An Open Science Platform

Built with an open source software community

Jonas Sølvsteen, Aimee Barciauskas, Vincent Sarago
Development Seed/NASA IMPACT

FOSS4G 2023 June 26, 2023

Credits:

VEDA Leadership Team: Manil Maskey, Brian Freitag, Anthony Boyd, Jonas Sølvsteen

VEDA Development Team: Daniel da Silva, Alexandra Kirk, Greg Corradini, Hanbyul Jo, Ricardo Mestre, Erik N, Isayah Vidito, Leo Thomas, David Bitner, Vincent Sarago, Julia Signell, Abdelhak Maroune, Vishal Gaur, Nathan Zimmerman, Slesa Adhikari, Justin Polchlopek, Saadhiq Mohiuddun, Siddharth Chaudry

Indiana University, EIS, 2i2c, SMCE

Agenda

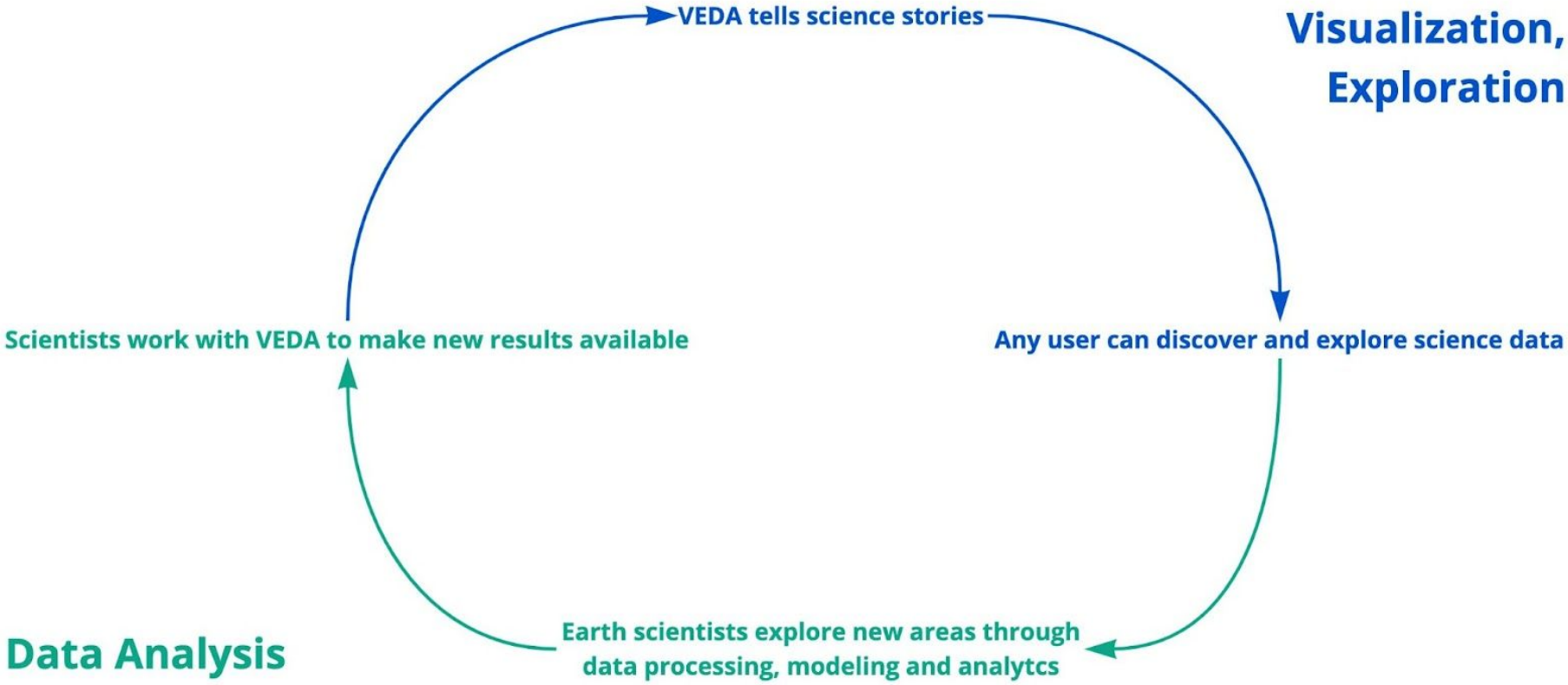
- VEDA in relation to EO Dashboard and Euro Data Cube
- Overview of VEDA
- Demo of creating content in the VEDA Dashboard: Fire Tracking Story
  - VEDA is an open science platform, not just a story telling dashboard
  - Scientists can:
    - Scale their science
    - Contribute documentation
    - Contribute datasets
    - Contribute discoveries
  - Users can reproduce analysis via VEDA's Jupyterhub.
- How to deploy VEDA
  - eoAPI + veda-ui + veda-config = Dashboard
  - Dashboard + JupyterHub = VEDA as an open science platform
- Questions?

VEDA in relation to EO Dashboard

- Through a trilateral collaboration ESA-JAXA-NASA, the 3 agencies exhave been exploring how to share content and reach users to benefit from the combined wealth of Earth observation information.
- VEDA is the NASA platform for data storytelling with open-source software and open-standard interfaces.
- VEDA provides NASA data to render in the EO Dashboard via STAC and dynamic tiling (TiTiler).

