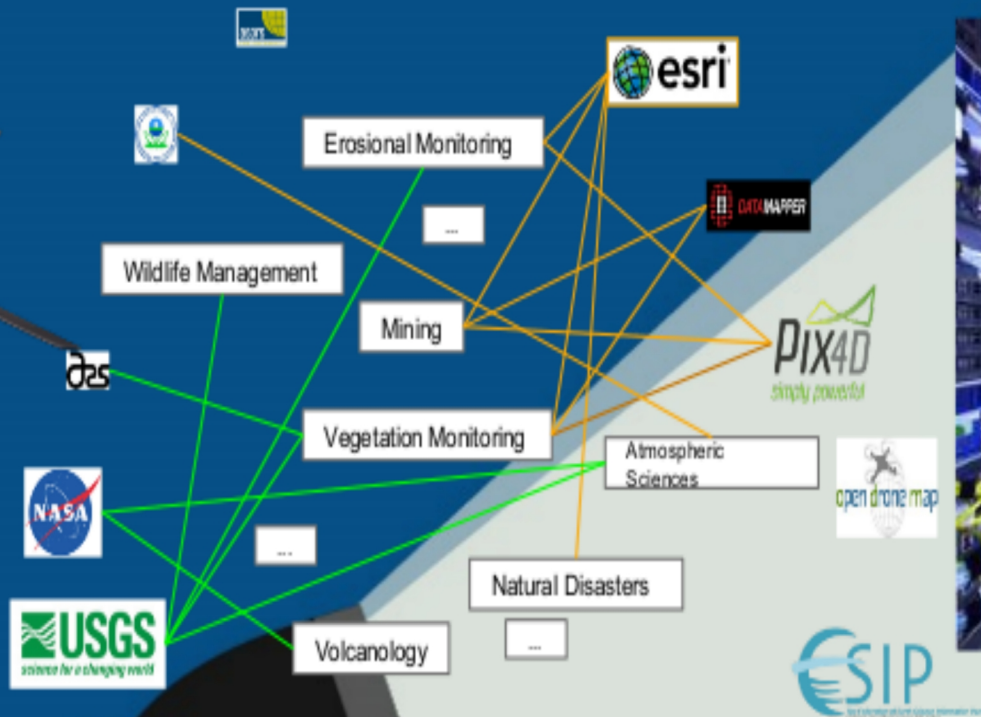# Science sUAS data standards

Introduction

- Research Data Alliance introduction

- sUAS in science use survey

- sUAS data management

- RDA sUAS data Interest Group

Research Data Alliance

- Research Data Alliance:

  - Builds the social and technical bridges that enable open sharing of data.

  - Vision: Researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society.

  - Launched as a community-driven organization in 2013 by the European Commission, the United States National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation.

    - >4,780 members, from 117 countries (January 2017),

    - A neutral space to develop and adopt infrastructure that promotes data-sharing and data-driven research

