ORNL/TM-2024/3449

MAPSTER: Automated Geospatial Data Sharing – version 1.4.0

Matthew Larson
Justin Cazares
Tyler Waddle
Aaron Myers
Alec Hamaker

July 2024

OAK RIDGE National Laboratory

ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

DOCUMENT AVAILABILITY

Online Access: US Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via https://www.osti.gov.

The public may also search the National Technical Information Service's National Technical Reports Library (NTRL) for reports not available in digital format.

DOE and DOE contractors should contact DOE's Office of Scientific and Technical Information (OSTI) for reports not currently available in digital format:

US Department of Energy
Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831-0062
Telephone: (865) 576-8401
Fax: (865) 576-5728
Email: reports@osti.gov
Website: www.osti.gov

ORNL/TM-2024/3449

Geospatial Science and Human Security Division

MAPSTER: AUTOMATED GEOSPATIAL DATA SHARING – VERSION 1.4.0

Matthew Larson
Justin Cazares
Tyler Waddle
Aaron Myers
Alec Hamaker

July 2024

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

LIST OF FIGURES

SUMMARY

The US Department of Energy's (DOE) Oak Ridge National Laboratory (ORNL) developed MAPSTER which is a geospatial data management tool that aggregates, organizes, and shares data from dispersed sources such as unmanned aerial systems (UAS). Built specifically for use in environments where communications may be limited, MAPSTER utilizes two key technologies to effectively manage data in the field and enable easy data sharing with authorized partners: Observer and Checkpoint. Observer is a lightweight software package on an edge device, such as a laptop, that automatically detects newly processed UAS data and sends to a central server called Checkpoint. Checkpoint is a centralized server at ORNL that receives and manages data from all Observer instances. Even in a very low bandwidth environment, Observer can still send information about the UAS data product almost instantly as it generates its own metadata package on the size of KB (kilobytes). MAPSTER is not only for UAS data but for any geospatial data collected at the austere edge and dispersed sources.

1. BACKGROUND

The most common geospatial intelligence (GEOINT) data products are from land surveys, satellite, and airborne imagery. However, with the proliferation of low-cost UAS, GEOINT can be captured by the warfighter at the edge more frequently and with data products that are centimeter resolution. As the technical capabilities of capturing GEOINT have increased, the National Geospatial Intelligence Agency (NGA) has expanded the mix of services and support they provide. In 2018 the NGA Warfighter Support Office created a UAS GEOINT kit called the Aerial Reconnaissance Tactical Edge Mapping and Imagery System (ARTEMIS). Additionally in 2022, NGA in partnership with ORNL created the Intelligent Reconnaissance Imagery System (IRIS) which added vertical takeoff capabilities, obstacle avoidance, and rapid 2D orthomosaic generation in less than 10 minutes. These UAS kits enable warfighters to gather rapid GEOINT at the edge in a disconnected environment. As the functional manager for GEOINT within the intelligence community, NGA requires users of UAS kits to send back collected data.

Traditional GEOINT products such as satellite imagery are collected and processed by NGA and then pushed out to the warfighter. It is the reverse case for UAS collection systems as the data is collected at the edge and then pushed back to NGA. Pushing UAS data back to NGA from field can be difficult as there can be limited communications as well as warfighters forgetting to send the data back after the mission. At the edge, it is difficult to enforce data upload requirements and data quality.

The old workflow from data collection, processing, and sending back the data to NGA is on the magnitude of hours on a best-case scenario. The majority of the UAS data collected at the edge is lost due to the manual upload process. After conducting a UAS survey flight, the user must process the data into a 2D orthomosaic or 3D model. Depending on the amount of data collected, processing can take hours to approximately a day on a laptop. The newest IRIS kit reduces the processing time significantly as most datasets can be processed into a 2D orthomosaic in under 10 minutes. However, the user still must manually upload the data to an unclassified data sharing site such as Protected Internet Exchange (PiX), then NGA manually downloads the data back at the enterprise. Quite often NGA personnel must reach out to the warfighter or end user to upload their data to PiX which is time consuming. Additionally, there is no efficient way to directly send data back to the NGA enterprise, store, and view all deployed UAS collected data. Automating the sending and sharing of small UAS data being collected is critical for not only reducing the overall workflow for the warfighter but also NGA receiving the most up to date GEOINT from the edge and reducing the amount of data being lost. MAPSTER allows for an automated process of managing and sharing data from the tactical user. Compared to the manual upload method,

MAPSTER fully automates the process, and the user only has to focus on the mission and processing the data. MAPSTER additionally gives a place for the user to store and manage the information they are collecting and keep a historical record of at the edge GEOINT data.