Overview of VEDA

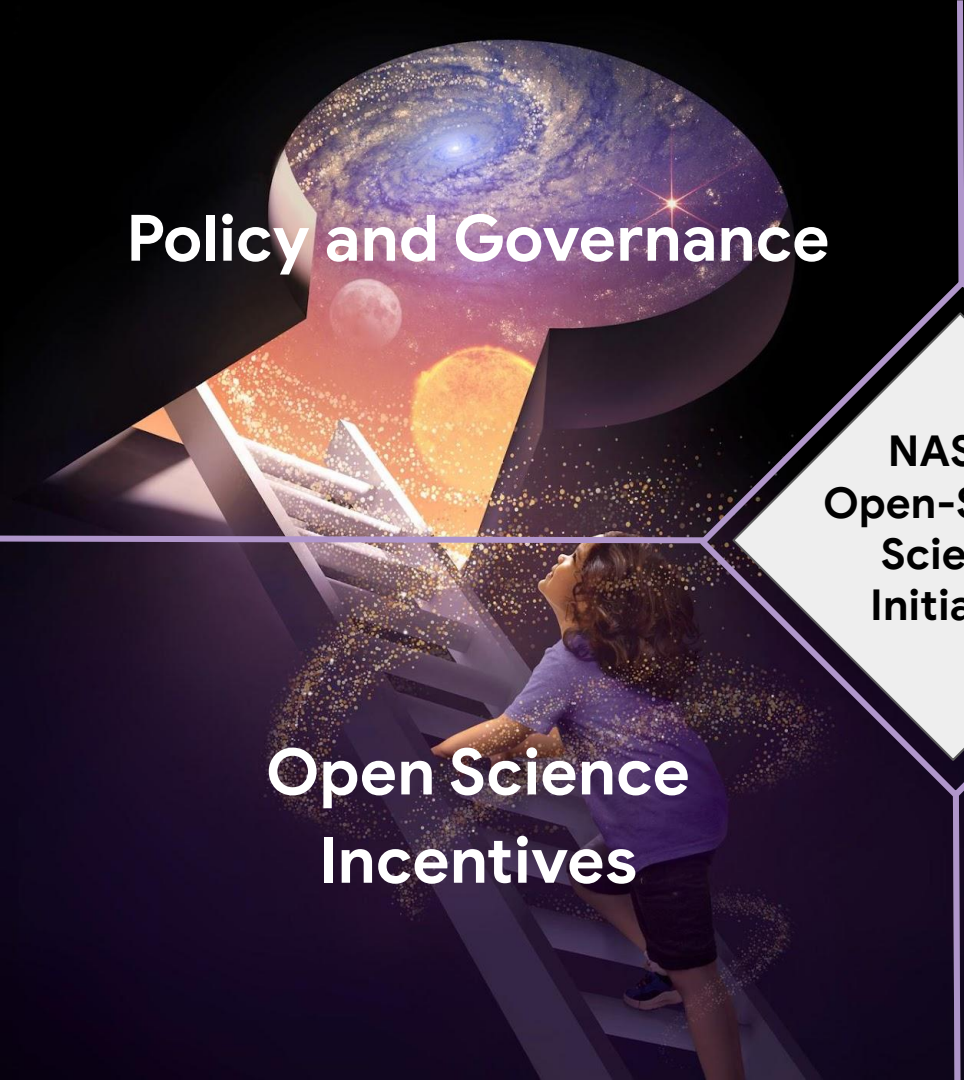
The VEDA Science Lifecycle



NASA's Open-Source Science Initiative

NASA's approach for putting Open Science into practice.

https://science.nasa.gov/open-science-overview



Policy and Governance

Core Data and Computing Services

NASA's Open-Source Science Initiative

Open Science Incentives

Community Engagement

Overview of VEDA

4-Pillars of NASA open science

1. **Policy and Governance:** Implement policies that advance open science, support Science Mission Directorate (SMD) open science activities.
2. **Core Data and Computing Services:** Develop SMD-wide data and computing infrastructure.
3. **Open Science Incentives:** Grants, prizes and challenges to enable groundbreaking scientific discoveries using open science principles and tools.
4. **Community Engagement:** Advance open science practices in the SMD community and build strategic partnerships for innovation in open science.

VEDA puts these initiatives into practice.

NASA has always had open data policies.

VEDA is moving NASA forward in its open science policies.

VEDA attempts to make NASA's earth science data mean more. Data has been openly accessible for anyone to use, but NASA hasn't exposed it in friendly interfaces or analytics platforms.

VEDA is making Earth data computing more accessible.

This presentation is a part of our community engagement - we want to know from you!

Demo creating content in the VEDA Dashboard.

VEDA is an open science platform, not just a story telling dashboard.

The VEDA team worked with the EIS Fire team to automate and scale the Fire Event Data Suite (FEDS), near-real time fire perimeter tracking for the continental united states (CONUS).

Learn more: Nature article describing the algorithm developed by the EIS Fire team

“We develop a novel object-based system for tracking the progression of individual fires using VIIRS active fire detections”

EIS Fire FEDS Team: Douglas Morton, Melanie Follette-Cook, Yang Chen, Tempest McCabe, Elijah Orland, Jim Randerson, Lesley Ott, Alexey Shiklomanov, Shane Coffield

scientific data

Explore content | About the journal | Publish with us

nature > scientific data > data descriptors > article

Data Descriptor | Open Access | Published: 30 May 2022

California wildfire spread derived using VIIRS satellite observations and an object-based tracking system

Yang Chen, Stijn Hantson, Niels Andela, Shane R. Coffield, Casey A. Graff, Douglas C. Morton, Lesley E. Ott, Efi Foufoula-Georgiou, Padhraic Smyth, Michael L. Goulden & James T. Randerson

Scientific Data **9**, Article number: 249 (2022) | Cite this article

3686 Accesses | **5** Citations | **219** Altmetric | Metrics

Abstract

Changing wildfire regimes in the western US and other fire-prone regions pose considerable risks to human health and ecosystem function. However, our understanding of wildfire behavior is still limited by a lack of data products that systematically quantify fire spread, behavior and impacts. Here we develop a novel object-based system for tracking the progression of individual fires using 375 m Visible Infrared Imaging Radiometer Suite active fire detections. At each half-daily time step, fire pixels are clustered according to their spatial proximity, and are either appended to an existing active fire object or are assigned to a new object. This automatic system allows us to update the attributes of each fire event, delineate the fire perimeter, and identify the active fire front shortly after satellite data acquisition. Using this system, we mapped the history of California fires during 2012–2020. Our approach and data stream may be useful for calibration and evaluation of fire spread models, estimation of near-real-time wildfire emissions, and as means for prescribing initial conditions in fire forecast models.

VEDA supported the scaling of EIS fire tracking ability to near real time

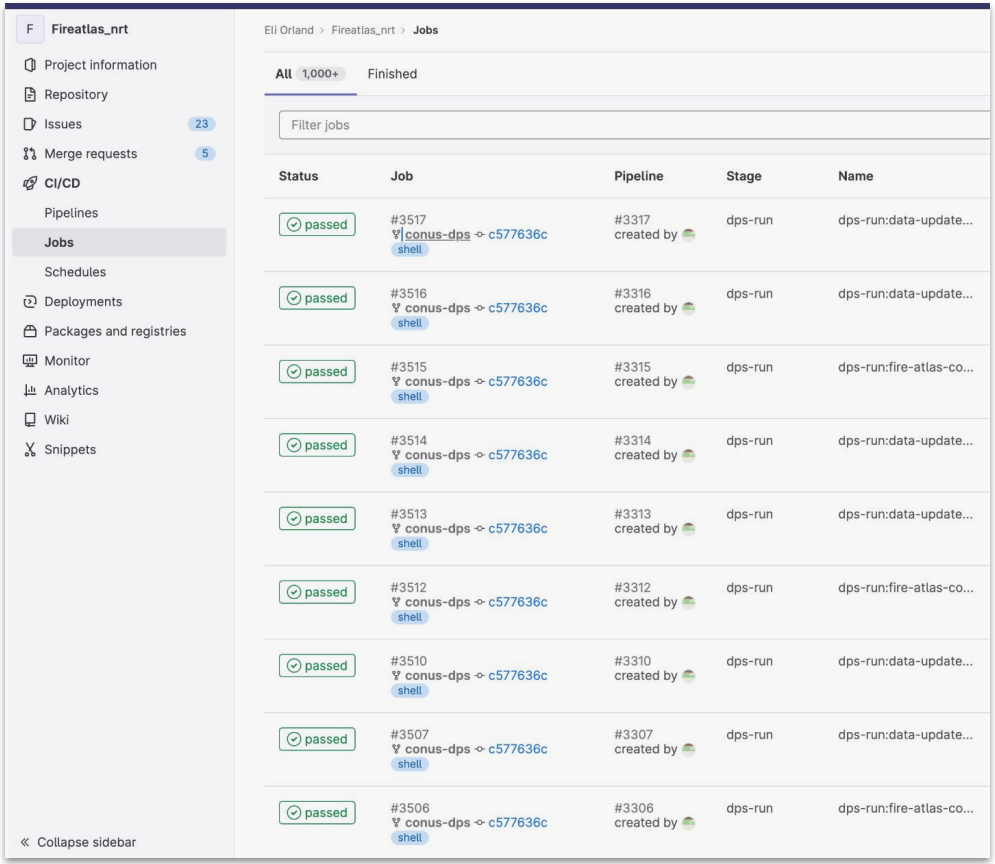
EIS fire team had an algorithm for tracking fire events in California.

They wanted to scale this algorithm to 1) run routinely in near real time and 2) expand the spatial domain to cover the Continental United States.