EUROPE

UNITED STATES

Research Data Alliance

- Interest Groups
  - Geospatial IG
  - Mapping the Landscape IG
  - Biodiversity Data Integration IG
  - Marine Data Harmonization IG
  - Metadata IG
  - Data Foundations and Terminology IG
  - Legal Interoperability IG
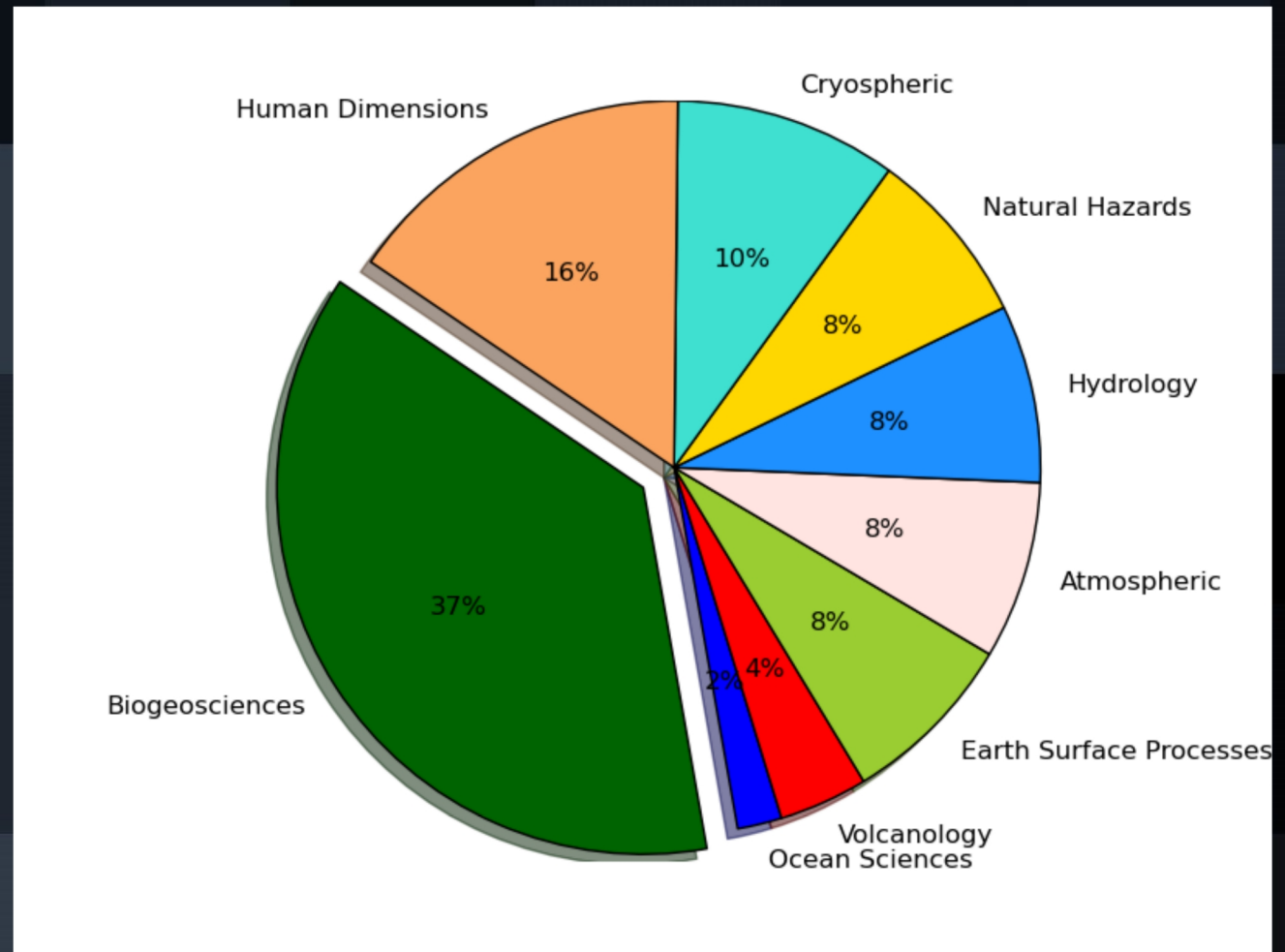 - Weather, climate and air quality
  - Structural Biology IG

- Working Groups
  - BioSharing Registry: connecting data policies, standards & databases in life sciences
  - Metadata Standards Catalog WG
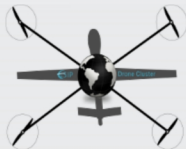 - Rice Data Interoperability

