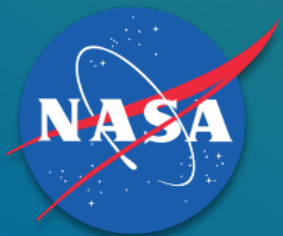


The Joint ESA-NASA Multi-Mission Algorithm and Analysis Platform (MAAP): Enabling Open Science for the Global Aboveground Terrestrial Carbon Dynamics Research Community

Amanda Whitehurst¹, **Kaylin Bugbee**², Laura Jewell³, Björn Fromnknecht⁴,
Clement Albinet⁴, Rahul Ramachandran⁵, Kevin Murphy⁶, Henri Laur⁴

*(1) ASTS/NASA (2) University of Alabama in Huntsville (3) Jet Propulsion Laboratory, California
Institute of Technology (4) European Space Agency (5) NASA MSFC (6) NASA Headquarters*



Importance of Open Science

- Open science allows collaboration that creates beneficial synergies
- Improves teamwork
- Encourages both higher quality work and community feedback on work
- Work has greater impact across the scientific community
- Open science creates new opportunities to leverage data, resources
- Encourages reuse of algorithms and computational workflows which lead to greater reproducibility



Challenges to Open Science

Data friction

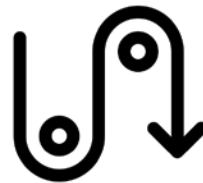
- “costs in time, energy, and attention required simply to collect, check, store, move, receive, and access data”
- Every time data is moved or transformed there is a risk of data loss/corruption
- Questions of data trustworthiness, data quality



Created by Template
from Noun Project

Computational friction

- “Every calculation requires time, energy, and human attention.”
- Physical and cost limits of data processing
- Human work of programming, debugging, etc...
- Effort involved in getting others to accept results



Created by Jems Mayor
from Noun Project

Science Limitations

- Lack of availability of data
- Need for transparency of methods and algorithms
- Need for data at correct temporal and spatial resolution



Created by Flatart
from Noun Project

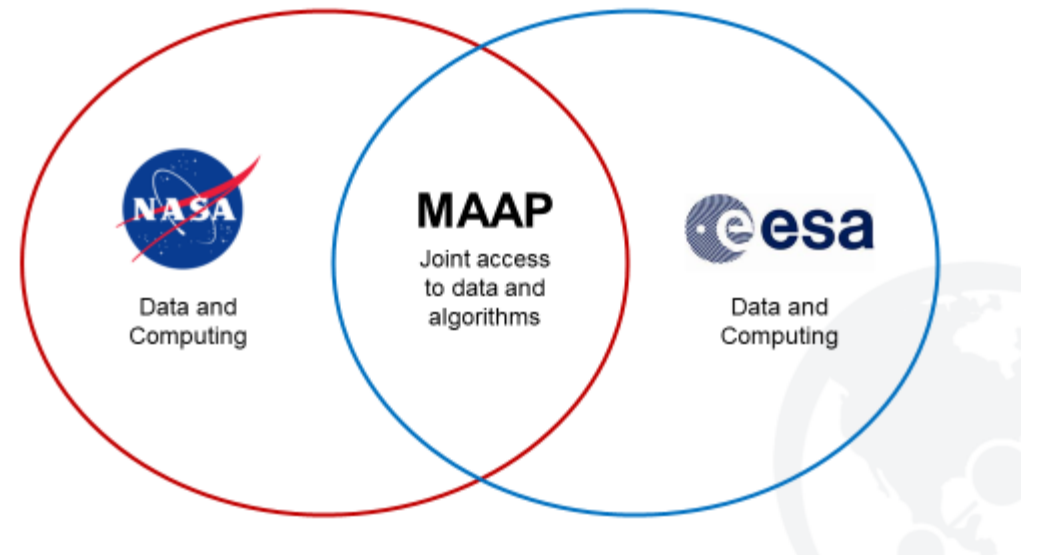
What is MAAP?

- The MAAP is a virtual environment dedicated to the unique needs of sharing and processing data from relevant field, airborne and satellite measurements related to ESA and NASA missions
 - Jointly managed by ESA and NASA and accessible to designated ESA and NASA scientists.
 - Initially populated with pre-launch and complimentary data from other projects.
- Science focus is to improve the understanding of global terrestrial carbon dynamics & to support algorithm development
- Addresses a need expressed by the science community to more easily share and process data collected by NASA and ESA activities



How is the MAAP Enabling Open Science?