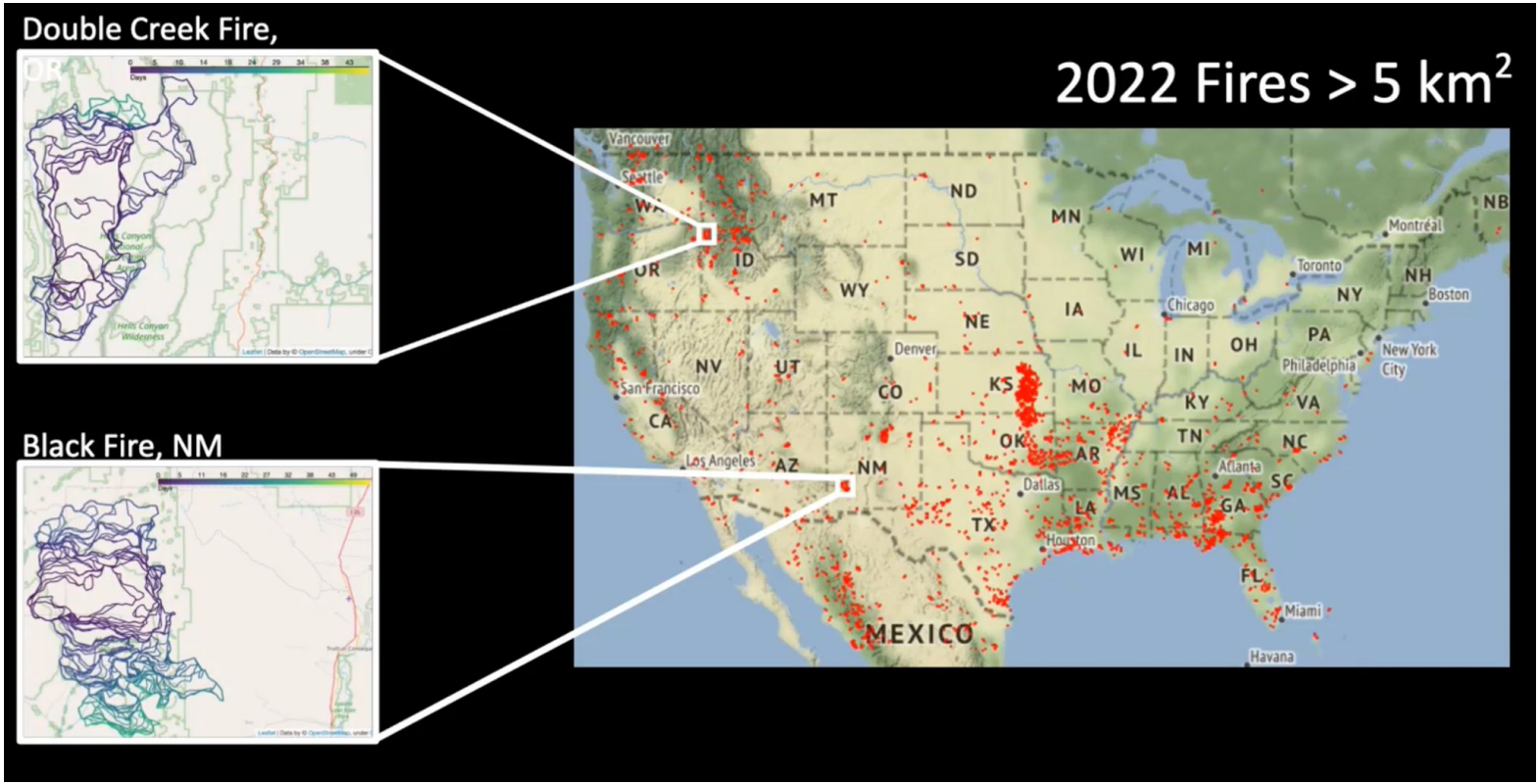
https://svs.gsfc.nasa.gov/4992

VEDA Analytics Platform Team supported the scaling of EIS fire tracking to near real time

- Crontab schedulers schedules DPS runs every 4 hours
- All of CONUS
- Output to the VEDA features API
- Kudos Greg Corradini, Alex M and Julia Signell



Where EIS Fire is now: Routinely producing fire perimeters in near real time for the continental United States (CONUS).



What was the impact of being able to scale?

FEDS data can be routinely access through an API. This enables the EIS Fire team to share the data with FEMA. It may be added to FIRMS as an experimental product.

The team is working with FEMA, USGS and Forest Service to develop value-added products such as fire direction, speed of fire spread, and fire severity metrics.

Science impacts: better understand of factors which contribute to fire intensity and air pollution.

- Melanie Follette-Cook

Cross-functional expertise was critical

“We would not have been able to do this without VEDA and its team.”

“This would not have been possible without research scientists across disciplines, data systems and cyber infrastructure experts.”

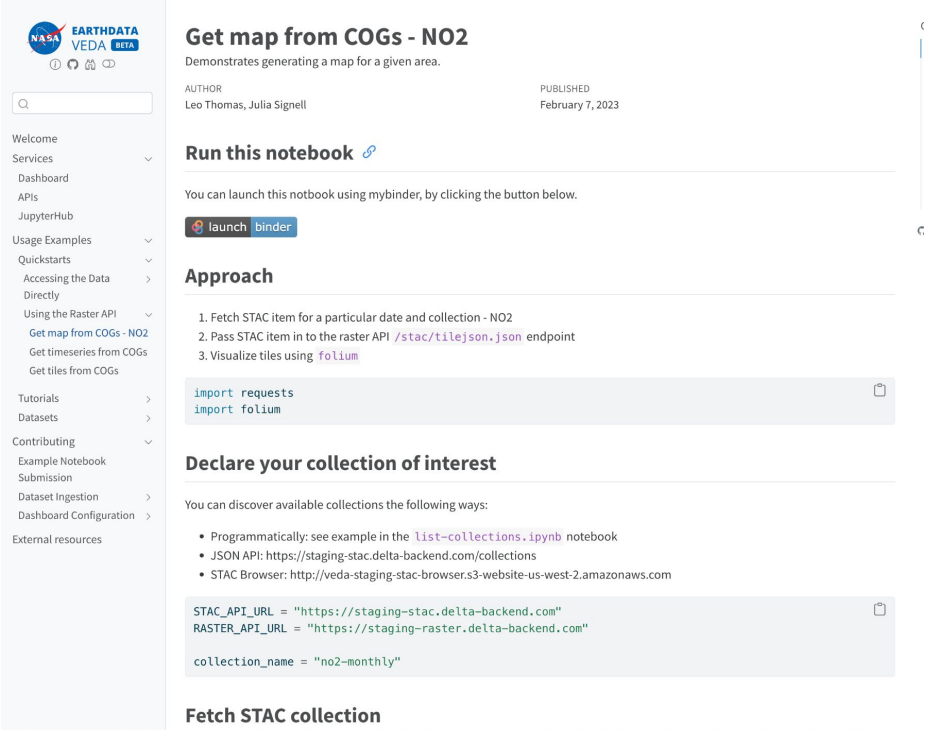
- Melanie Follette-Cook

Scientists can contribute documentation about their analysis to veda-docs

Tess McCabe submitted this mapping fires notebook: https://github.com/NASA-IMPACT/veda-docs/pull/79 which is previewed https://nasa-impact.github.io/veda-docs/pr-preview/pr-79/notebooks/tutorials/mapping_fires.html.