A quick survey
sUAS use in the Earth Sciences by domain
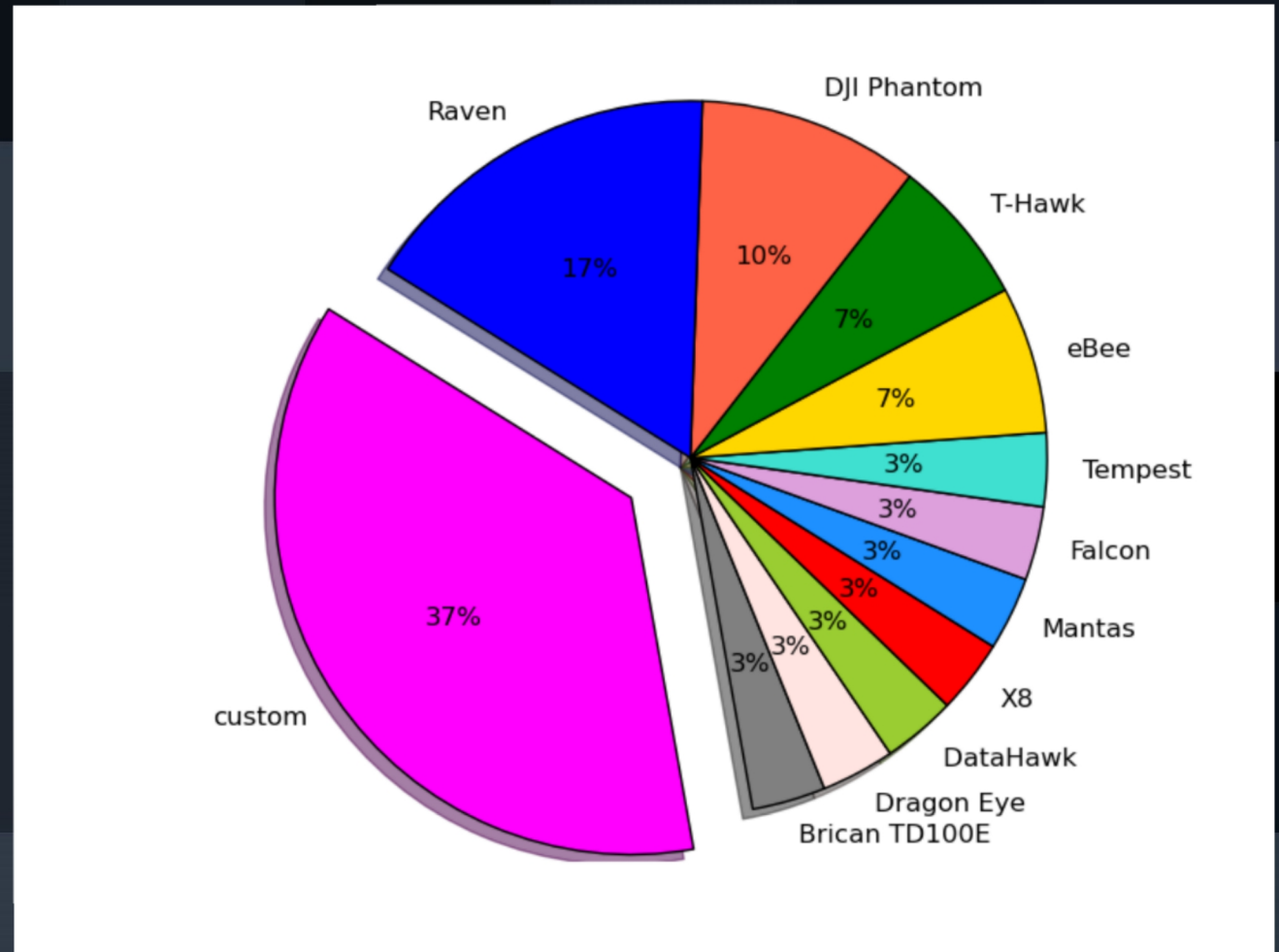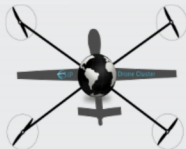
- 37% Biogeosciences