Along with the need of getting collected at the edge data back from the tactical environment, the same issue persists in emergency response. After a natural disaster such as a tornado, earthquake, hurricane, or flood, up-to-date information is needed quickly to make actionable response plans. Many disaster response agencies deploy UAS to map the extent of the damage. However, with all the UAS deployed in a disaster situation, it is difficult to share and manage data across agencies. MAPSTER allows for that immediate automatic sharing of that data not only to the agencies on the ground but additionally back to headquarters and the decision makers. It also helps reduce data duplication efforts between agencies and fosters and open sharing data environment. Specifically, DOE's Cybersecurity, Energy Security, and Emergency Response Office (CESER) and Emergency Support Functions (ESF) are rapidly mobilized when a major disaster or emergency causes damage to energy systems. These ESF units use UAS to map the energy infrastructure at a high resolution. MAPSTER allows for proliferation of that data through DOE to help make quick decisions to get energy infrastructure back up running again. Whether there is a low or high bandwidth internet connection after a natural disaster. MAPSTER aids in immediately sharing and managing at the edge geospatial data.

2. MAPSTER OVERVIEW

MAPSTER is a geospatial data management tool that aggregates, organizes, and shares data from dispersed sources such as UAS. Built specifically for use in environments where communications may be limited, MAPSTER utilizes two key technologies to effectively manage data in the field and enable easy data sharing with authorized partners: Observer & Checkpoint (Figure 1).

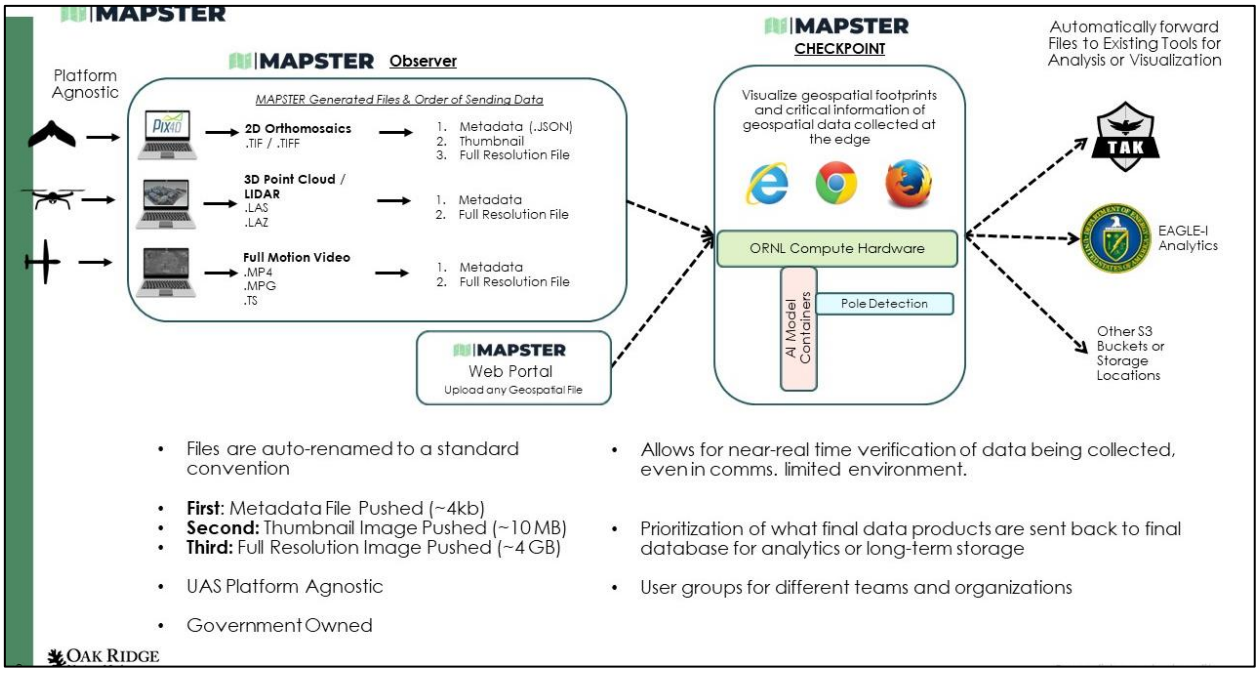


Figure 1.MAPSTER workflow and features of version 1.4.0.

2.1 OBSERVER

MAPSTER Observer is a light-weight software package that can be installed on field deployed computers which automatically manages data and creates metadata packages designed to organize large datasets, making them easy to retrieve and share. These small metadata files are sent first to the Checkpoint server before the actual data product as not all deployed UAS teams will have high bandwidth internet connections suitable for large data transfers. This structured sending approach ensures the agency collecting the data knows what data products are being collected as soon as possible.

The Observer runs in the background and therefore needs no interaction by the user. Currently, Observer continuously looks for new 2D (.tif/.tiff), 3D (.las/.laz), and video (.mp4/.mpg/.ts) files within the MAPSTER Files folder or a folder/s the user selects. The 2D and 3D data files types are the most common output type from Pix4D or other UAS image processing software. MAPSTER is UAS platform and software agnostic so it can easily be integrated into any workflow. Observer can be modified to look for any geospatial file type as MAPSTER evolves.