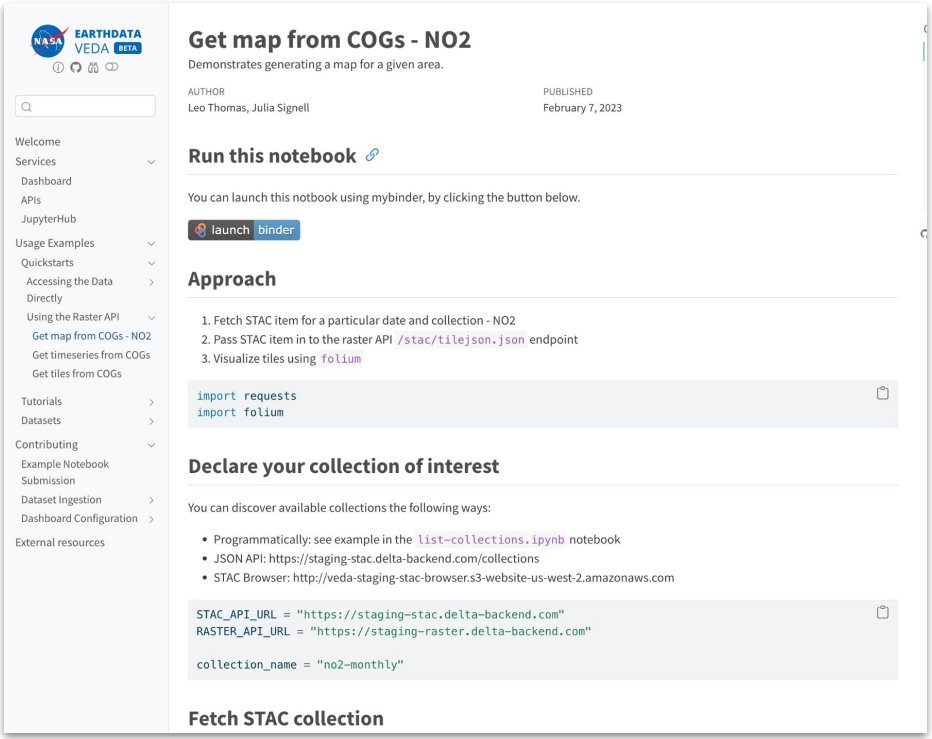
Try it out

- Go to https://nasa-impact.github.io/veda-docs and launch one of the notebooks with a mybinder link, such as https://nasa-impact.github.io/veda-docs/notebooks/quickstarts/no2-map-plot.html. Warning: mybinder.org is not reliable.
- **Solution?** Request access to the VEDA 2i2c jupyterhub. (https://nasa-impact.github.io/veda-docs/services/jupyterhub.html) Many tutorials have links to launch from there.



Try it out

SCAN ME



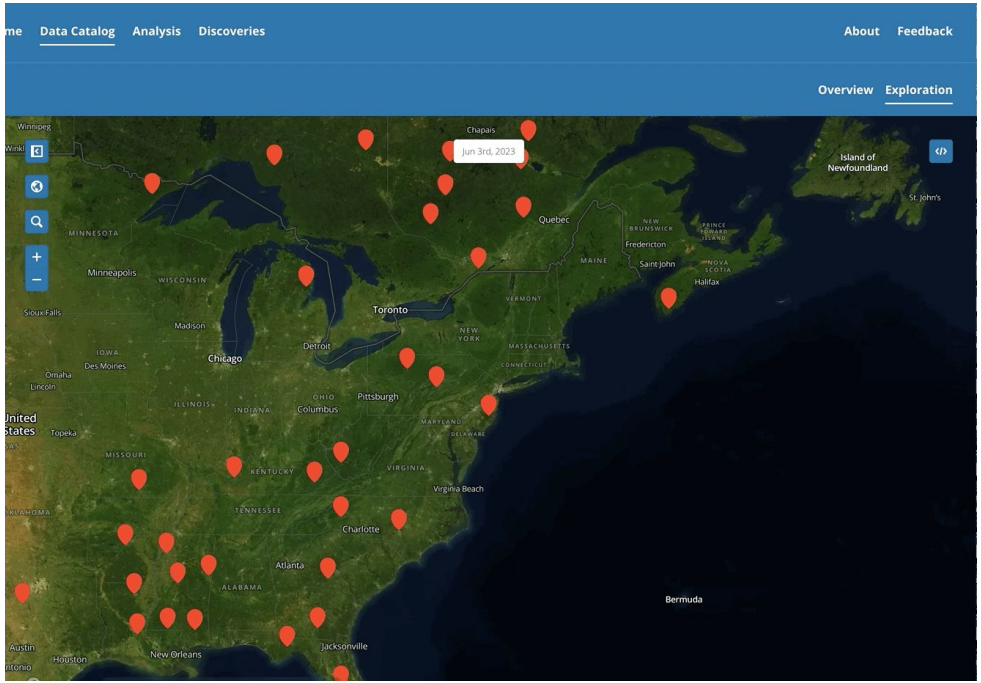
https://nasa-impact.github.io/veda-docs/

Scientists can contribute their datasets and discoveries to the dashboard.

Tess McCabe also submitted a PR to veda-config to add the fire perimeters dataset: https://github.com/NASA-IMPACT/veda-config/pull/260/files. Which can be seen in the staging site: https://visex.netlify.app/data-catalog/fire/

Try it out: Click “Explore Data”

SCAN ME



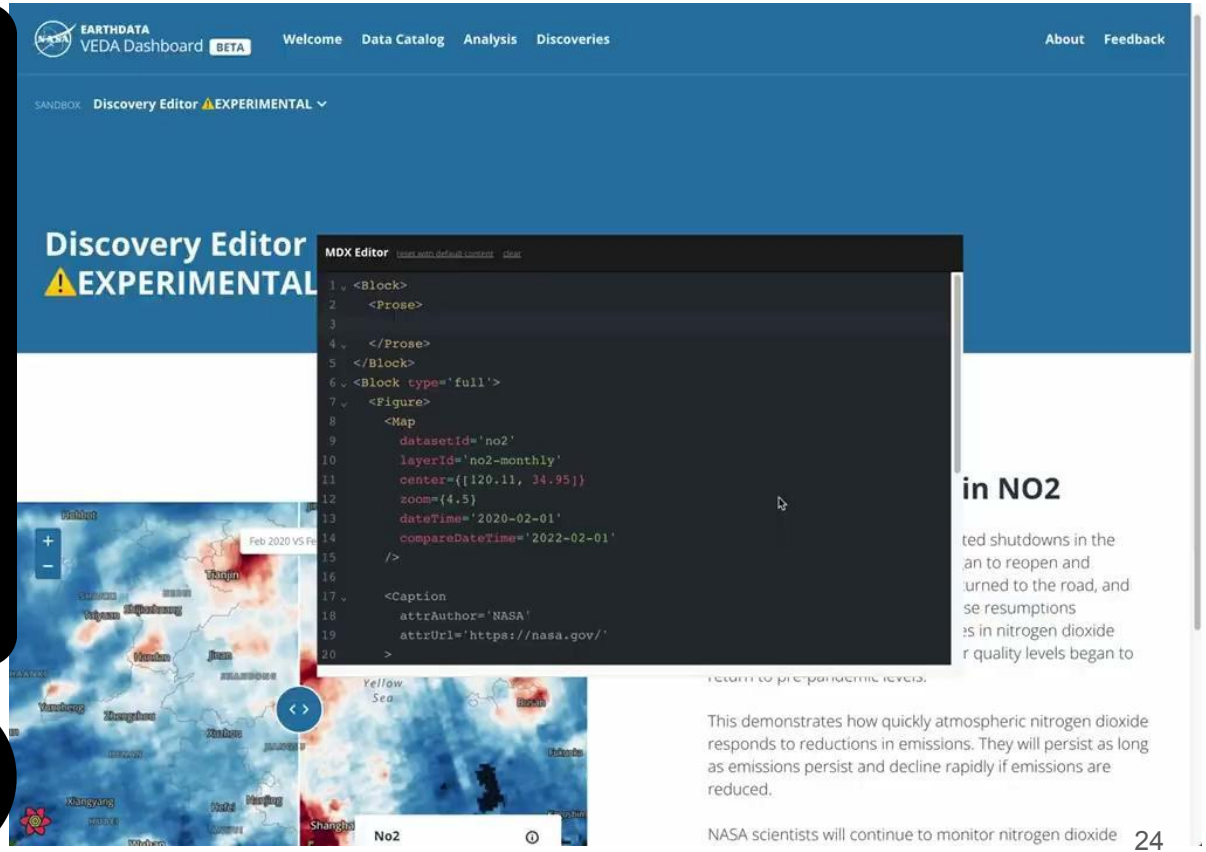
https://visex.netlify.app/data-catalog/fire/

However, MDX hard to edit for less technical audience not used to markup so we have an **experimental** live editor.

