RDA Plenary 10: sUAS Data IG

Link to this page: http://bit.ly/2xpjNRk

Slides:

https://docs.google.com/presentation/d/1Jue9BAnWoezEFHcEsYW4xwhijLcdF0OQlgh-UJegBWo/edit?usp=sharing

Participants:

Name	Email	Organization	Interested because … AND other relevant RDA IG or WG
Jonathan Petters	jpetters@vt.edu	Virginia Tech Libraries (Data Services)	Some researchers we support at VT use UAS; background in meteorology - RDA domain repository interoperability group might be relevant
Pascal Bonnet	pascal.bonnet@cirad.fr	CIRAD France	We support an emergent community of research (interdisciplinary) using sUAS with sensors attached, we would like to know if there are developments in specific standards etc.. regarding data they produce (flight records etc..)
Marcel Gangwisch	marcel@gangwisch.eu	University of Freiburg	General interest in spatial geodata, Background in Computer Science
Andrea Thomer	athomer@umich.edu	University of Michigan	Working w/ESIP drone cluster, interested in data mgmt and provenance
Howard Silver	Hsilver@mit.edu	MIT	libraries are supporting drone

			use and collecting researcher data
Tim Barnes	tdba@bas.ac.uk	British Antarctic Survey	We use a lot of UAVs in the Antarctic
Charles Vardeman	cvardema@nd.edu	University of Notre Dame - Center for Research Computing/Computer Science and Engineering	Ontology, Vocabulary and ontology patterns. Post-disaster reconnaissance drone data.
Jane Wyngaard	jwyngaar@nd.edu	University of Notre Dame - Centre for Research Computing	Keen for global conversation on making sUAS data FAIR
Cyndy Parr	cynthia.parr@ars.usda.gov	US National Agricultural Library	Preparing for cataloging and managing UAV data and metadata for reuse
Andrew Treloar	andrew.treloar@ands.org.au	Australian National Data Service	Drones as instrument platforms, place in research data ecosystem
Hamish Holewa	hamish.holewa@csiro.au	Atlas of Living Australia	Integration of biodiversity data from devices including Drones, Camera Traps, Bioacoustics and other sensors. Interested in workflows that develop occurrence, abundance, trait data.
Kimberly Silk	ksilk@crkn.ca	Canadian Research Knowledge Network	Data Librarian interested in emerging areas of research data management. RDA newbie.

Useful Links:

Earth Science Information Partners (ESIP)
→ Wiki: http://wiki.esipfed.org/index.php/Drone_Cluster
→ Open Science Framework: https://osf.io/nuvem/

Research Data Alliance (RDA)
→ sUAS Data IG:
https://www.rd-alliance.org/groups/small-unmanned-aircraft-systems%E2%80%99-data-ig

- ESIP Winter meeting (11-13Jan)

Challenges:

1. Best practices regarding the long term storage of sUAS data.

In pursuit of ensuring data *Accessibility,* and given the *BigData* nature of sUAS data, best practices around which data product levels and what versions should be stored needs to be investigated and discussed. Along with the practical considerations of storage capacity, the many potential future data reuse and unanticipated data fusion applications a given data set might have value in should also be considered.

2. Standard data and metadata formatting best practices

These are needed in order for sUAS captured data to be easily made Findable, Accessible, Reusable, and Interoperable FAIR, per standard good scientific data management guidelines. Having such will also; lower the barrier to entry in using sUAS for science, and open the door wider to science sUAS specific software and hardware development by making it easier for developers.

3. Flight procedures for different data need to be defined and articulated in a common language

Well tested and defined best practice methodologies and procedures for flying various sensors on sUAS. For practical purposes and because we need to be able to document the data for reusability

4. Standard sensor calibration and use procedures need to be defined and articulated in a common language

Calibration and use procedural best practices and standards are needed. Sensor specific guidelines should be provided by the manufacturer, however, the community should define

sensor operation and maintenance, and metadata requirements for different sensor genesis. For practical purposes and because we need to be able to document the data for reusabaility

5. Best practices regarding data post processing and error analysis methodology needs to be outlined and named:

So as to provide the correct metadata with higher level data products. It's relatively easy to introduce unintended error and artifacts into your data when processing it, we need to both describe best practices and quantify acceptable error tolerances for the primary sensor taxonomy branches and the associated processes.