When Observer starts for the first time it creates a local database with a list of all the supported file types in the directory. When a new file is detected, it adds it to the database to eliminate duplicating data product uploads. For the new file, Observer creates a JSON (JavaScript Object Notation) file which includes critical metadata information about the data product such as date, time, sUAS platform, camera, and location. This metadata JSON file is very small in size and is only a few kilobytes (kb). The purpose of generating these metadata JSON files is to always send metadata to the checkpoint first as not all end-users will have high-bandwidth internet at every location. Sending these kb-sized JSON files, even over satellite communication systems is on the order of seconds.

After the metadata file is created, a thumbnail .PNG file is created as well from the .tiff file. This thumbnail is a lower resolution version of the data product and is considered a “quick look” image. The thumbnail is on the order of ~10 MB and is sent to checkpoint after the metadata file. Even in a low-bandwidth environment, this file is still sent quickly. The order that data is sent to Checkpoint is as follows: first the metadata JSON file is sent, second is the quick look thumbnail, and third is the full resolution .tiff file. This allows for near real time verification of data products being collected in the field. For 3D and video data files, no thumbnail is generated.

Before any of the files are sent to Checkpoint, the JSON, thumbnail, and full resolution data products are automatically renamed to a standard naming convention. Automatically naming the files allows for standardization across all data products and provides easy data management. Currently in MAPSTER 1.4.0, the data product naming convention is as follows:

YYYY-MM-DD_TIME_ProcessingSoftware_CoordinateSystem_UniqueIDNumber

Example: *2024-03-05T163330_PIX4DMATIC1.54.3_WGS84UTMzone18S_eb743.tiff*

There is an Observer user interface (UI) (Figure 2) which allows for troubleshooting, monitoring data upload, and additional advanced features. Observer UI does not need to be opened by the user in an operational scenario. One of the advanced features in the Observer UI is a low-bandwidth button. If a user is in a low-bandwidth environment (i.e. satellite communication), and has many data products to upload, instead of trying to upload many GB sized .tiff files, the user can toggle on the low-bandwidth button which only sends the metadata JSON files to checkpoint and pauses the upload of the full resolution files. This allows the user not to stress the already low-bandwidth connection especially if it is needed for other

tasks. The button has a timeout function so when the user is back in a high-bandwidth environment, the full resolution files will automatically begin uploading to Checkpoint.

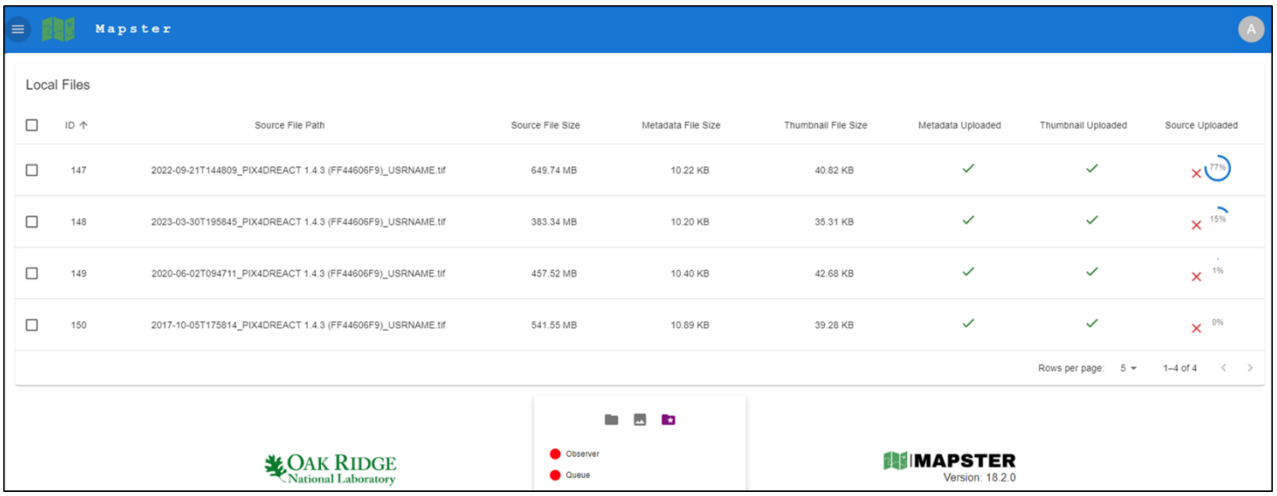


Figure 2. MAPSTER Observer.

2.2 CHECKPOINT

The MAPSTER Checkpoint is a central server located at ORNL where all at the edge instances of Observer software will send their data. Within Checkpoint, users can view, manage, and download data products. To access Checkpoint, users can go to https://mapster.ornl.gov on any browser to create an account and log into the system. Currently, MAPSTER accounts are only available to U.S. Government agencies.

2.2.1 Dashboard

After a user logs into MAPSTER Checkpoint, they arrive at the main dashboard (Figure 4). The dashboard allows the user to search for data products within a specific date range or area as well as download the latest version of Observer. To search for data products in an area, the user draws a bounding box on the map and only data products in that area are displayed. The resulting data products from the search methods are listed in the results section of the dashboard. Data products are shown on the map as polygons. Each polygon represents a single data product and the geospatial extent that it covers. More information about the data can be viewed by either clicking on a polygon or by clicking the “View Data” button next to the data product (Figure 3).