Try it out

https://tinyurl.com/veda-mdx-editor

SCAN ME

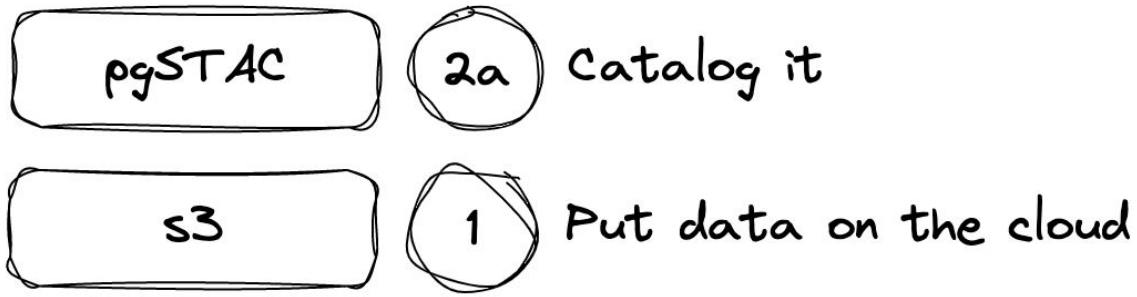


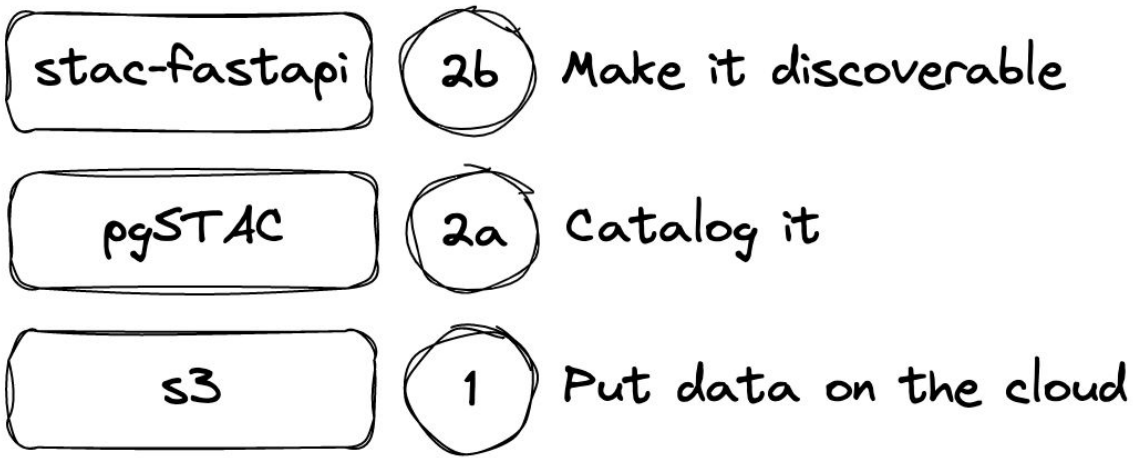
How does it all work?

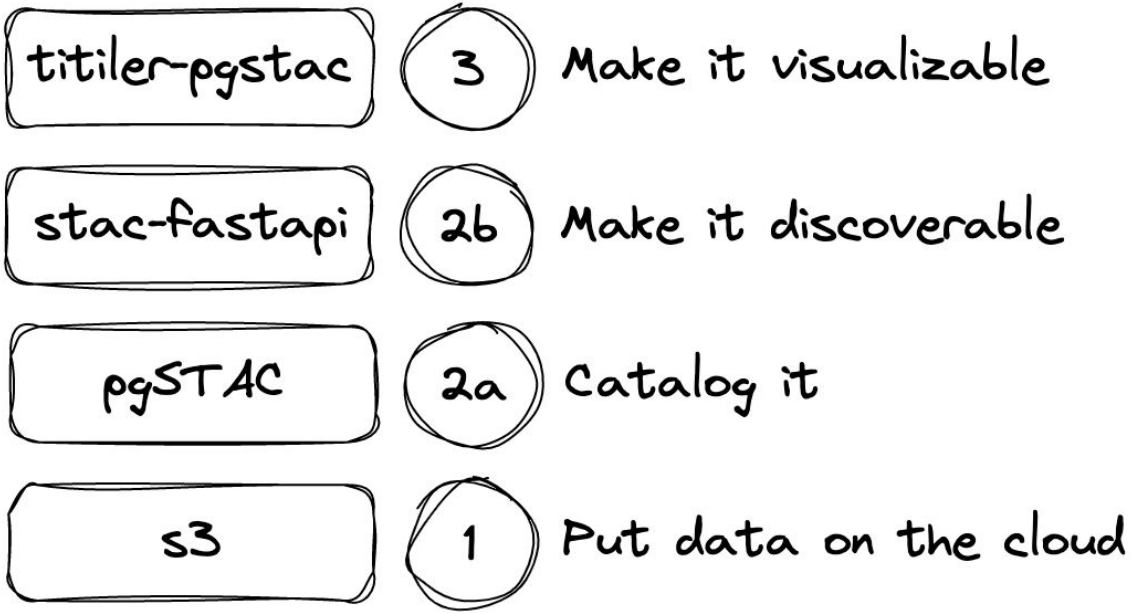
NASA IMPACT and Development Seed develop a modern STAC tool set. Storing STAC in postgres with pgSTAC means item discovery via the STAC API can be exactly transferable to rich visualization and data APIs: stats, OGC features and tiles for both raster and vector data.