https://osf.io/nuvem/

A quick survey
sUAS flight platforms in the Earth Sciences

- 37% Custom build
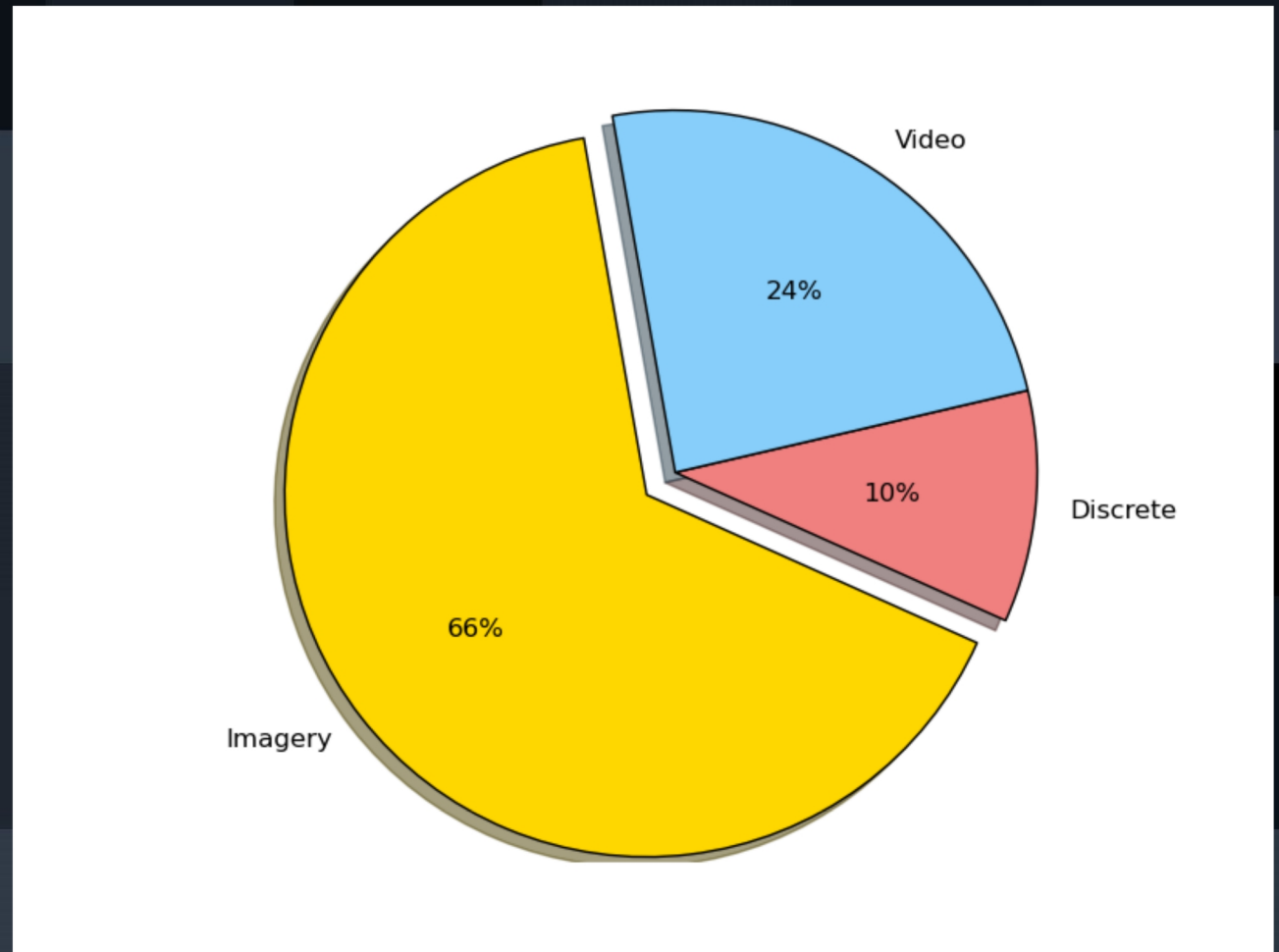

https://osf.io/nuvem/

A quick survey
sUAS sensory data output type in the Earth Sciences

- 66% Imagery

https://osf.io/nuvem/

sUAS Data Management

Value of Small Unmanned Aerial Systems [sUAS] to science:

- Data on demand / Data currency
  - Event-reactive data capture
  - High frequency resampling
- Greater accessability:
  - Higher resolutions
  - Novel locations
- Improved safety
- Lowered environmental impact
- Cost savings

Value of making an entirely open sUAS for science stack:

- Data standardisation >> FAIR data
- Enable citizen science drones
- Facilitate academic-commercial-government collaboration
- It's already 90% there
  - Improve ease of use
  - Improve sensor integration
  - Improve mission planning
- Increase innovation potential
- Enable scientific work globally

"Drones will become a standard tool for Earth Scientists within the decade" - Unknown, heard in a USGS public telecon cica mid-2015

sUAS Data Managment

sUAS data is:

- chaotic by definition:
  - Temporally
  - Spatially (and globally sparse)
  - By type (highly heterogeneous)
- Geotagged (or it should be)

- Doesn't require reinventing the wheel...

