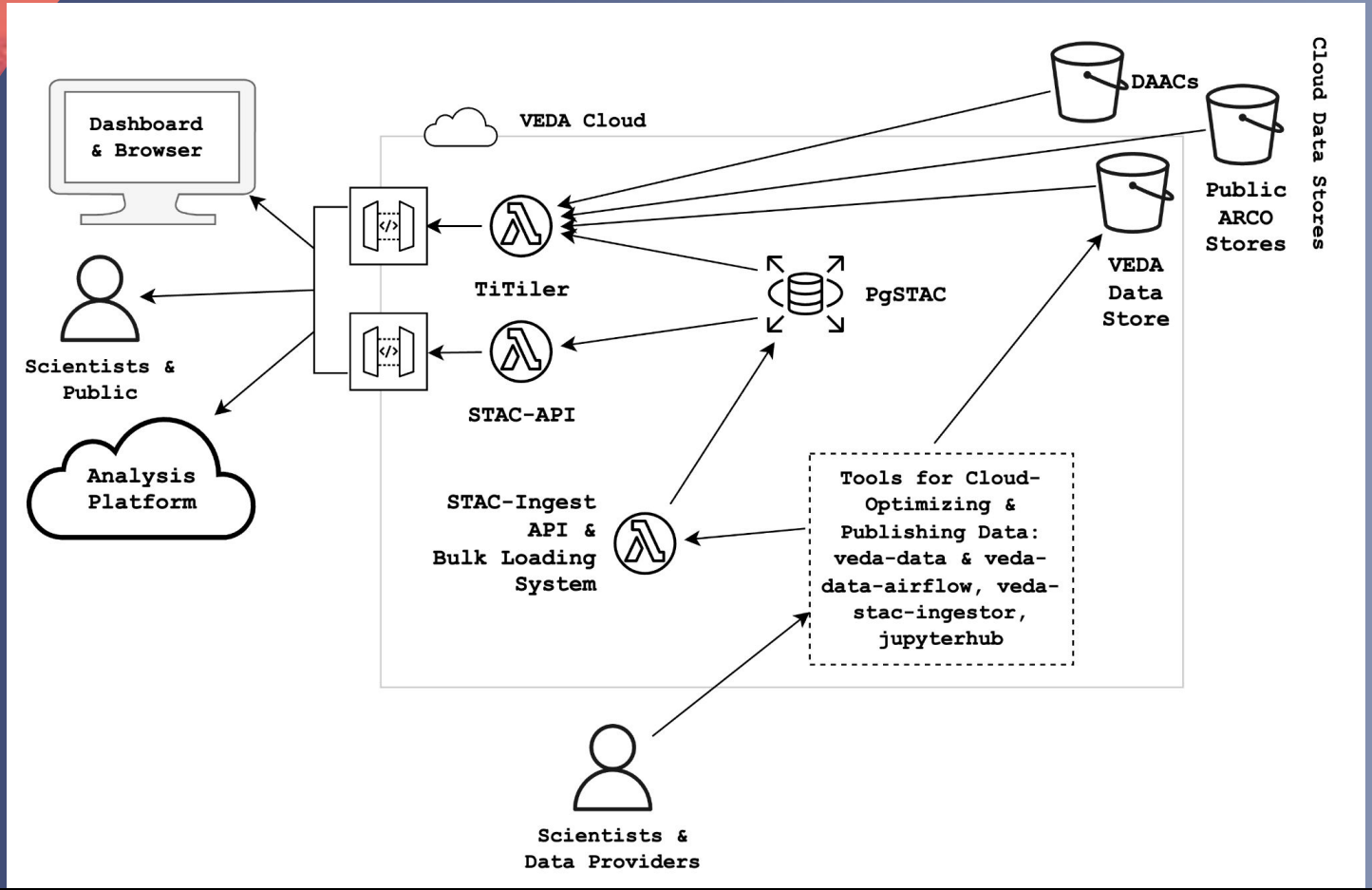


System Demonstration

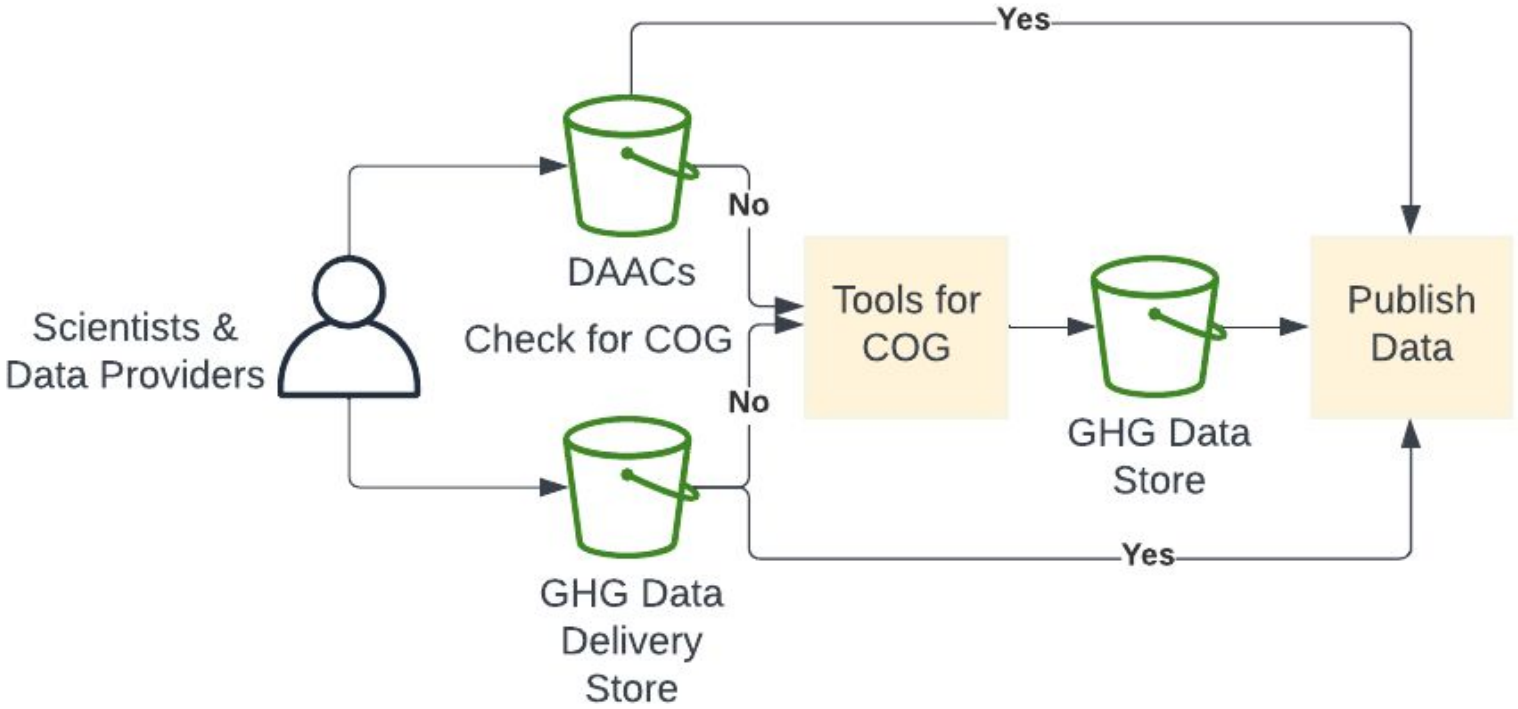
1. Home Page
2. Browse Data Catalog (Filters, Use case, Tags)
3. Curated Data (Intro about data)
4. Access Data
5. Explore Data (Comparison, legend, Integration with QGIS)
6. Data Insights (1 story)
7. Analysis
8. Hub
9. Contact Us