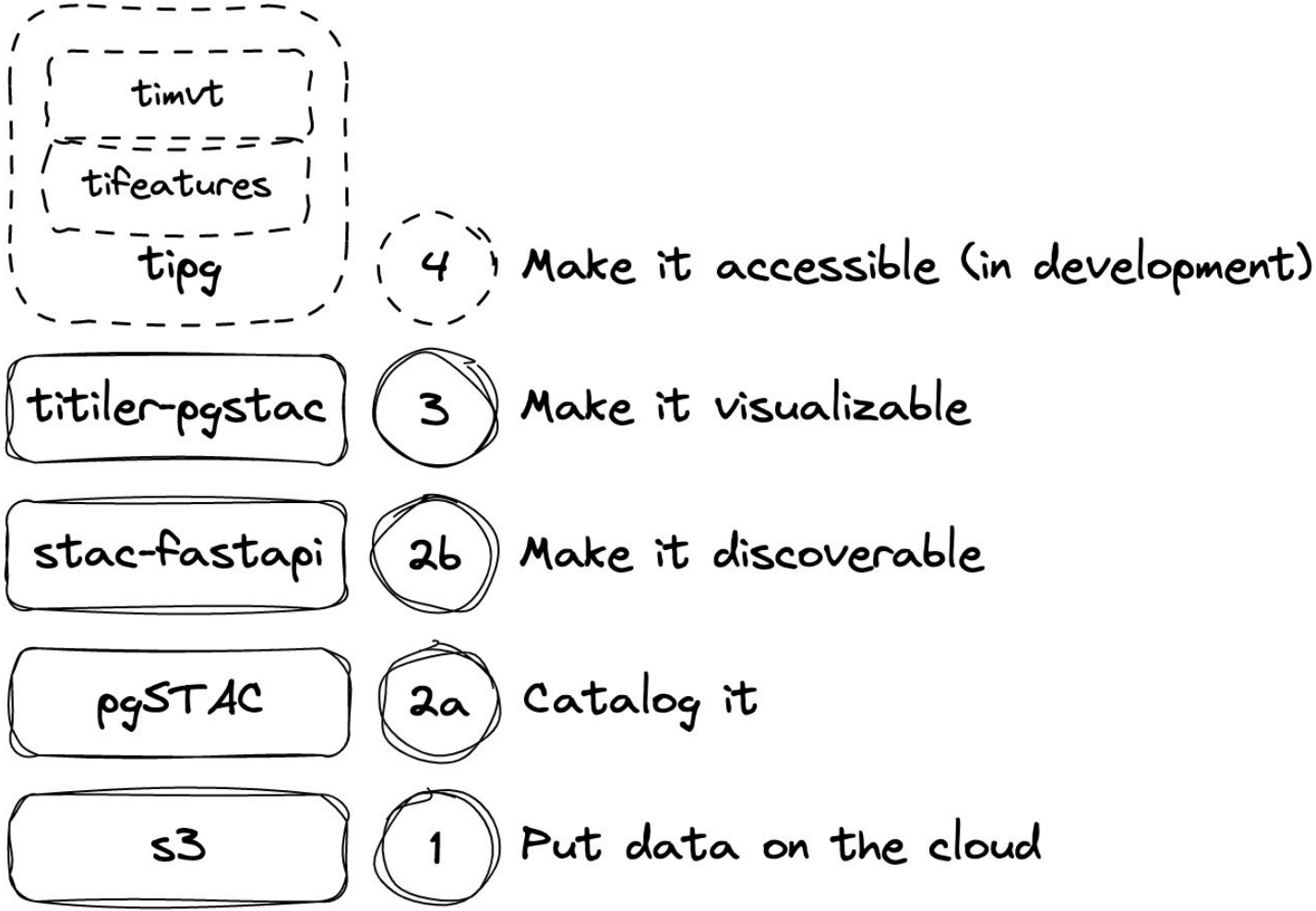
developmentSEED


s3
1
Put data on the cloud

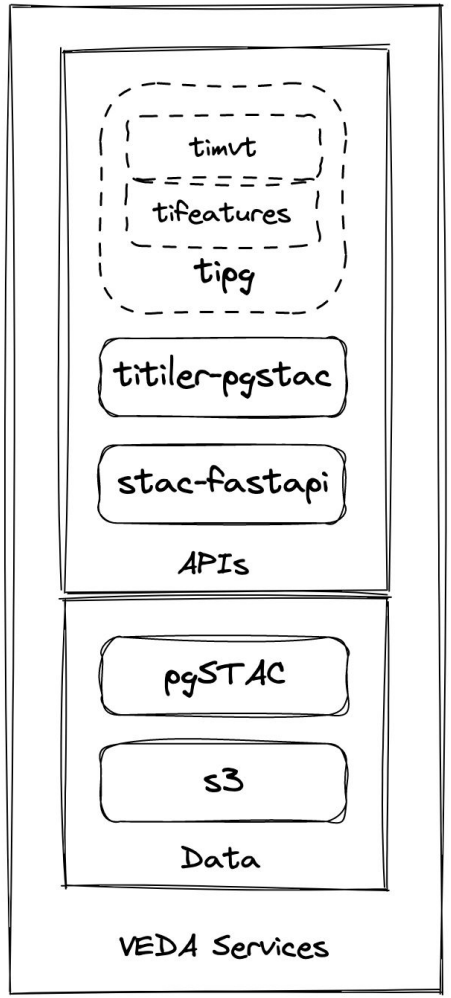

pgSTAC
2a
Catalog it
s3
1
Put data on the cloud

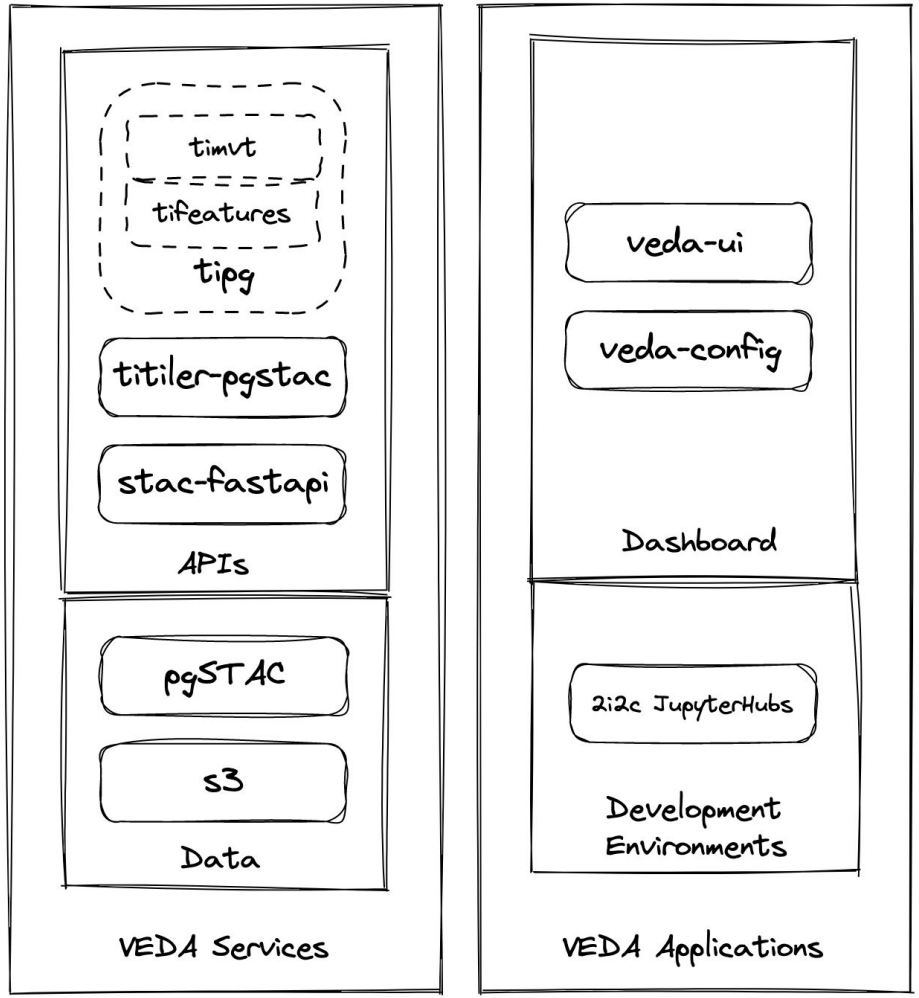

stac-fastapi
2b
Make it discoverable
pgSTAC
2a
Catalog it
s3
1
Put data on the cloud


titiler-pgstac
3
Make it visualizable
stac-fastapi
2b
Make it discoverable
pgSTAC
2a
Catalog it
s3
1
Put data on the cloud


timvt
tifeatures
tipg
4
Make it accessible (in development)
titiler-pgstac
3
Make it visualizable
stac-fastapi
2b
Make it discoverable
pgSTAC
2a
Catalog it
s3
1
Put data on the cloud

