Managing your drone data through the data life cycle: RDA guidelines for FAIR and responsible RPAS data use

Authors: Alice Fremand[1], Jens Klump[2], Sarah Manthorpe[3], Mari Whitelaw[1], France Gerard[4], Wendy Garland[5], Thabo Semong[6], Charles George[4], Hugo Ricketts[7], Nina Grau[8]

[1] *UK Polar Data Centre, British Antarctic Survey, UK*
[2] *Commonwealth Scientific and Industrial Research Organisation, Australia*
[3] *University College Cork, Ireland*
[4] *Centre for Ecology and Hydrology, UK*
[5] *Centre for Environmental Data Analysis, UK*
[6] *Botswana International University of Science and Technology, Botswana*
[7] *University of Manchester, UK*
[8] *CODATA, France*

Introduction

The use of Remotely Piloted Aerial Systems (RPAS), also referenced as Uncrewed Aerial Vehicles (UAVs) and more generally as drones, is increasingly prevalent across various scientific disciplines, enabling the collection of large volumes of data for diverse research applications. However, the volume of data generated and the absence of standardised workflows often complicate data sharing and publication. As part of the Research Data Alliance (RDA) Small Uncrewed Aircraft and Autonomous Platforms Data Working Group, we have developed guidelines on how best to improve the Findability, Accessibility, Interoperability and Reusability (FAIR, Wilkinson et *al.* 2016) of these data while taking into account legal, privacy and ethical considerations. In this document, we summarise our working group activities and present our recommendations with a particular focus on mission planning, data collection, ethical considerations, data pipeline description, data storage, data format, and interoperability aspects.

Presentation of the RDA small Uncrewed Aircraft and Autonomous Platform Data Working Group

The Research Data Alliance (RDA) Small Uncrewed Aircraft and Autonomous Platform Data Working Group, established in June 2024, is dedicated to enhancing data sharing and interoperability for RPAS and autonomous platforms. These technologies are revolutionising data collection by offering higher temporal and spatial resolutions and enabling data collection in hazardous areas. As part of its activities, the working group has compiled use cases showcasing RPAS applications across various research disciplines. This compilation has been instrumental in documenting best practices and identifying gaps and challenges researchers have while handling their RPAS-derived data. Drawing on these insights, the group has now developed guidelines and recommendations to improve RPAS data management throughout the research life cycle.

General Considerations

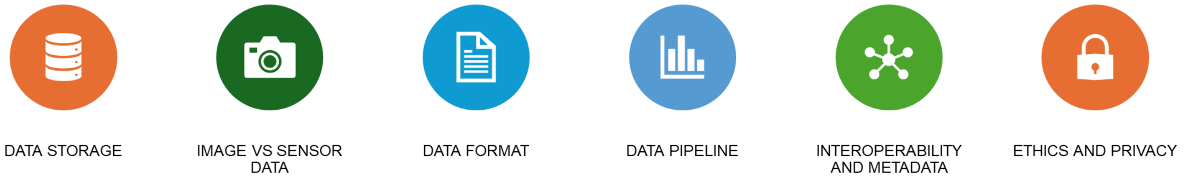
The recommendations are designed to support researchers, scientists, and data practitioners working with Remotely Piloted Aircraft Systems (RPAS) data across all disciplines. By “RPAS,” we include the full spectrum of remotely piloted systems—whether very large, complex platforms or small, lightweight units. We developed the recommendations with a data management perspective and more specifically with both the FAIR (Findable, Accessible, Interoperable, and Reusable) principles and the CARE (Collective Benefit, Authority to Control, Responsibility, and Ethics) principles in mind. They aim to promote responsible data stewardship and interoperability while promoting safe and ethical use of RPAS. While the recommendations are very general, we integrated some examples and references to allow users to further develop their knowledge on relevant associated topics. The open and accessible format of the recommendations is intended to lower entry barriers for Small and Medium-sized Enterprises (SMEs) acting as service providers in RPAS and autonomous platform missions. The growing use of RPAS in environmental monitoring and infrastructure maintenance is expected to contribute meaningfully to the achievement of the following United Nations Sustainable Development Goals (UN SDGs):

- **SDG 2:Food security and nutrition and sustainable agriculture**
  RPAS are being used in weed detection, plant phenomics measurements, and agricultural management, contributing to increasing crop yields and sustainable agriculture.
- **SDG 6: Clean Water and Sanitation**
  RPAS are being used in monitoring water bodies and wetland ecosystems, contributing to a better understanding of the state of freshwater ecosystems and water quality.
- **SDG 14: Life in Water. Ocean and Seas**
  RPAS are being used in marine monitoring, contributing to a better understanding of the state of the ocean and marine ecosystems.
- **SDG 15: Life on Land. Biodiversity and Ecosystems**
  RPAS are being used in terrestrial environmental monitoring in numerous ways, including species and weed detection, contributing to a better understanding of terrestrial ecosystems.