- The MAAP's long term vision:
 - Clearly connect data, algorithms, software and results to support the global aboveground terrestrial carbon dynamics research community
 - Encourage community collaboration by
 - Providing collaborative work environments
 - Making it easy to share data, algorithms and software to collaborators and the MAAP
 - Reducing data and computational friction
 - Support ESA and NASA's commitments to open data, open software

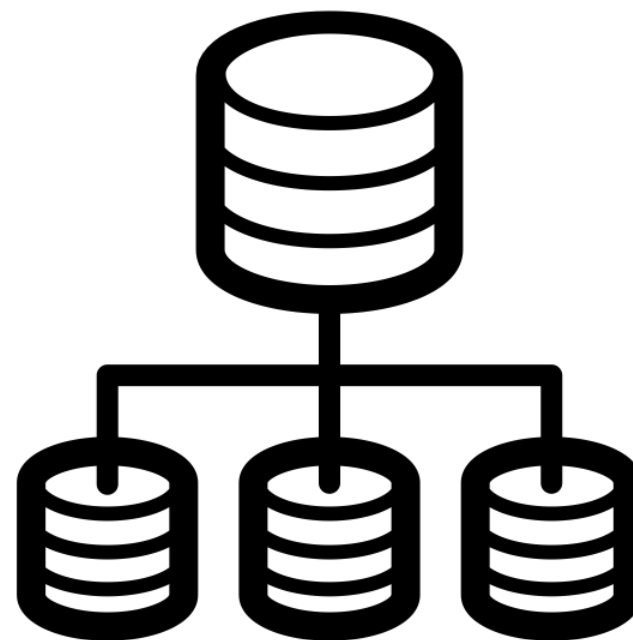




Data

Easing Data Friction: Data

- Facilitating discovery of biomass relevant data via a centralized location
 - Allows data from various organizations to be quickly discovered
 - Data includes primary mission data and supporting ancillary data
 - Highly curated data holdings encouraging data reuse
- ESA and NASA are contributing metadata to a single repository
 - Common Metadata Repository (CMR)
 - Meta(data) may be discovered via an API or the Earthdata Search client
 - Additional metadata information provided to support biomass search needs

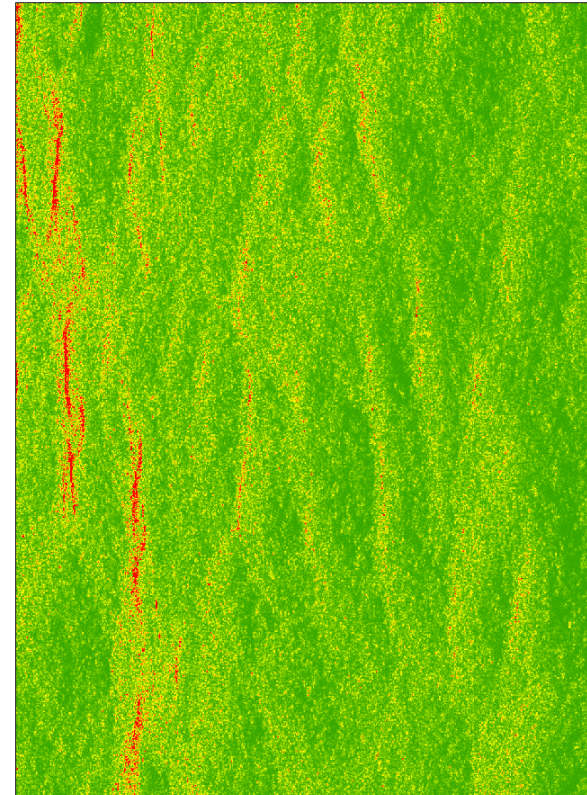


Created by H Alberto Gongora
from Noun Project



Easing Data Friction: Data Format

- Identified data formats relevant to biomass user community
- ESA and NASA are providing primary data in agreed upon data formats to lower data use barriers
 - GeoTIFF
 - HDF5
- Exploring analysis ready data options to make data use even easier
 - Exploring efficient formats for data with high volume and high spatial/temporal resolutions
 - Exploit cloud infrastructure