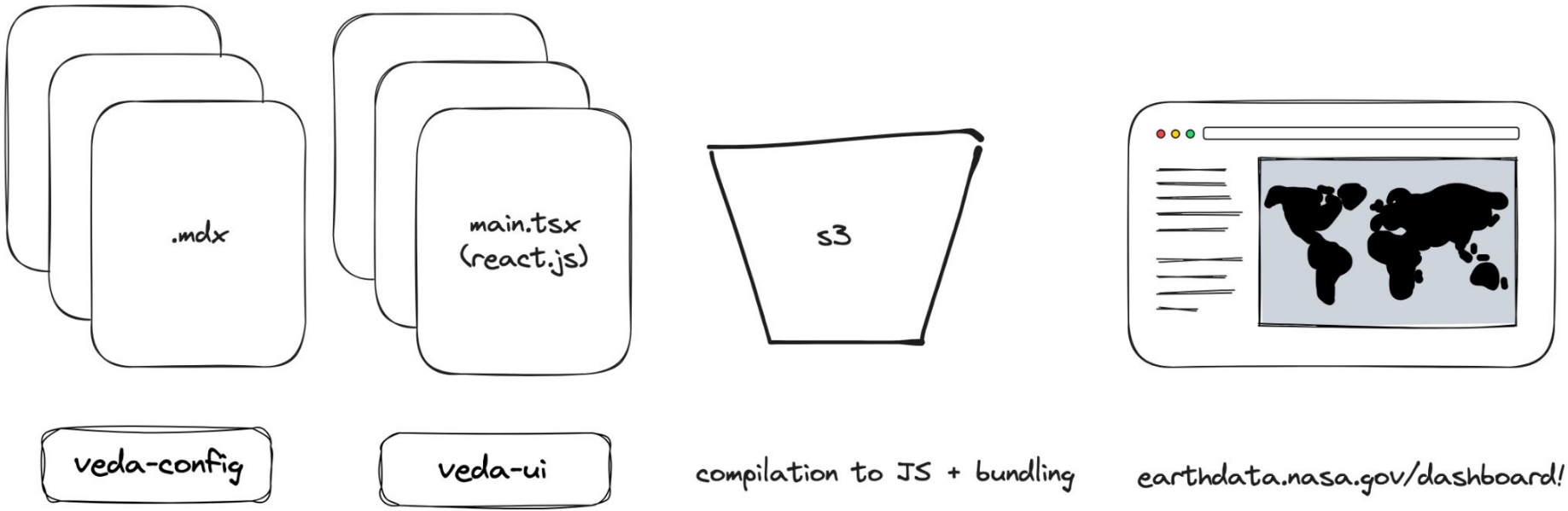
With these core data services, you can layer on applications.

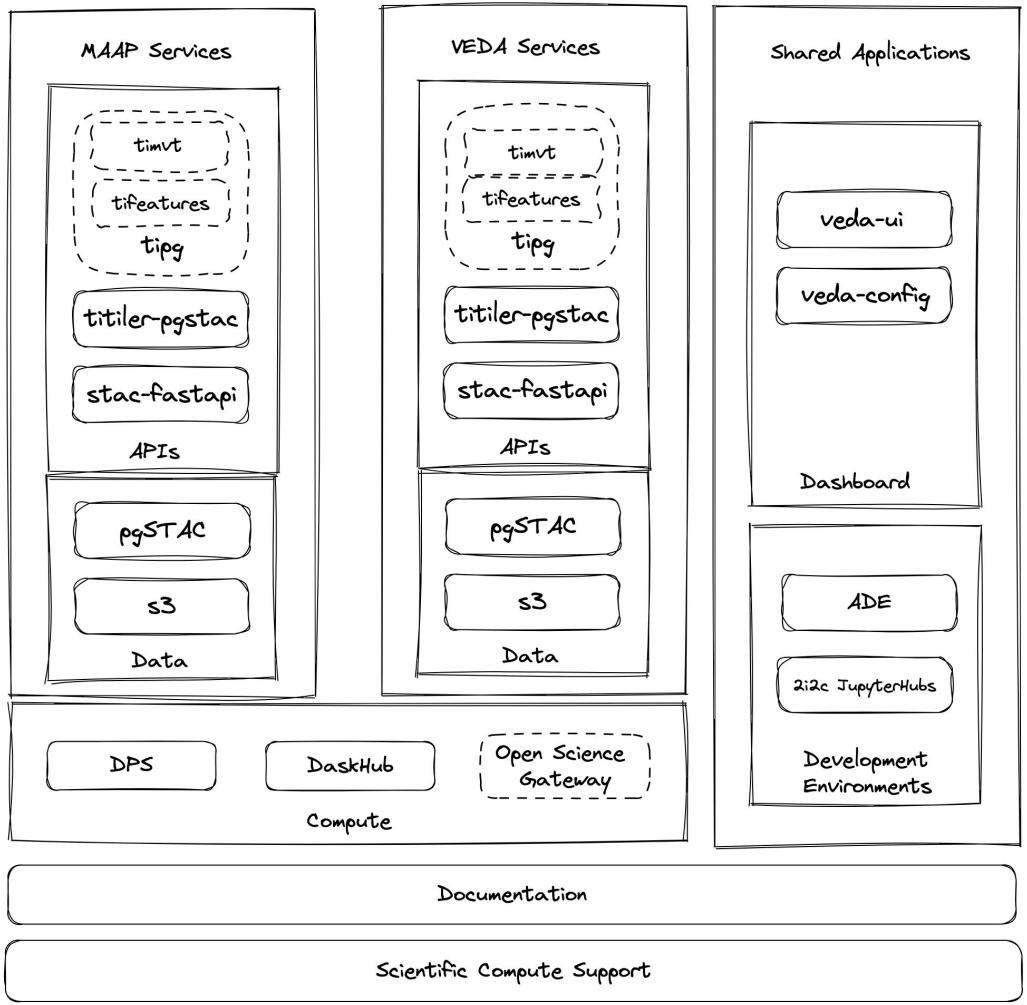

timvt
tifeatures
tipg
titiler-pgstac
stac-fastapi
APIs
pgSTAC
s3
Data
VEDA Services



veda-ui is configurable react.js components
veda-config stores the configuration: markdown files which point to STAC collections, items and queries.

