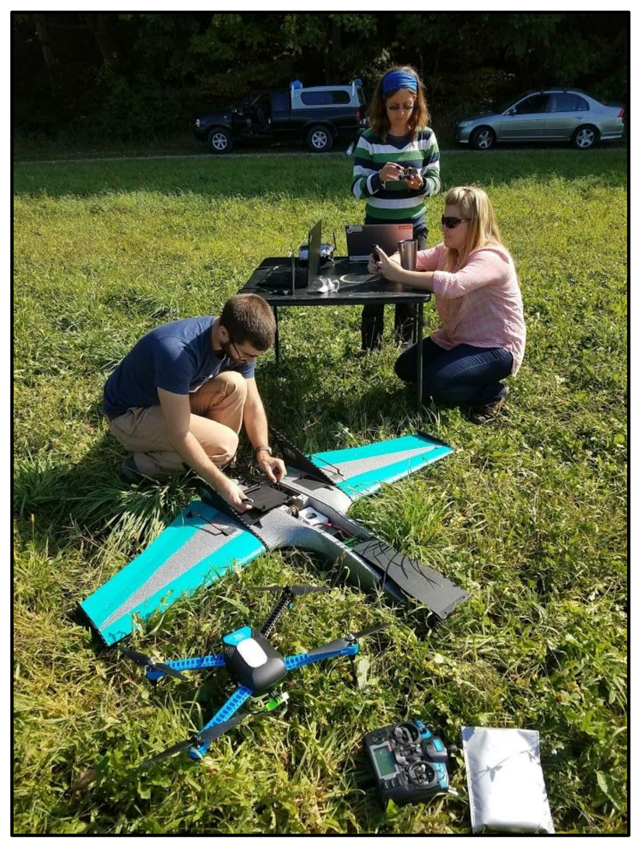
Minimal Information Framework for Scientific Data Collection from Remotely Piloted Aircraft Systems (RPAS)

Authors: Jane Wyngaard (jwyngaar@nd.edu), Lindsay Barbieri (lkbar@uvm.edu), Charles Vardeman II (cvardema@nd.edu), Bryan Leahy (bleahy@umich.edu), Sarah Swanz (spswanz@umich.edu), Andrea Thomer (athomer@umich.edu)



Abstract. The image and sensor data collected by Remotely Piloted Aircraft Systems (also known as “drones”, or small Unmanned Aircraft Systems) are rapidly changing the way researchers in many domains collect data. However, in order for these data to be fully utilised they must be made FAIR [Findable, Accessible, Interoperable and Reusable]. An initial step towards achieving FAIRness is to both augment them with machine-readable, semantically-rich metadata, and to annotate them in ways that make their provenance (the record of the processes that created the data) explicit. In RPAS-based research this is particularly challenging given the many agents (e.g. people, RPAS, sensors, controllers, computers, software systems), and complex processes (e.g. pre- and post- correction processing, data integration) with often inexplicit relationships. We are working to draft a **minimum information framework for data collected by RPAS,** as well as **a drone data ontology.** The MIF outlines a core set of parameters all RPAS data should be accompanied by. The drone data ontology draws on existing ontologies and semantic web efforts.

Approach: collecting and analyzing scientific RPAS workflows.

Through workshops and semi-structured interviews, we are developing use cases of RPAS data collection. A general workflow, with details from each of three cases follows.

I] Pre flight

1. Science Question / Campaign Planning | **2.** Selection of Platform & Sensors | **3.** Sensor Integration on Platform | **4.** Pre-Flight Check / Sensor Calibration

1. Define project hypothesis and research questions
   a. **Agriculture emissions case:** Does x farming method reduce greenhouse gas emissions?
   b. **Crop cover biomass case:** Does Biomass correlate with nutrient run-off?
   c. **Pika habitat case:** How large is the population/extent of x flora/fauna?
2. Select platform and flight sensors
   a. Multirotor, CO2 and CH4 sensors
   b. Fixed wing, RGB camera, RTK GPS
   c. Fixed wing, Hyperspectral camera
3. Perform integration accounting for how aircraft will impact sensor sampling such as the positioning of collection ports, air and heat flow, angle of sensor relative to ground, and potential EMF interference. Final solution may require post-processing data correction.
4. Preflight checks: Aircraft, GPS, sensor, and environmental safety checks
   a. Ground samples: wind, gas flux rates, soil moisture, soil temperature
   b. Camera calibration images, ground control points, water samples
   c. Camera calibration images, ground control points.