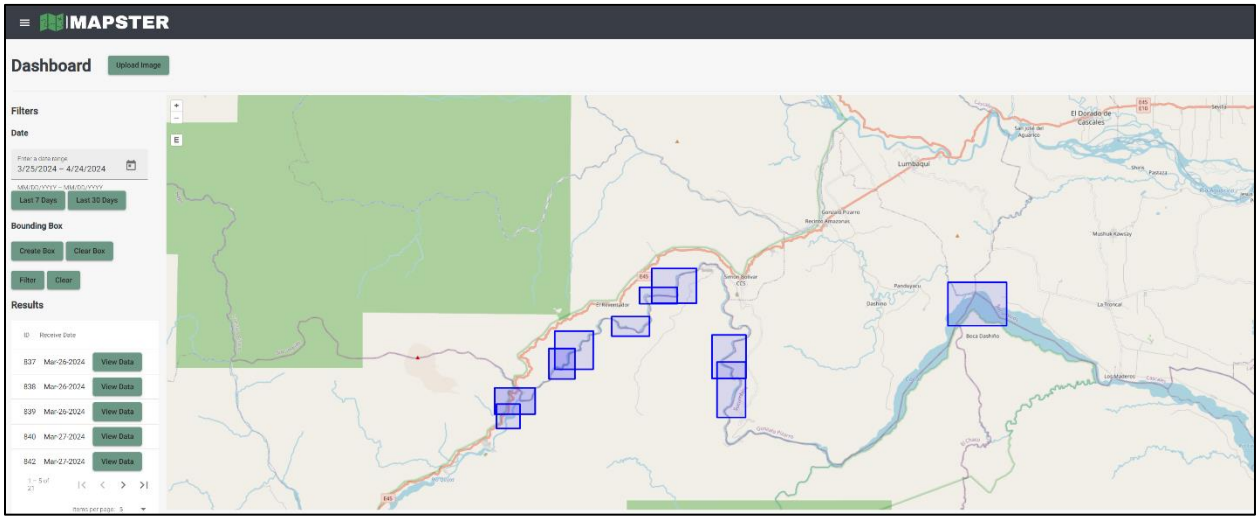


Figure 3. MAPSTER Checkpoint dashboard showing available data products.

2.2.2 Data Product Baseball Card

When viewing more information about a data product, each has its own "baseball card" (Figure 4). The baseball card lists all the critical information about the data product allowing the user to download it to their computer. The metadata JSON file is what fills out the baseball card, and the image displayed is the thumbnail. If a baseball card does not display a thumbnail, it means that the thumbnail or full resolution .tiff file has not been uploaded yet. In the baseball card, the following information is listed about the data product: file name, date of collection, country location, state and county (if in U.S.), center coordinate of image, processing software used, and user ID. The actual username is hidden from the baseball card to protect the identity of the user. Each user has an ID number that can be looked up by an administrator.

2.2.3 Web Upload

If a user does not have Observer installed on their laptop, they can still upload data products via the "Upload Image" button on the dashboard. A user can upload a supported file type directly to Checkpoint. Once the file is uploaded, it is auto-named to the standard convention and associated metadata file is generated.