How the dashboard works



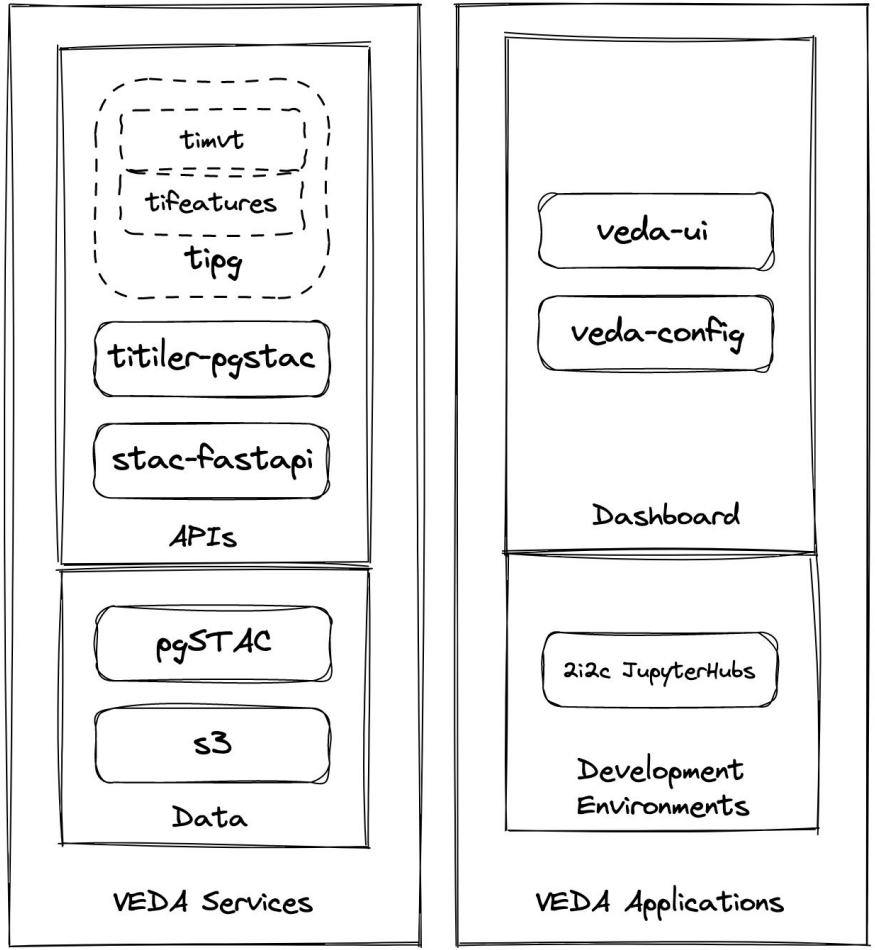


An Earth Observation and Exploitation Platform

data services + compute + documentation + support

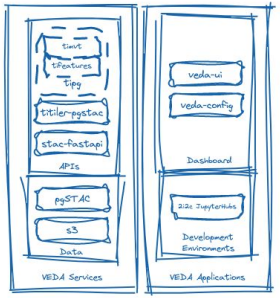
→ end user experiences

Reusable in part or in whole



VEDA Dashboard

earthdata.nasa.gov/dashboard

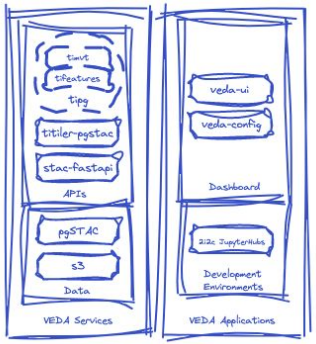


MAAP STAC Catalog

Greenhouse Gas (GHG) Center - July 2023

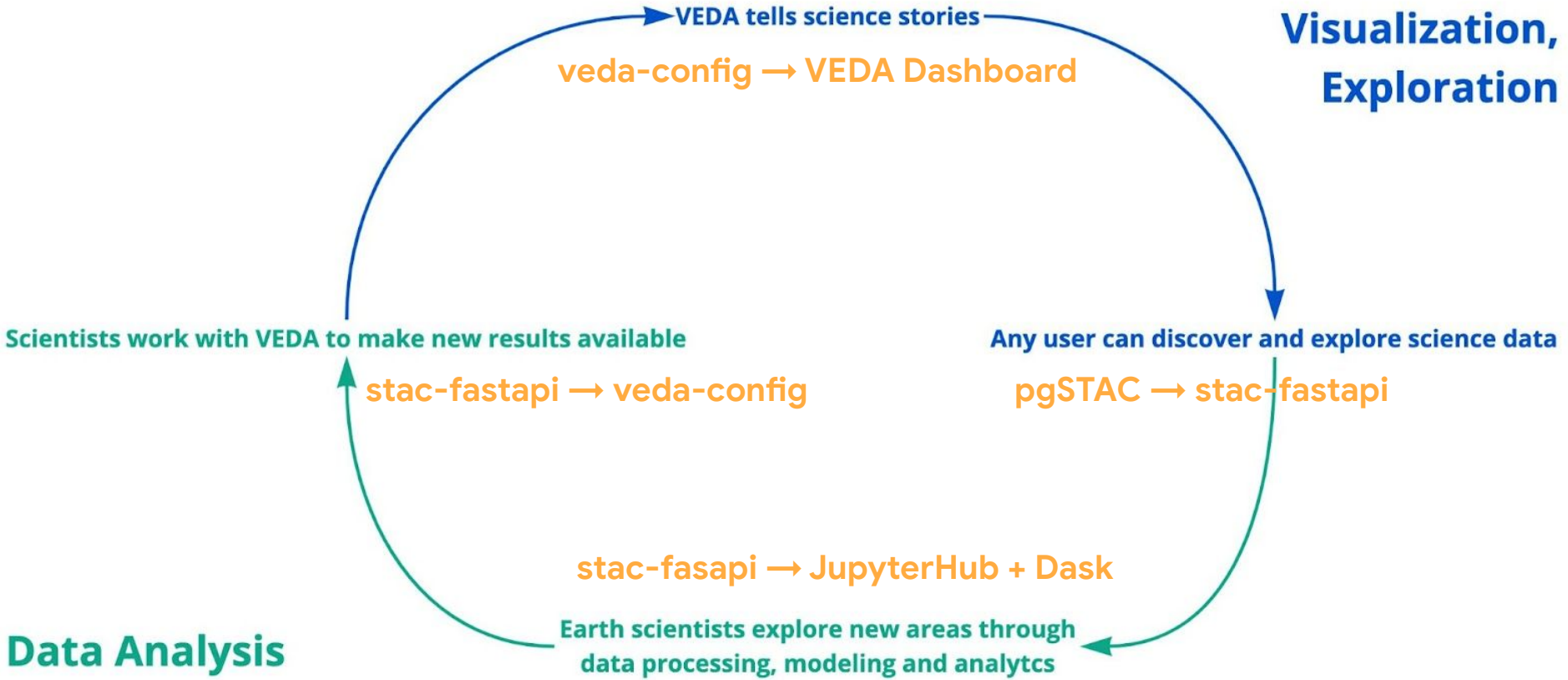
GIBS/Worldview Tiler

FIRMS titiler



Reusable components

Reusable Components



Reusable Components

- VEDA Data System
  - Database: GitHub - stac-utils/pgstac: Schema, functions and a python library for storing and accessing STAC collections and items in PostgreSQL
  - API: GitHub - stac-utils/stac-fastapi: STAC API implementation with FastAPI.
  - Image tiling: Github - stac-utils/titiler-pgstac
  - Features API: GitHub - developmentseed/tipg: Simple and Fast Geospatial OGC Features and Tiles API for PostGIS.
  - AWS Deployment: GitHub - developmentseed/cdk-pgstac
- VEDA Dashboard
  - Dataset and discovery configuration: GitHub - NASA-IMPACT/veda-config: Configuration template for the Dashboard Evolution Project
  - UI Code: GitHub - NASA-IMPACT/veda-ui: Frontend for the Dashboard Evolution project
- VEDA Data Analysis
  - JupyterHub: 2i2c Hub Deployment Guide

Get involved

- Request access to VEDA 2i2c JupyterHub
- Redeploy the VEDA stack of tools for application-specific science projects
- Help us
  - Become a part of the community
  - If you redeploy or need help redeploying, we want to help. This helps us scale the impact.

developmentSEED

2i2c
the international interactive
computing collaboration
2i2c.org

INDIANA UNIVERSITY

Element 84

Thank you!