CO2 illustration
Data reuse example

Data-use1:

Provide farmers with the information needed to judge the effectiveness of a GHG emissions mitigation strategy
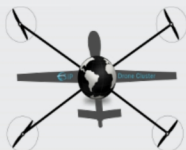
Potential Data-reuse3:

IPCC and governmental regulators estimates of regional agricultural GHG emissions for policy decision making.

Potential Data-reuse2:

Provide farmers with data to justify carbon tax credit due to emissions reduction

Potential Data-reuse4:

GHG emmisions estimates as model input.

CO2 Flux
- 51.61084064
- 58.35990019
- 74.89051927
- 82.07326018

Value
High : 475

Low : 420

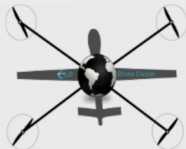# CO2 illustration
Data questions example

Data:

- Discrete CO_2 point data
  - XML/CSV/ Geojson/NETCDF /Shapefile/?
- RGB Imagery
  - GeoTiff/Jpeg/RAW/?
- Elevation
  - DEM/ESRI/?
- Static chambers
  - XML/CSV/txt/?

Metadata:

- Temperature /Wind speed /Humidity /….?
  - Units? C/F, Kts/kmph/mph/ , …?
- Flightlog?
  - Format?
  - Params?

Sensor:

- Last calibration /sampling frequency, warm up period / …?

CO2 Flux
- 51.61084064
- 58.35990019
- 74.89051927
- 82.07326018

Value
High : 475

Low : 420

UNIVERSITY OF NOTRE DAME

The UNIVERSITY of VERMONT

ESIP
The Federation of Earth Science Information Partners
Fostering connections to make data matter

Examples to learn from
Underwater Gliders

Integrated Ocean Observing System (IOOS) Underwater Gliders

- CF-compliant NetCDF

- Preserve the original resolution of the data sets

- FDGC and ISO 19115 metadata

- Designated file name schema:

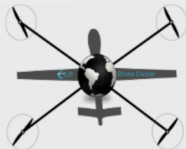glider_yyyymmddTHHMMSSZ_rt.nc (ISO 8601 date format)

- OGC SOS, OpenDAP

https://github.com/ioos/ioosngdac/wiki

UNIVERSITY OF NOTRE DAME
VITA · DULCEDO · SPES

The University of VERMONT

Examples to learn from
Data Portals

American Society for Photogrammetry and Remote Sensing (ASPRS)

- Aerial Data Catalog: http://dpac.asprs.org/

Examples to learn from
Current sUAS data management practices

- United States Geographical Survey (USGS)
  - PROJECT/
    - Project.meta
    - MISSION1/
      - Mission.meta
      - COLLECT1/
        - Collect1.meta
        - SENSOR/
          - RAW
          - DEM
          - ORTHO
          - SUPPORT FILES

http://uas.usgs.gov/pdf/UAS_Data_Mgt_Plan_V1_0_Final.pdf

- University of Vermont Spatial Analysis Lab (SAL)
  - UAV/Missions/ProjectName*/
    - Data/Date/
      - Flight#**/
        - Subdirectories: IMG, KML, Pix4D
        - Bbx file (flight log)
      - Mosaics/Flight#**/
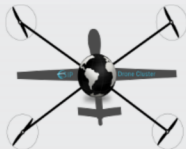
*87 as of August 2016