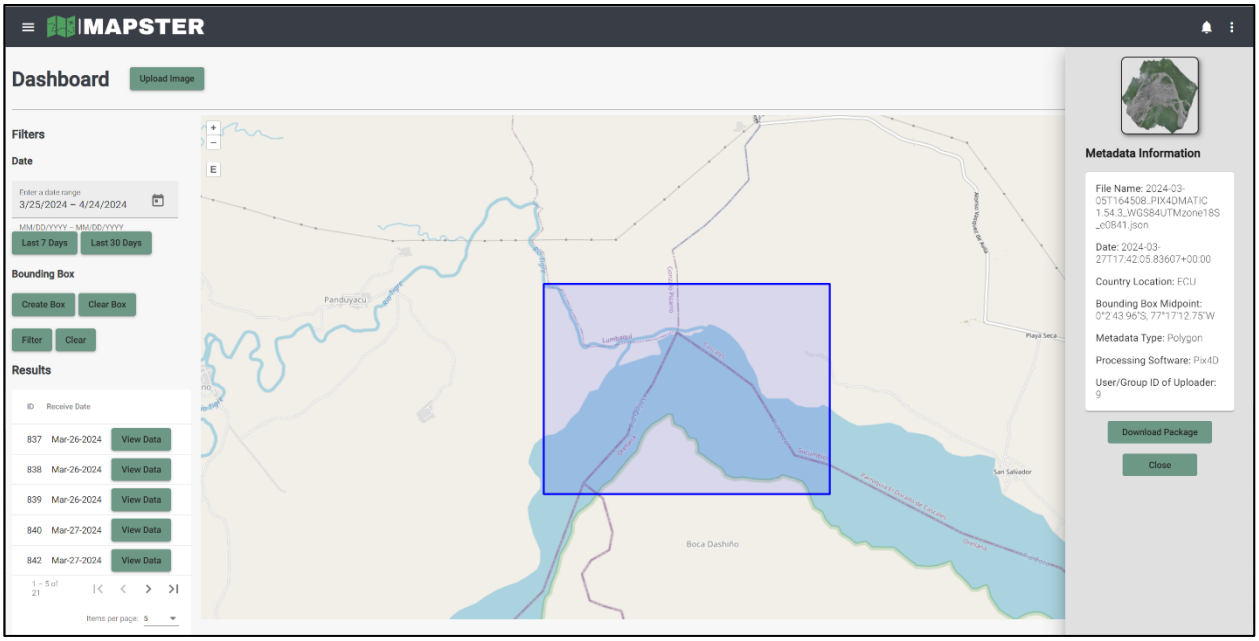


Figure 4. Data product on MAPSTER and its associated baseball card with critical information.

2.2.4 Account Management and User Groups

Checkpoint allows for users to manage their account and reset their password as needed. A new feature that was added to Checkpoint was user groups. The default group that data is uploaded to is the Public Group in where all MAPSTER users can view and download the data. If a specific user group is created, only users in that group can upload, view, and download that data. If desired, data products in a user group can be migrated over to the Public Group at any time. As many different agencies will be using MAPSTER, there are instances where some data products may only want to be shared within a group and not with other agencies

2.2.5 AI/ML Containers

Checkpoint also has a section for running AI/ML (Artificial Intelligence / Machine Learning) models on uploaded data. In support of DOE ESF-12 mission of monitoring the energy infrastructure, a damaged utility pole detection AI/ML model was developed. For this particular model, the user uploads the original UAS data, and the model will detect all utility poles in the image and if they are damaged or undamaged. The model will additionally output the coordinate of each utility pole and its confidence percentage on if it is damaged (Figure 5).

While the Pole Detection Model is the only AI/ML model currently on Checkpoint, MAPSTER was designed to easily add additional models that can be selected by the user. In future versions of MAPSTER, additional AI/ML models will be added to the library for the user to select.

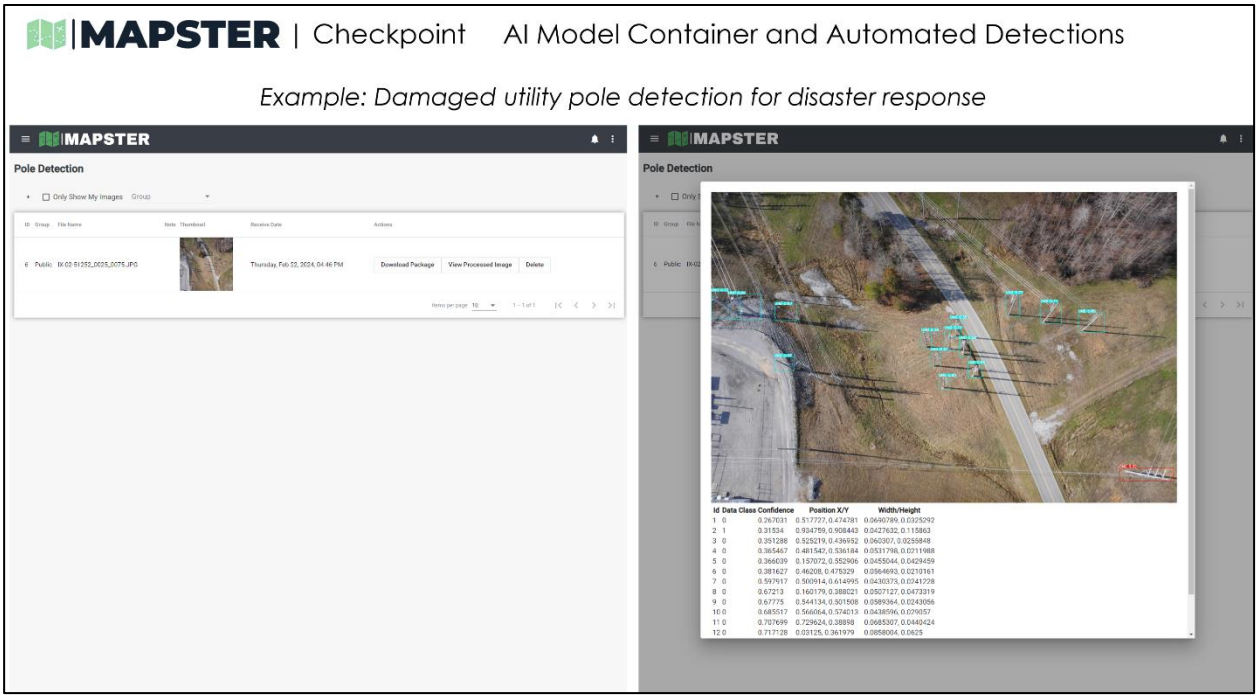


Figure 5. Damaged utility pole detection AI/ML model on Checkpoint.