II] Flight

5. Mission Planning / In field | **6.** Flight & Data Collection | **7.** Download / Stream data

5. Mission planning:
   a. Create grid flight pattern over field of concern to follow the terrain and set the sampling rate (time or GPS triggered)
   b. Create grid flight pattern over field of concern accounting for ground control points and desired image overlap for given altitude.
   c. Similar to (b). May adjust to manual control if monitoring moving targets.
6. Check, Launch, Observe, Land
7. Download full raw data and flight logs for data alignment post processing

III] Post Flight

8. Post Processing | **9.** Secondary Data Products / Analysis | **10.** Fusion / Integration | **11.** Reuse

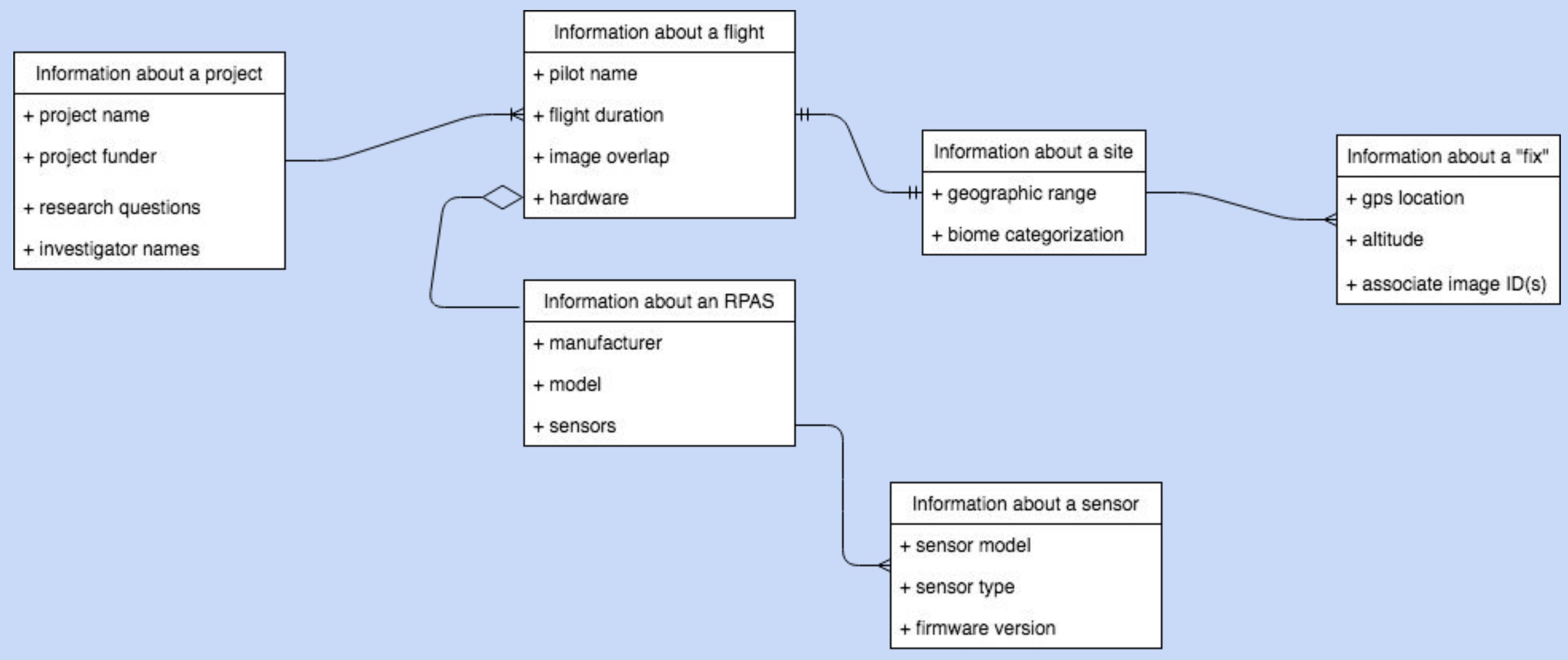
8. Post processing:
   a. Remove noise, align with flight logs GPS values, combine samples from multiple gas sensors, correct for motor noise, carry out statistical analyses and create data visualisations
   b. Align and stitch imagery into DEM, subtract from bare earth DEM, calculate field total biomass from volume and biomass sample. Compare with nutrient run-off data over time (weeks of flights)
   c. Use computer vision, a trained AI, and/or indices to detect and quantify species of interest in images