**Ebee non unique identifier

Examples to learn from
Further potential references

- Open Imagery Network (OIN)
  - Open Aerial Map
    - Data:
      - Satellite/Balloon/Kite/UAS/Aircraft: GeoTiff or RGB
      - Imagery only: 2901 images (as of 16:46 EST, 12 September 2016)
    - Metadata: Title, Sensor name, Date, Location, Provider, License
    - QGIS plugin

- ScienceBase

- EarthExplorer (USGS)
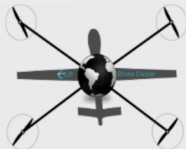
- GeoNetwork

- Industry sUAS operators
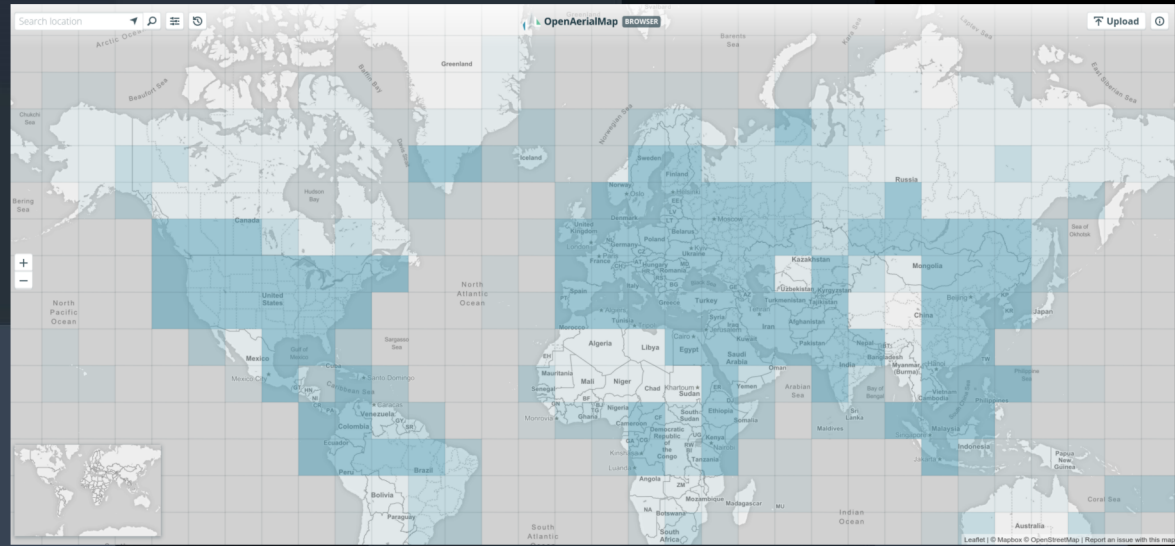
A straw-man
A data management plan for a scientific sUAS

- Raw Data archived
  - Imagery: GeoTiff*
  - Video: MXF*
  - Discrete: HDF5**
  - Flight log: human readable DroneCode
  - Procedure: JSON
- Metadata: ISO19115/ (xml)
- Sensors: Sensor MI***
- License
  - Creative commons attribution

* Sensor native data format while onboard drone

** Geojson while onboard

*** JSON while onboard

- And many more points:
  - Sensor metadata (Sensor MI)
  - File naming convention?
  - Variable naming convention?
  - Higher tier data sets?
    - Point clouds
    - Mosaics (DEM/etc)
    - Fused data sets
  - Data Sharing?
    - OGC SOS/OpenDAP/etc
  - Citation mechanisms?

RDA sUAS IG

RDA sUAS IG

- Motivation

  - Standards will lower the barrier to entry and innovation

  - Science sUAS is currently hampered by no standards inhibiting development

  - To enable good scientific practice: sUAS captured data is - for the most part – currently not FAIR

RDA sUAS IG

Objectives:

- Provide a venue for conversation around data standards and best practices for scientific sUAS

- Identify common and divergent data needs across sUAS implementations in different domains.

- Be a community aggregation point for others in the field who are currently isolated.

- Identify community partnerships, including with industry, tech companies/manufacturers, and computing organizations and infrastructures.

- Provide a venue for ongoing community discussion around the legalities, logistics and opportunities governing sUAS use, given that sUAS are a relatively new data collection platform.