3. USE CASES AND TESTING

3.1 AMORY, MISSISSIPPI EMERGENCY RESPONSE

In March 2023, a tornado destroyed most of the city of Amory, Mississippi. ORNL responded shortly after the disaster with two UAS kits to map the extent of the damage and status of the energy infrastructure. Utilizing MAPSTER, minutes after the data was collected and processed into 2D orthomosaics, data was able to be viewed back at ORNL and at DOE HQ in Washington, DC. This was a good operational test, to easily share the data in a disaster response situation. Without MAPSTER, it would have taken days, to share the data with the appropriate decision makers.

3.2 AMAZON RAINFOREST, ECUADOR

ORNL and NGA's QX office were asked to map over 100 km of the Rio Coca, in the Amazon Rainforest, Ecuador due to extensive erosion damaging the energy and transportation infrastructure. UAS were used to map the region as satellite imagery is non-existent due to the constant cloud cover of the rainforest. 2D data products were generated and MAPSTER was used to share that data with scientists back in the U.S. Due to the austere environment of the Amazon Rainforest, the only communication system was a Starlink panel. Within minutes of processing the data, MAPSTER was able to share the metadata and thumbnail of the collected imagery with scientists back in the U.S. After approximately 20 min, the full resolution imagery was also shared. This scenario proved the functionality of MAPSTER in an austere environment and sending data through a low-bandwidth connection.

4. ADDITIONAL FEATURES

Additional features that MAPSTER can include are the following:

- Sending any geospatial file (e.g. AI/ML detections, phone images, etc)
- Automatic image processing on Checkpoint – Users can send raw .Jpg UAS files to Checkpoint where it will automatically be processed into 2D/3D data products.
- Federated storage – allows for users to plug in their own storage hardware and not have all data hosted at ORNL. This also includes a MAPSTER monitoring client which tracks what files are available on the user's storage and can pull that file to the main Checkpoint if a user wants to download it.

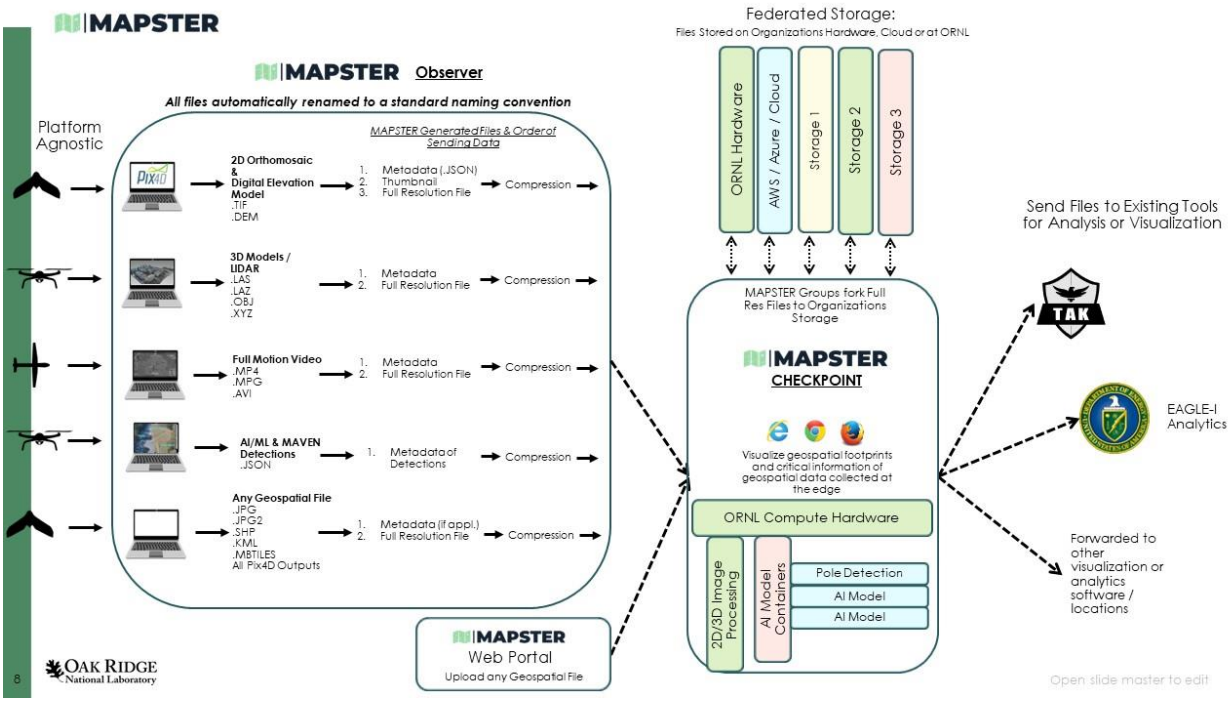


Figure 6. Additional MAPSTER features.