9-11. Secondary data products/ fusion
   d. Gas contour plots, time series of GHG emissions through annual agriculture cycles, proof of mitigation for governmental carbon trading programs. Combine gas data with satellite and governmental data for larger scale inferences and reporting.
   e. Compare biomass and spectral signatures with: satellite analyses, farmer reported yields, reported and satellite detected downstream algae blooms (vs what nutrient and biomass would predict)
   f. Long term monitoring of animal population trends as monitored from other data (manual sampling, vegetative impact, feces sampling). Or in the case of flora: long term monitoring of spread/reduction of rare/alien flora data are combined with intervention program data/goals

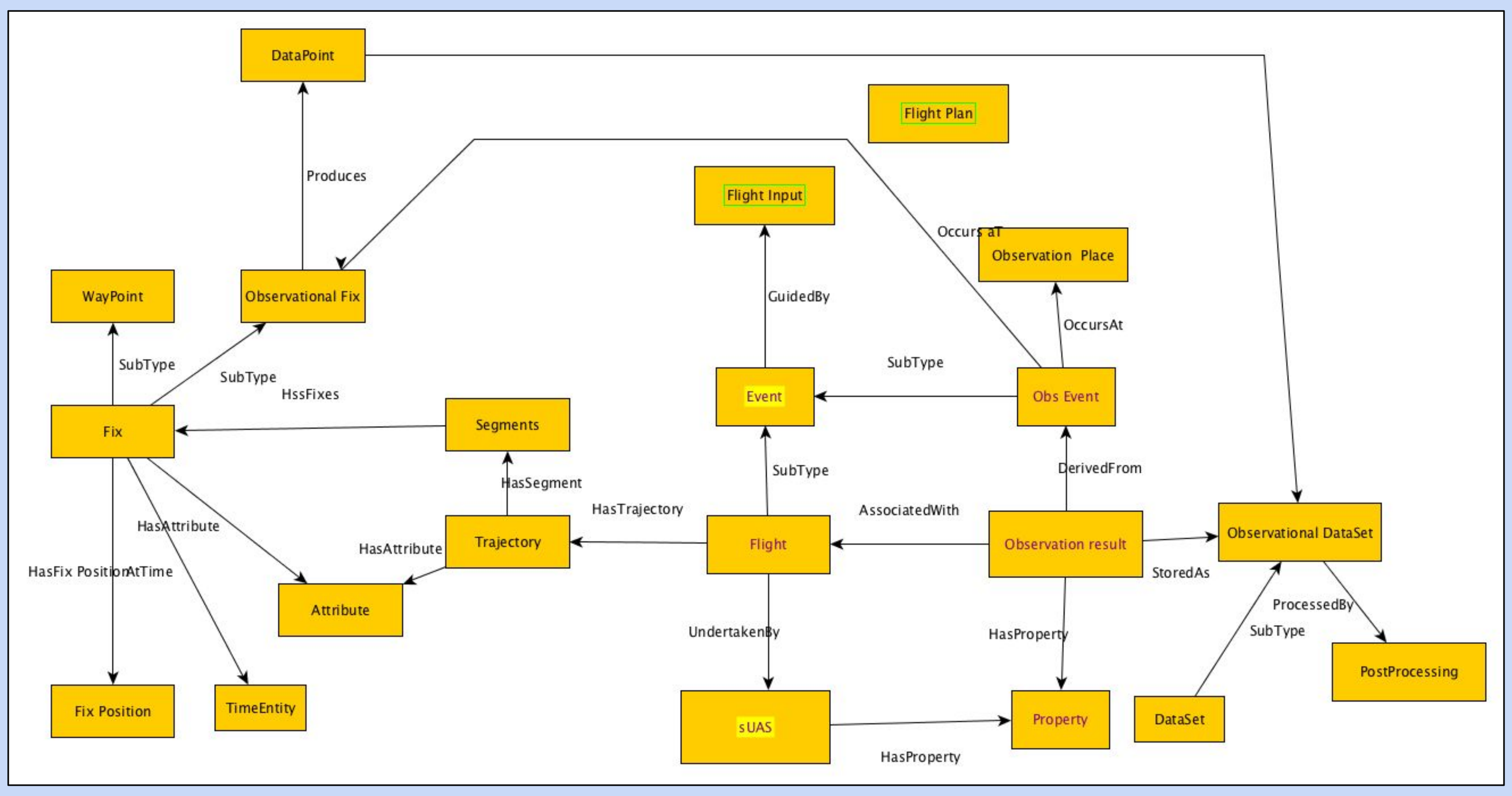
Results to date: draft MIF and drone ontology patterns