Pr-recorded video

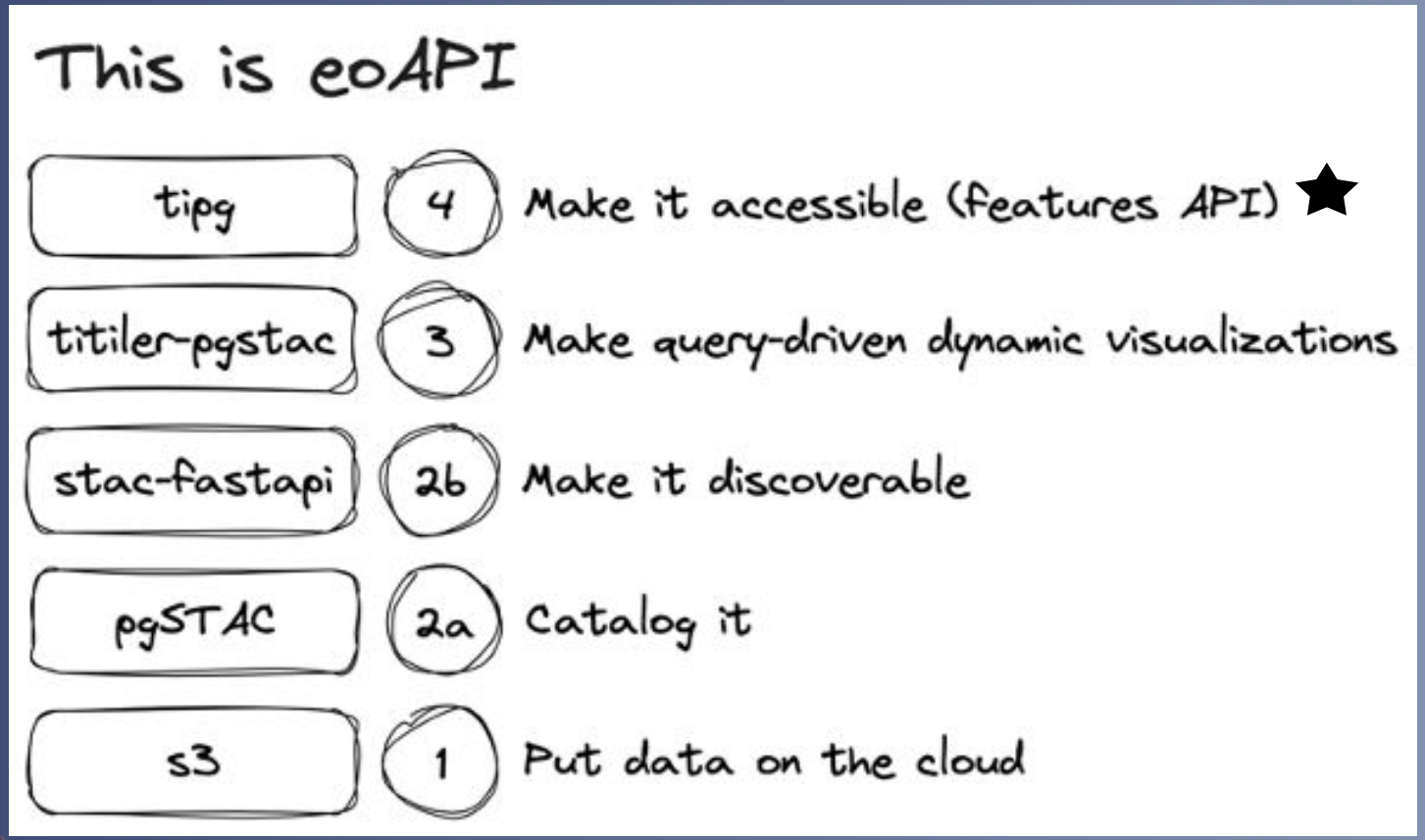
U.S. GHG Center Infrastructure



Cloud Data Store



eoAPI: an open-source framework for working with Earth Observation data



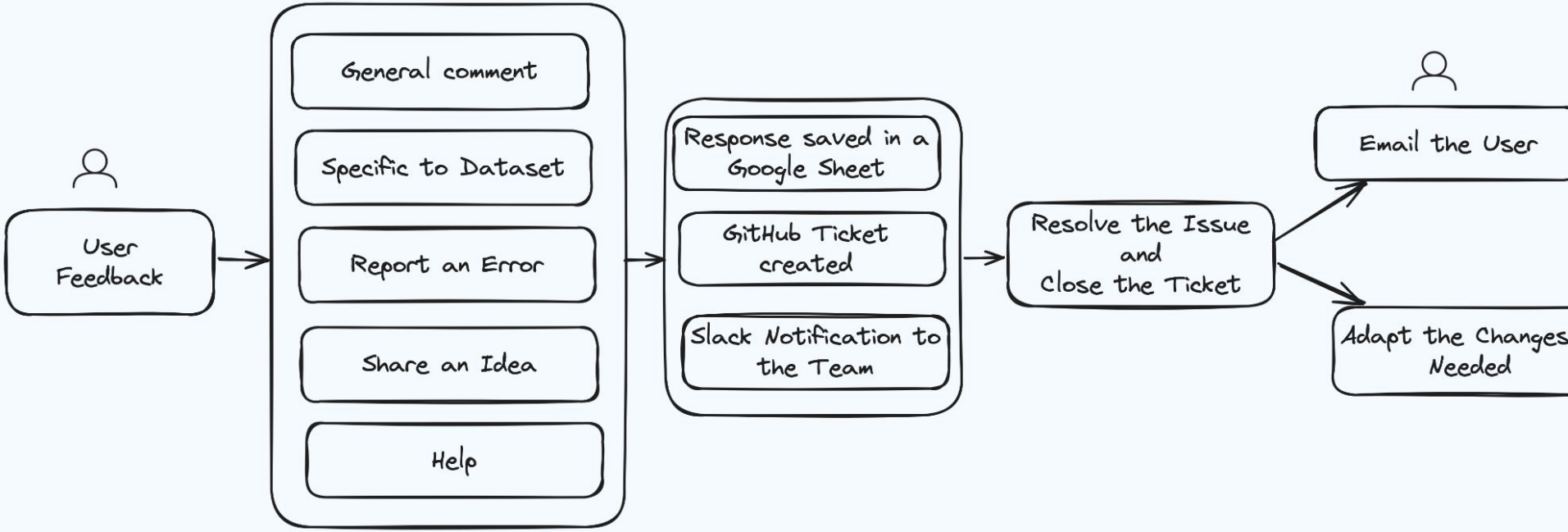
Data Analysis Hub - Computational Capabilities

GIF

US GHG Center supports Open Science

- Data is open
  - Data, it's visualization and analysis, - everything is open!
- Methods are open
  - Data transformation code
  - Data validation reports
  - Data analysis code
- Infrastructure is open
  - The computing infrastructure to do scientific analysis is open - via JupyterHub!
- Software is open
  - Via public GitHub repositories!
- Educational resources are open
  - Via the web portal, via notebooks, via workshops like these!

User Feedback



Upcoming Features

1. Update in Explore and Analysis Feature
2. Datasets for different use cases in release 2

Thank You

siddharth.chaudhary@nasa.gov

This research is supported by NASA Grant NNM11AA01A as part of the IMPACT project.