RDA WGs
RDA Working Groups

WGs:
- Ideally, three or more continents
- 2-4 co-chairs leading the initiative
- Measurable outcomes
- 12-18 months

Some current WGs:
- BioSharing Registry: connecting data policies, standards & databases in life sciences WG

- Wheat Data Interoperability WG

- Research Data Repository Interoperability WG

- RDA / TDWG Metadata Standards for attribution of physical and digital collections stewardship

- Metadata Standards Catalog WG

- RDA Transportation WG

RDA WGs
RDA Working Group Review Process

Details:
http://tinyurl.com/hjsanou
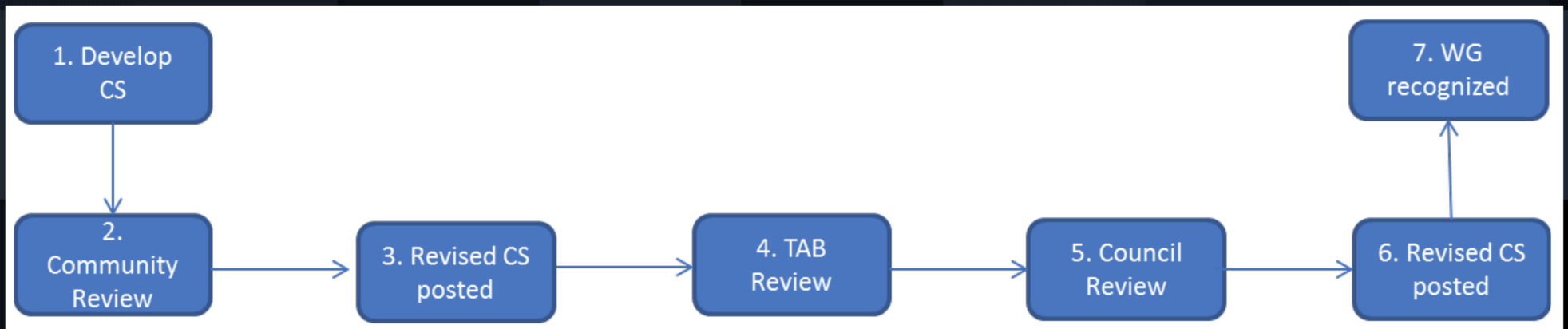
Interactive

- Some possible WGs

  - Minimum metadata parameters for science sUAS data capture.

  - Best practice data formats (while onboard, during analytics, for long term publishing,...).

  - Flight-log publication best practices: formats, readers, parameters

  - Naming convention adoption in such a heterogenous domain?

  - Procedures for gas sensing

- Post-its

- Gdoc: tinyurl.com/jqtvbu5

- RDA wiki

Interactive

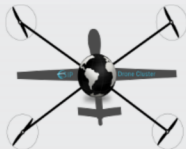tinyurl.com/jqtvbu5

Questions

- What does this/another of your communities need in terms of community data discussion?

- What non optical data do you most often fuse optical with (what standards do you need to be interoperable with)?

- Who else in the EU should we be talking to?

- What data formats/standards do you use?

- What metadata formats/standards do you use?

Conclusions:

Contact us:

Jane Wyngaard, PhD
jwyngaar@nd.edu
@jrwyngaard

Lindsay Barbieri
lkbar@uvm.edu
@barbieriiv

RDA Small Unmanned Aircraft Systems' Data IG:

tinyurl.com/z5gf4zr

ESIP Drone Cluster (Open Science Framework)
https://osf.io/nuvem/

RDA 9th Plenary Meeting
Data Infrastructures for Open Science
5-7 April 2017, Barcelo Sants Hotel, Barcelona, Spain

Organised by Barcelona Supercomputing Center (BSC) with the support of RDA Europe

RDA RESEARCH DATA ALLIANCE

ESIP The Federation of Earth Science Information Partners — Fostering connections to make data matter

UNIVERSITY OF NOTRE DAME — VITA DULCEDO SPES

The UNIVERSITY of VERMONT