*Image: AfriSAR ONERA SLC VV
GeoTIFF*

Easing Data Friction: Open Data Access Policy

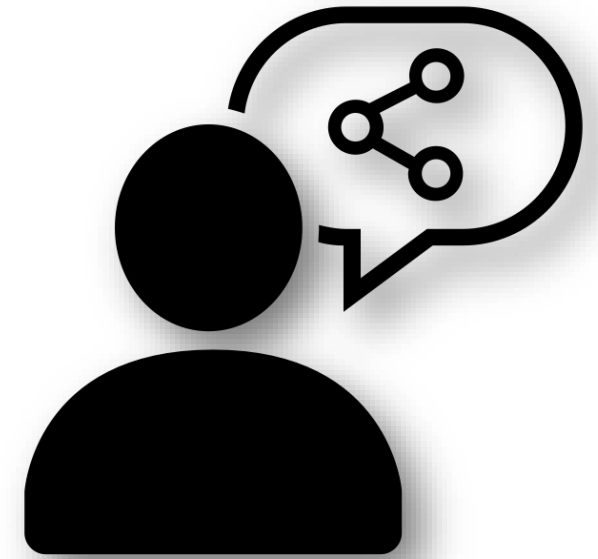
- MAAP encourages free flowing data by promoting full and open sharing of all data
 - Standard MAAP data products will abide by NASA and ESA's open data policies
 - User uploaded and generated data will also be required to adhere to the MAAP open data policy
- Data is exchanged between NASA and ESA without restriction
- Data access for users is likewise non-discriminatory
 - All users treated equally



Created by Gregor Cresnar
from Noun Project

Easing Data Friction: User Data

- Enabling quick and easy data sharing with MAAP users
 - Users can share data with select collaborators
 - Can share data more broadly to MAAP CMR so users can discover it
 - Supports MAAP open data policy
- To make data sharing easier, MAAP will leverage creative ways to capture metadata info
 - Capturing information from the data itself
 - Streamlining metadata needs to lower burden on user



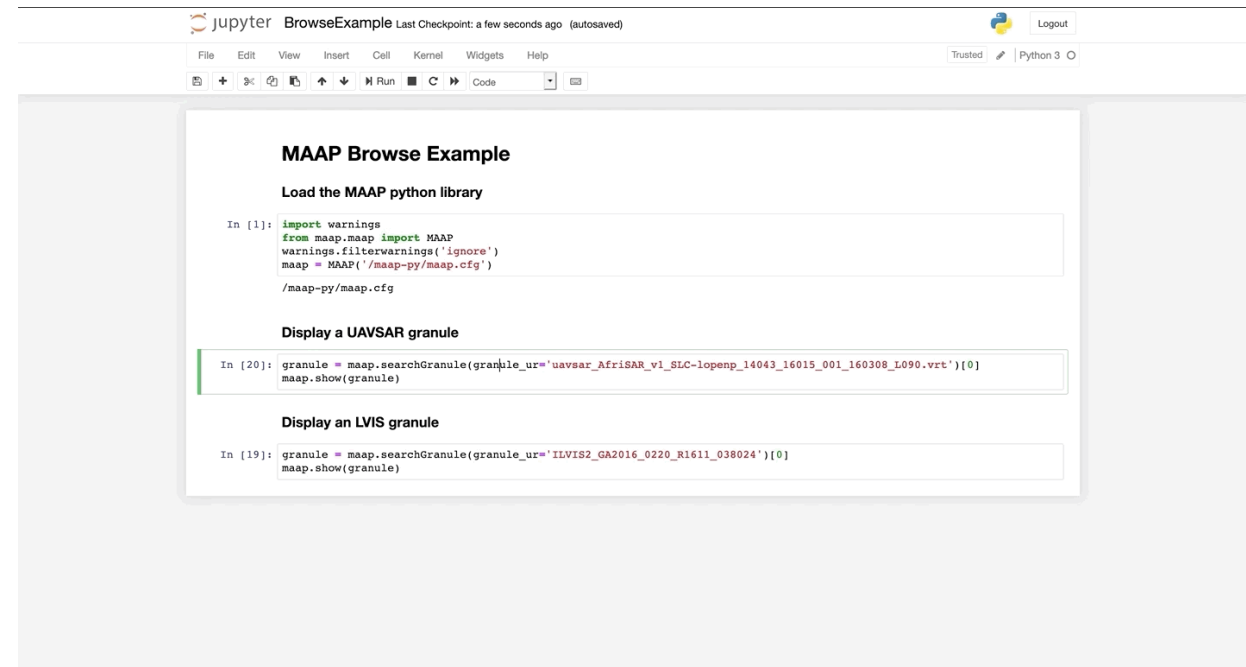
Created by Thomas' designs
from Noun Project