Methodology

This work has been led by the Research Data Alliance (RDA) Small Uncrewed Aircraft and Autonomous Platform Data Working Group. To better identify the challenges and needs of the community, the work has been based on previous work undertaken by the associated Interest Group and chair/members activities, as well as community-led activities and workshops. The full list of associated publications and engagement activities is provided in the Appendix 1.

From this comprehensive consultation, the following recurring challenges were identified:



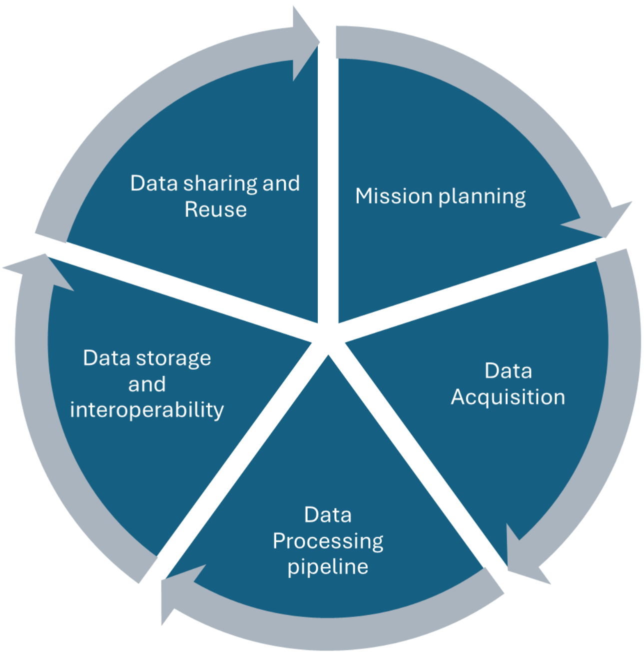
- **Data Storage:** RPAS generates high volumes of imagery and sensor data, often exceeding traditional storage capacities.
- **Image and sensor data:** Management of sensor and image RPAS-derived data is very different and each type of data will require specific metadata.
- **Data format:** Due to the variety of disciplines, there is a very wide range of data formats used. It is to be noted that much of the information related to the RPAS, such as the position, altitude and attitude, is often saved in a manufacturer proprietary format and can't be easily read by open software.
- **Data pipeline and processing description:** Absence of best practices in describing complex pipelines, often combining open and proprietary software.
- **Metadata and Interoperability:** Inconsistent metadata and a lack of standards prevent discovery and reuse.
- **Ethics and Privacy:** RPAS often captures sensitive data requiring ethical consideration.

These challenges are addressed throughout the data life cycle recommendations that follow.

Recommendations

We present here the recommendations for RPAS data management best practices. This work aims to guide researchers, data managers, and institutions in adopting ethical and interoperable RPAS data practices.

We have structured our recommendations around the following research data life cycle:



1. Mission planning

Mission planning refers to any pre-flight activities required for a successful survey.

2. Data Acquisition

The data acquisition step refers to all the activities surrounding the collection of RPAS data. It also covers legal/ethical requirements that need to be taken into consideration when on the field.

3. Data Processing Pipeline

The data processing pipeline refers to all the steps required to create a data product after the data acquisition. This includes all the steps from pre-processing to derived products, including calibration and quality assessment tests.

4. Data Storage and Interoperability

Data storage and interoperability refer to information regarding storage, but also ontologies and metadata recommendations to promote better interoperability.

5. Data Sharing and Reuse

The data sharing, access and reuse section refers to any requirements related to the publication of the output dataset in order to foster their reuse.

Each stage is addressed with:

- A recommendation section. This section gives general principles to follow. Here we distinguish data and metadata specific recommendations to privacy, legal and ethics recommendations.
- A resource and additional information section. This section includes more practical and actionable recommendations for specific domains or external resources of interest.
- A use case section that proposes examples of implementation.