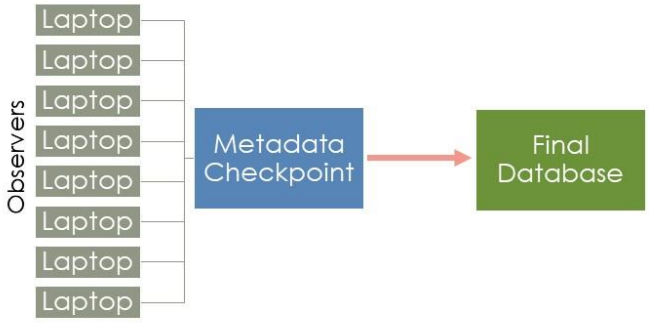
5. MAPSTER SUMMARY

MAPSTER is an automated geospatial sharing software designed to share geospatial data collected from dispersed sources such as UAS back to a central location. MAPSTER consists of two components: the MAPSTER Observer and Checkpoint. The Observer is a light-weight software package that is installed on an edge device (e.g. laptop) while the Checkpoint is a central server where all the data is sent to. The Observer continuously looks for processed 2D orthomosaics, 3D point clouds, and video data, and automatically pushes both metadata, thumbnail, and full resolution file to Checkpoint. A user can login to Checkpoint on any web browser and view, manage, and download full resolution files from dispersed, at the edge sources. Additionally, Checkpoint also supports direct upload of these data types if they do not have the software on their laptop. MAPSTER allows for ensuring quality and standardization of geospatial data collected from dispersed sources such as UAS and sending that data to a central location where it can easily be shared and managed.

6. APPENDIX

Summary

- MAPSTER is a lightweight software package that automatically searches deployed laptops for geospatial data and compiles metadata (GPS coordinates, file size, etc) at central checkpoint.
- The *Checkpoint* allows for an organization to see what geospatial data is being collected in the field by multiple assets. It can prioritize what geospatial data is sent back to the final database.
- Mapster supports the following file types
  - 2D data
    - .TIF / .TIFF
  - 3D data
    - .LAS / .LAZ
  - Video
    - .MP4 / .MPG / .TS