Computation & Open Source Policies

Easing Computational Friction: Open Source Policy

- User developed implementations of algorithms will be subject to MAAP's open source policy
 - Permissive licenses to support use, modification and redistribution will be strongly encouraged
- Software developed for the MAAP itself is open source
 - Both NASA and ESA have completed open source process
 - Source code will be available on GitHub once pilot MAAP is complete



```
jupyter BrowseExample Last Checkpoint: a few seconds ago (autosaved)
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3
+ ↺ ↻ ↵ ↶ ↷ ↸ ↹ ↺ ↻ ↵ ↶ ↷ ↸ ↹ Code

MAAP Browse Example
Load the MAAP python library
In [1]: import warnings
from maap.maap import MAAP
warnings.filterwarnings('ignore')
maap = MAAP('/maap-py/maap.cfg')

/maap-py/maap.cfg

Display a UAVSAR granule
In [20]: granule = maap.searchGranule(granule_ur='uavsar_AfrisAR_v1_SLC-lopemp_14043_16015_001_160308_L090.vrt')[0]
maap.show(granule)

Display an LVIS granule
In [19]: granule = maap.searchGranule(granule_ur='ILVIS2_GA2016_0220_R1611_038024')[0]
maap.show(granule)
```



Enabling Science

Science

- MAAP will make more in situ data available needed for reference, calibration, validation
 - GEDI Cal/Val Database
- MAAP will act as a centralized location of several biomass-focused spaceborne observations
 - Encourage the development of fusion biomass estimation products
 - Promote collaboration across organizations
 - Increased geographic coverage, detail from combined data



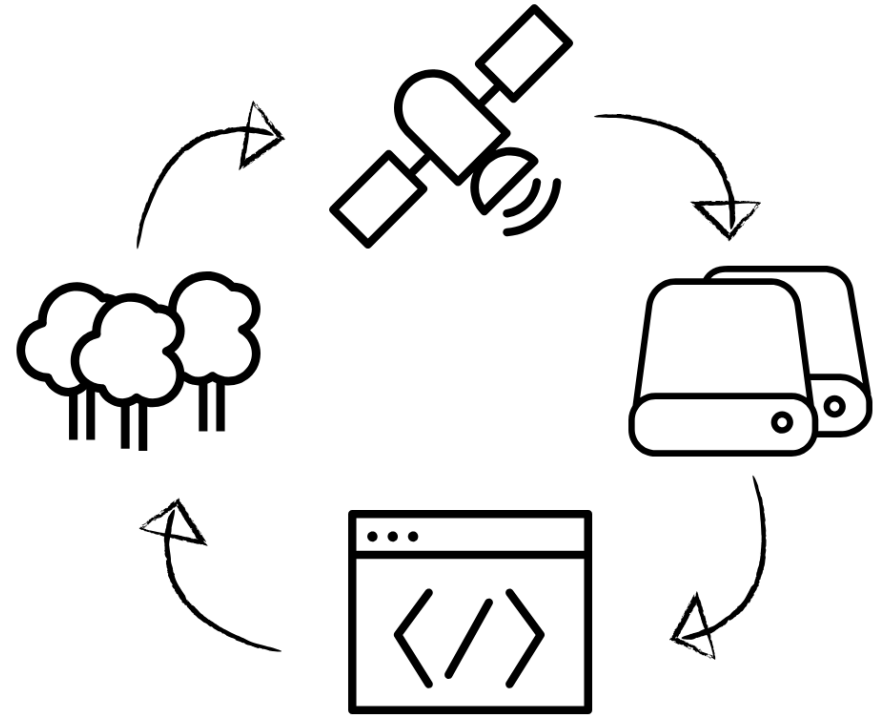
<https://www.jpl.nasa.gov/images/earth/africa/20160225/earth20160225d.jpg>



<https://www.jpl.nasa.gov/images/earth/africa/20160225/earth20160225c.jpg>

Conclusions

- The MAAP will make connections between data, algorithms, software and results
- Should make it easier to reproduce results and build from existing work
- Encourage collaboration between scientists and data scientists
- Help users match the tool to the question
- Bring together data from various spaceborne missions from various organizations to support development of biomass maps



Questions?

Contact me at:
Kaylin.m.Bugbee@nasa.gov