The diagram below gives an overview of these recommendations:

5 STEPS
To improve Remotely Piloted Aerial System data management

MISSION PLANNING

- Choose a platform and sensors that are in line with your survey requirements.
- Plan your survey lines considering the restrictions of the environment and define your Standard Operating accordingly.
- Make sure you prepare all the required equipment and storage for acquisition before you are in the field.
- Write a Data Management Plan that documents:
  - Which data to keep for long-term storage
  - Data volume, file formats, and metadata requirements

- Define clear purposes for data collection and communicate your plan with local communities
- Get familiar with legal requirements and civil aviation rules, and ensure any institutional procedures are adhered to
- Apply CARE principles if working with indigenous communities

DATA ACQUISITION

- Use the data collection checklist to record all the required information for future processing
- Calibrate your instruments following best practices

- Adhere to relevant national civil aviation rules, institutional procedures and access restrictions to undertake safe operations
- Build trust with local communities through consistent ethical behaviour
- Follow best practices regarding environmental and wildlife protection.

DATA PROCESSING PIPELINE

- Document the processing workflow following the data processing checklist
- Separate raw (level 0), pre-processed (level 1) and processed data (level 2 onwards) and use version control
- Provide a quality statement giving insights about quality assessment and control (QA/QC) applied to your data

- Blur or anonymise personal data if applicable
- Reduce the precision relative to sensitive location (e.g. of endangered species) if applicable

DATA STORAGE AND INTEROPERABILITY

- Use recognised metadata standards used in your disciplines and follow the recommendations for specific RPAS metadata
- Use open formats, where available.
- Adopt ontologies to enhance semantic clarity and interoperability.
- Utilise Persistent Identifiers for Instruments (PIDINST) for accurate referencing and traceability.

- Establish secure storage with access restrictions in alignment with requirements

DATA SHARING AND REUSE

- Share your data through a trusted data repository using open formats and all the necessary metadata required for reuse

- Adhere to privacy laws and ethical guidelines, particularly regarding sensitive data by:
- Following the principle 'as open as possible, as closed as necessary'
- Define a license in line with the requirement of privacy

1. Mission planning

Before heading to the field, it is essential to prepare carefully for your RPAS mission. From legal requirements, institutional rules, to ethical consent, RPAS operations need to be planned in advance. This includes the creation of a data management plan to define how your data will be handled as part of your project.

Recommendations:

Data and metadata

- Choose a platform and sensors that are in line with your survey requirements.
- Plan your survey lines, taking into account the restrictions of the environment, and define your Standard Operating Procedure (SOP) accordingly.
- Make sure you prepare all the required equipment and storage for acquisition before you are in the field.
- Write a Data Management Plan that documents:
    - Which data to keep for long-term storage
    - Data volume, file formats, and metadata requirements

Privacy, legal and ethics considerations

- Define clear purposes for data collection and communicate your plan with local communities
- Get familiar with legal requirements and civil aviation rules, and ensure that any institutional procedures are adhered to
- Apply CARE principles if working with indigenous communities

Resources and additional information

➢ The first step to undertake a successful RPAS mission is to choose your RPAS platform in accordance with your survey objectives. Your choice will depend on external factors such as the characteristics of your survey area, the spatial resolution needed, the flight duration or the sensors you will need. For example, the use of a multirotor drone is usually preferred for small-scale or detailed surveys as it offers precision and maneuverability, while you might want to choose a fixed-wing platform for covering large areas efficiently. Like the choice of a platform, the selection of the sensors must be aligned with the requirements of your research.

Specific recommendations around RPAS legislation, platform choice, and planning software are provided in the following research paper:

- Tmušić, G.; Manfreda, S.; Aasen, H.; James, M.R.; Gonçalves, G.; Ben-Dor, E.; Brook, A.; Polinova, M.; Arranz, J.J.; Mészáros, J.; Zhuang, R.; Johansen, K.; Malbeteau, Y.; de Lima, I.P.; Davids, C.; Herban, S.; McCabe, M.F. Current Practices in UAS-based Environmental Monitoring. *Remote Sens.* **2020**, *12*, 1001.https://doi.org/10.3390/rs12061001

- To undertake safe and consistent RPAS surveys, it is highly recommended to describe your protocol into a SOP. This document will detail all the specific steps needed for undertaking your survey safely. To plan your survey lines, it is important to take into account the environmental constraints of the survey, such as terrain, wildlife, weather, and airspace legislation. Before flying, it is essential to check for powerlines, visual obstructions, towers, buildings, trees, shrubbery, people or wildlife. You might also want to identify alternate take-off/landing locations in case of issues. Additional information specific to your survey might also be documented, such as the weather conditions under which your platform can safely operate, as this helps prevent damage and ensures reliable performance. To assist you in developing your SOP, the following resources might be of interest:

  - Victorian Coastal Mapping Project. (2022). *Standard Survey & Operating Procedures for UAV Mapping Coastal Erosion* (p. 54). Melbourne, VIC, Australia: Deakin University. Retrieved from https://ars.els-cdn.com/content/image/1-s2.0-S0278434322001534-mmc1.pdf
  - National Severe Storms Laboratory (2023). UAS Standard Operating Procedures. https://inside.nssl.noaa.gov/uas/wp-content/uploads/sites/35/2023/07/NOAA_UASD-Standard-Operating-Procedures-NSSL_071923.pdf
  - https://www.airhub.app/resources/news/drone-standard-operating-procedures-in-high-stakes-operation

- Before heading into the field, it is essential to also prepare all the equipment required for your survey. In particular, you might want to check drone batteries, sensors, backup devices, and all the data storage needed for the data acquisition. Surveys are often undertaken in remote locations where access to specific equipment is limited. Bringing a small toolbox or small parts that are easily breakable might save your mission. Sometimes, test flights might need to be done prior to the main survey mission and will need to be planned accordingly. These can be for calibration or to test the equipment and ensure the sensors and platform are fit for purpose.

- A Data Management Plan (DMP) is a live document that explains how you and your colleague plan to collect, manage, store and share the data of your project. The DMP informs the expected data volumes, file formats, and which data products will be retained for long-term storage, including raw, pre-processed and final outputs. It allows you to discuss your strategy to make your data FAIR and CARE. To facilitate the creation of your DMP, online tools are available, such as:

  - **DMPTool** – dmptool.org: A free service primarily used in the U.S. to help researchers create DMPs that meet funder requirements.
  - **DMPonline** – dmponline.dcc.ac.uk: Developed by the Digital Curation Centre (UK), it supports researchers in creating DMPs tailored to institutional and funder guidelines.

- Additional tools might be available at your institute. The following website lists a few others and provides training materials if you want to learn more about DMPs: https://rdmkit.elixir-europe.org/data_management_plan

➢ To get the necessary permission to fly your RPAS, defining a clear purpose for data collection is essential as it will help you discuss with the local authorities and communities and get the necessary permissions for your flights. This is particularly important when imagery is captured in populated areas where additional consent might be necessary to enable compliance with data regulations such as the General Data Protection Regulation (GDPR).

➢ An important step of mission planning is understanding the legal and institutional requirements that apply to your RPAS survey. These rules vary depending on the country and specific area where the survey will take place. Obtaining necessary documentation such as registrations, licenses, airspace permissions, and consent is often time-consuming; it is thus advised to be proactive to respect the timeline of the mission. The following websites give links to different state civil aviation rules as a first step to prepare your mission:

- The International Civil Aviation Organisation (ICAO) provides links to several state guidances on its website and additional resources: https://www2023.icao.int/safety/UA/UASToolkit/Pages/State-Regulations.aspx https://www.icao.int/uas-toolkit-home.
- The European Union Aviation Safety Agency also has a good range of documentation related to RPAS: https://www.easa.europa.eu/en/domains/civil-drones

➢ Ethical considerations when using a RPAS: although the two following resources have been developed for domestic and humanitarian use of RPAS, they provide interesting insights into ethical considerations that are relevant for scientific use:

- Wang, N., Christen, M. & Hunt, M. Ethical Considerations Associated with "Humanitarian Drones": A Scoping Literature Review. *Sci Eng Ethics* **27**, 51 (2021). https://doi.org/10.1007/s11948-021-00327-4
- West, J. P., & Bowman, J. S. (2016). The domestic use of drones: An ethical analysis of surveillance issues. *Public Administration Review*, *76*(4), 649-659. https://doi.org/10.1111/puar.12506

➢ When operating RPAS in areas involving Indigenous communities, it is essential to follow the principles of Collective Benefit, Authority to Control, Responsibility, and Ethics (CARE). The CARE principles provide guidance on ensuring Indigenous Peoples' rights to govern the collection, access, use, and reuse of data about their communities, lands, and cultures. Using RPAS responsibly in these contexts means recognising the cultural significance of land and data, and working collaboratively in accordance with ethical standards.
We invite you to learn more about the CARE principles by following this list of resources:

- Global Indigenous Data Alliance (GIDA): https://www.gida-global.org/care

- Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., et al. (2020). The CARE Principles for Indigenous Data Governance. Data Science Journal, 19(43), 1–12. https://doi.org/10.5334/dsj-2020-043

Use case

Specificities of data acquisition on Traditional Owners' sites *By Jens Klump* *CSIRO*
In most cases, fieldwork is conducted on someone's land and requires appropriate access arrangements. Beyond obtaining the land owner's permission, it can be advisable to inform local communities about the planned work. In some countries, e.g. Australia, land rights and traditional ownership by Indigenous communities intersect. For example, sites of cultural significance might be located in a national park. Besides obtaining permission from the relevant national park authority, Traditional Owners also need to be consulted to ensure that the physical integrity of the sites is preserved and data and knowledge about the sites are shared in a culturally appropriate way. This could mean that access to the data is restricted to persons with cultural permission. In Australia, Traditional Owners can be approached through the Prescribed Body Corporate of the relevant site or Native Title Area.

2. Data Acquisition

Data acquisition is a key step that will directly impact the quality, consistency, and future (re-)usability of the data you collect. It is thus extremely important to record as much information as possible regarding your RPAS operations while maintaining safe operations in the field.

Recommendations:

Data and metadata

- Use the data collection checklist to record all the required information for future processing
- Calibrate your instruments following best practices

Privacy, legal and ethics considerations

- Adhere to relevant national civil aviation rules, institutional procedures and access restrictions to undertake safe operations
- Build trust with local communities through consistent ethical behaviour
- Follow best practices regarding environmental and wildlife protection.

Resources and additional information

- ➢ Recording information related to your RPAS survey is likely most accurately recorded in real time, while you are actively working in the field. Required information includes, for example, weather conditions, sensor settings, site details and any issues you have encountered during the flight. To best assist you in documenting your mission, we have developed a data collection checklist to help you systematically record all required information. You are invited to complete the logsheet during each flight to support future analysis and archiving.

- The logsheet and RPAS data collection checklist is available in the appendix or downloadable here: Fremand, A., Klump, J., Manthorpe, S., Whitelaw, M., Gerard, F., Garland, W., Semong, T., George, C., & Ricketts, H. (2026). RDA Recommendations on Remotely Piloted Aerial System data management: Log sheet and acquisition checklist. Zenodo. https://doi.org/10.5281/zenodo.18999774

While some information can be automatically extracted from the data files, we recommend that you also record wider information that is not likely to be retained otherwise, to aid data discovery and usability such as the purpose of the survey - as this can affect the interpretation of the measurements, any atmospheric or geophysical features or phenomena encountered, issues, data quality, why there are no data collected or other environmental context information.

➢ To maintain the accuracy of your sensors, calibration is a crucial step to correct for sensor drift, environmental influences, and platform-specific biases. The exact procedure might differ depending on the sensor; it is thus advised to refer to the manufacturer's recommendations. The following documents give additional guidance regarding calibration best practices:

- Guidelines for UAV imagery calibration by USGS: https://www.usgs.gov/publications/guidelines-calibration-uncrewed-aircraft-systems-imagery
- Tmušić, G.; Manfreda, S.; Aasen, H.; James, M.R.; Gonçalves, G.; Ben-Dor, E.; Brook, A.; Polinova, M.; Arranz, J.J.; Mészáros, J.; Zhuang, R.; Johansen, K.; Malbeteau, Y.; de Lima, I.P.; Davids, C.; Herban, S.; McCabe, M.F. Current Practices in UAS-based Environmental Monitoring. *Remote Sens.* **2020**, *12*, 1001. https://doi.org/10.3390/rs12061001
- Cramer, M., Przybilla, H.-J., and Zurhorst, A.: UAV cameras: Overview and geometric calibration benchmark, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W6, 85–92, https://doi.org/10.5194/isprs-archives-XLII-2-W6-85-2017, 2017.
- Daniels, L., Eeckhout, E., Wieme, J., Dejaegher, Y., Audenaert, K., & Maes, W. H. (2023). Identifying the optimal radiometric calibration method for UAV-based multispectral imaging. *Remote Sensing*, *15*(11), 2909. https://doi.org/10.3390/rs15112909
- Aragon, B., Johansen, K., Parkes, S., Malbeteau, Y., Al-Mashharawi, S., Al-Amoudi, T., ... & McCabe, M. F. (2020). A calibration procedure for field and UAV-based uncooled thermal infrared instruments. *Sensors*, *20*(11), 3316. https://doi.org/10.3390/s20113316

➢ When undertaking your mission, it is essential to comply with the civil aviation regulations of the country in which they are conducted. Additional resources are provided in the previous section.