6. Basic Standard Guidelines for Safe Operation

While many countries have begun to settle on regulations regarding operation of sUAS in their airspace, most organisations are still in the early stages of writing their own internal policies and protocols. Aside from liability legality issues, a self governing science sUAS community will go a long way to developing goodwill with the public, especially as privacy and military concerns around sUAS remain.

7. Need for Introductory information and education:

As the domain grows their in an increasing demand for introductory information regarding the multitude of new expertise needed to even begin entering this space. For to sUAS to become a ubiquitous and safely used tool of the trade, we need to begin equipping new members.

8. Technical challenges and research opportunities particularly given the BigData nature of much of sUAS data:

1. Image processing - Machine learning applied to classification, and enumeration
2. How to fuse insitu drone data (3D data) with satellite, sensor network, traditional ("2D" and "1D" data)
3. Real time dynamic mission response to data
4. Multiple sUAS coordinated data management

An opportunity not to miss

If we do this now we save our future selves decades of headaches because…

Push factors:

- We don't leave sUAS data open to extensive suspicion and challenges on the grounds of unspecified quality guarantees
- Won't have to retroactively make the above and then try to rescue all the data captured up until then
- We won't loose all the data that is currently being lost through poor management
- We won't have to keep on building as many of our own custom time consuming and expensive data management and analytics tools and pipelines.

Pull factors: and if we do do it we gain so much...

- We can go to commercial manufactures and require they provide the above as far as it's their domain
- We lower the barrier to entry for those not yet using sUAS
- We make our own jobs easier but having this already formulated and in place for our own campaigns
- We make it easy for the tech domain to make us the tools we want - think ready made data analytics and processing, think web portals for sUAS data hosting with built in quality checks, think cloud tools for analytics that also enables pulling in other data sets (sUAS and other)...
- We'll see more of our data reused and cited

Notes:

- OGC - have an embedded platform for machine vision that integrates with multiple Autopilots
- What metadata schema would fit (perhaps select by scientific intent)
- We can learn from Ag and Military (have some public metadata stds)
- Should we include non flying robots? OGC's group is UXS
- We should go to commercial community with things we want
- For open citizen science commenter (open source tools) - nothing fully open yet but, Modern Blanket Throw - drones and teenagers: https://www.alaska.edu/epscor/the-modern-blanket-toss/. And POTENTIAL with open SkyRocket : http://sky-viper.com/ ('Toy' category and open)
- In Ag, lots of hyperspectral/multiple sensors etc - but commercial platforms don't release everything by default

ESIP Winter meeting: 11-13 January Bethesda MD - will be in close collaboration with OGC and other Federal agencies based in MD area

RCN in works to run workshops on developing best practices- email us if you're interested in planning with us

- Comment from libraries community - have a workflow prototype
- New RDA IG: From observational data to information in a semantic way
- There's lots out there already for capturing workflows
- NATO created a meta-standard of "for this application" this is the standard of standards you should use. Have we decided to do this too and if so what organisation would 'own' that
- Interoperability - address **what** not **who** first
- Also need best practices in what's stored/shared/etc - given accumulating data sizes….
- Open Topography just starting to explore how to deal with drone data as they start being given it

Who's here and which other RDA groups should we work with

- Domain repository interoperability WG (asking what is common)
- Biodiversity IG
- Chemistry
- Sensor platforms
- Ag semantics group
- Array DB
- Science Gateway, Virtual research Env
- Collections group
- Precision Ag
- Vocab and Semantics Services group
-

HERE BE DRAGONS

Other Relevant Notes from RDA:

GLOBIS-B
Workshop - Observations/Primary Data to EBV usable Data (biodiversity)

Slides:

Measurements with comparable units , similar observations →

Agree on what Raw data is.

Structuring well-forming packaging, adding 3rd party detail

Designing Workflows, thought experience working on specific use cases.

Paper To read:
Kissling et al. 2017, doi: 10.1111/brv.12359
http://onlinelibrary.wiley.com/doi/10.1111/brv.12359/abstract

Marcus:

“Pattern”:

- Primary observational (sensor) data
- Data interpretation
- Derived information about observed environment
- Information is formal (machine readable)

3 Use Cases:

Aerosol:

- Event Ontologies (?)
- OGC Services -- observations and measurement in a standardized way.
  - Researcher and interpretation is heterogeneous…
  - RDF → describes processes

Interpretations of data on data cycles:
Coombs Theory of data

linked open event ontology rdf
LODE

environment ontology - Pier
process vs event

WPS - web processing service model processes