Key classes on information for ensuring the FAIRness of RPAS data include:

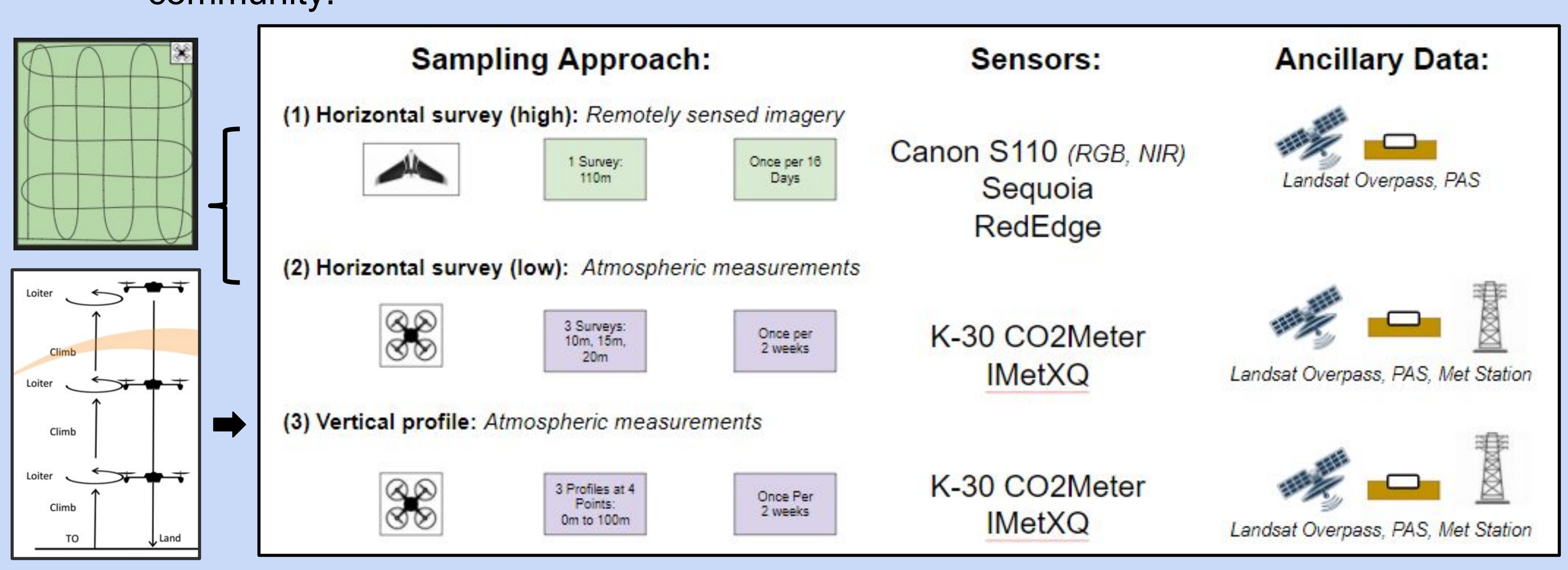
- *Project data:* e.g. project name, project funder, investigator names, and research goals/questions
- *Flight data:* pilot name, flight duration, flight path, image overlap, hardware specs (including RPAS make/model, sensor model, software details
- *Site data:* the geographic range and environmental context of a study
- *Observation or “fix” data:* data and metadata about individual observations, including specific gps location, altitude, other scientific parameters.



We are also working with the VoCamp community to develop ontology patterns to support the semantic annotation of drone data (https://github.com/Vocamp/dronedata). Below: a draft of a design pattern modeling flight data.



RPAS data (right) can be captured in a number of different ways. The following three examples (below) are from a use case provided by Lindsay Barbieri - future work will ensure that our models map to the many methods used by her and others in this community.



How to get involved at RDA:

- Thursday Breakout 6 15:30-17:00 - Room B95!
  - Joint meeting of: Small Unmanned Aircraft Systems (aka RPAS) data IG, Research Data Collections WG, and From Observational Data to Information IG
- Thursday Breakout 5 13:30-15:00 - Room A05!
  - Joint meeting of: Data for Development IG, Ethics and Social Aspects of Data IG, Small Unmanned Aircraft Systems (aka RPAS) Data IG, Health Data IG, RDA/NISO Privacy Implications of Research Data Sets IG

ESIP The Federation of Earth Science Information Partners — University of Michigan — University of Notre Dame — The University of Vermont — Research Data Alliance