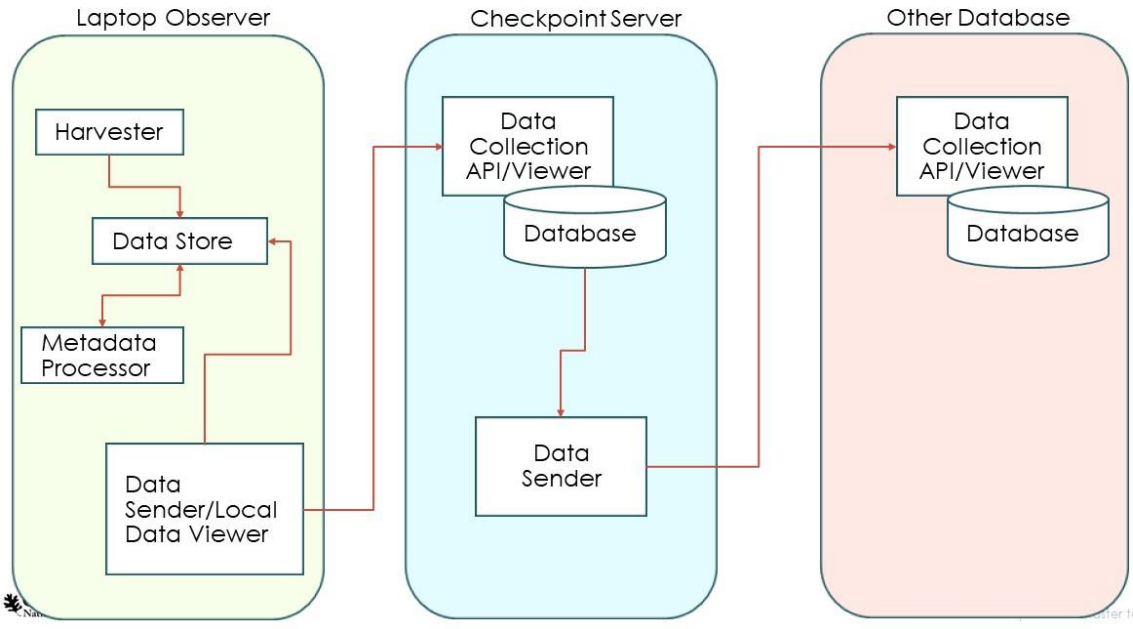


OAK RIDGE National Laboratory

2

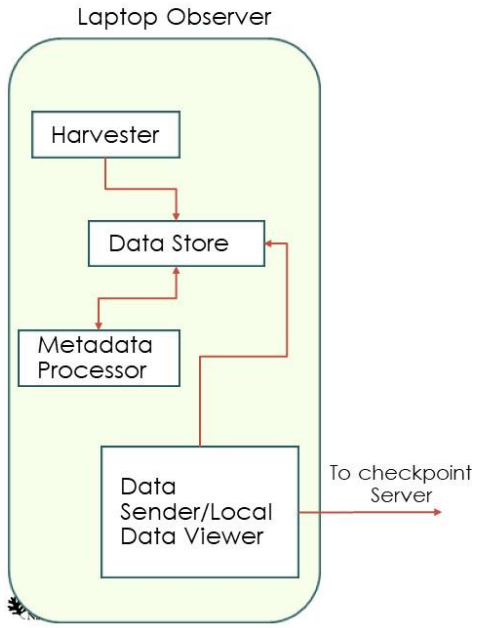
Open slide master to edit

Detailed MAPSTER Data Flow Chart



3

Laptop

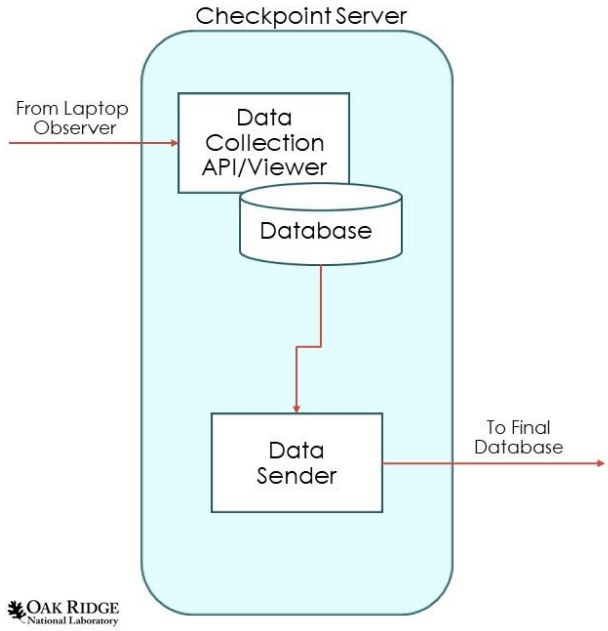


1. Harvester searches laptop for any Geospatial data files (e.g., .TIF / .TIFF / .LAS / .LAZ / .MP4 / .MPG / .TS)
2. *Observer* creates a data store with all the found supported files on the user defined directories.
3. *Observer* creates a metadata .JSON file for each found file.
4. Metadata, thumbnail, and original file are automatically renamed to a standard convention.
5. Metadata is sent to a central *Checkpoint* server first
   1. In the case of 2D data, a thumbnail is generated and sent second
6. Original file is sent to checkpoint last
7. *Observer* can also wait for a data request from the Checkpoint

4

Open slide master to edit

Checkpoint Server



1. *Checkpoint* Server receives Geospatial metadata information and original files from many laptops running the observer
2. The *Checkpoint* builds a list of all Geospatial Data collected from the laptops and what is available for download
3. *Checkpoint* can request and prioritize Geospatial downloads
4. Since the *Checkpoint* is receiving metadata from many laptops, it can show on a map where the Geospatial Data is being collected and its geospatial footprint.

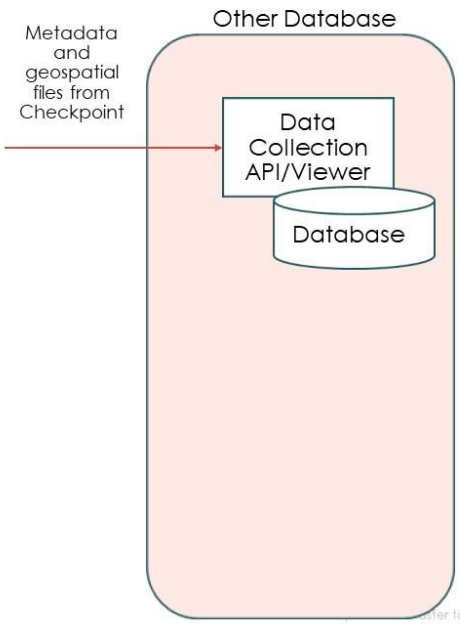
5

OAK RIDGE National Laboratory

Open slide master to edit

Other Database

1. Other database is optional
2. Data products can be forwarded to another location
3. This database can be set by the organization
   Examples:
   1. AWS S3 Bucket
   2. PixToday
   3. Azure Cloud



OAK RIDGE National Laboratory

6

Open slide master to edit

OBSERVER

Supported File Detected

File recorded in data store with unique UUID

Generation of supporting files (UUID is assigned to each one to keep them as a set)
- Metadata (.JSON)
- Thumbnail (.PNG)

Auto-naming of file to standard naming convention
- Metadata (.JSON)
- Thumbnail (.PNG)
- Original File

Files sent to Checkpoint in order of lowest file size to highest.
1. Metadata (.JSON)
2. Thumbnail (.PNG)
3. Original File

CHECKPOINT

Checkpoint receives the files in the following order
1. Metadata (.JSON)
2. Thumbnail (.PNG)
3. Original File

Checkpoint links all the files together based on their UUID as a "Data Set"

Metadata file is used to outline area on map of the data products spatial coverage. Metadata also used to fill out information about the data product on "baseball Card"

Thumbnail is displayed on baseball card

The entire dataset can be downloaded from the baseball card as a package (Metadata, Thumbnail, Original File)

OAK RIDGE National Laboratory

7

Open slide master to edit