In addition to national and international aviation rules, operators must follow the SOP developed prior to the mission and any other specific institutional procedures. For example, institutions often recommend having at least two trained personnels present during drone operations: one acting as the pilot and the other as an observer. This configuration will help enhance situational awareness and allow for immediate response to technical or environmental issues. The following article gives insights into successful operation procedures used in the UK:

  - Andrew M. Cunliffe, Karen Anderson, Leon DeBell & James P. Duffy (2017). A UK Civil Aviation Authority (CAA)-approved operations manual for safe deployment of lightweight drones in research, International Journal of Remote Sensing, 38:8-10, 2737-2744, https://doi.org/10.1080/01431161.2017.1286059

- Building trust with local communities will ensure you can undertake your mission activities in the best conditions. This is done by undertaking safe and transparent RPAS surveys in alignment with ethical rules. As part of your survey, it is important to keep the local community updated with your progress and respond to any concerns promptly. The mission should follow the plan communicated with locals and all documentation regarding permits and consents available on request.

- Flying a RPAS might disturb the local wildlife; it is thus important to follow best practices and locally applicable laws to avoid disturbance. Best practices for RPAS use around wildlife are detailed in the following documents:

  - Edney, A.J.; Hart, T.; Jessopp, M.J.; Banks, A.; Clarke, L.E.; Cugniere, L.; Elliot, K.H.; Martinez, I.J.; Kilcoyne, A.; Murphy, M.; Nager, R.G.; Ratcliffe, N.; Thompson, D.L.; Ward, R.M.; Wood, M.J.. 2023 Best practices for using drones in seabird monitoring and research. Marine Ornithology, 51 (2). 265-280. https://doi.org/10.5038/2074-1235.51.2.1544

  - Harris, C. M., Herata, H., & Hertel, F. (2019). Environmental guidelines for operation of Remotely Piloted Aircraft Systems (RPAS): experience from Antarctica. *Biological Conservation*, *236*, 521-531. https://doi.org/10.1016/j.biocon.2019.05.019
  - Marine life:

    - https://www.fisheries.noaa.gov/insight/viewing-marine-life

    - https://www.fws.gov/refuge/oregon-islands/tips-responsible-drone-use

  - Hodgson, J. C., & Koh, L. P. (2016). Best practice for minimising unmanned aerial vehicle disturbance to wildlife in biological field research. Current Biology, 26(10), R404-R405. https://doi.org/10.1016/j.cub.2016.04.001

  - Bierlich KC, Hewitt J, Bird CN, Johnston DW, Dale J, Pirotta E, Schick RS, Stewart JD, New L, Chimienti E, Goldbogen JA, Friedlaender AS, Cantor M, Torres LG. 2025. A workflow of open-source tools for drone-based photogrammetry of marine megafauna. *PeerJ* 13:e19768 https://doi.org/10.7717/peerj.19768

Use cases

RPAS surveying of marine mammals and seabirds in Ireland

By Sarah Manthorpe
University College Cork and MaREI - SFI Research Centre for Energy, Climate and Marine

RPAS use has expanded at University College Cork to survey seabirds and marine mammals more widely in recent years. Surveys are carried out for detection, population censusing and monitoring health via body condition. Data is acquired on both land and from vessels, and is often around sensitive wildlife populations in harsh weather conditions. Thus, data acquisition involves a pilot and observer, where both have trained together prior to fieldwork on deployment, recovery and wildlife disturbance and response.

Fieldsheets record information on general flight conditions, as well as additional environmental parameters such as wind speed, cloud cover, wave height and swell. Further, any disturbance observed is recorded alongside the pilot's response with the RPAS. The pilot and observer are to adhere to local aviation authority legislation, which includes maximum flight heights and restricted areas, while also managing university procedure, which will soon include parameters on signage and Irish Aviation Authority (IAA) certifications present at the field site.

Flights in these conditions are often manual, but in some cases are pre-programmed. Survey height is a critical variable in the effect of disturbance and the subsequent quality of data (ground sampling distance), and thus, great care is taken in following guidelines and height restrictions based on the surveyed species. Additionally, surveys may also record in both RGB and thermal, and so the pilot has to ensure the RPAS is collecting imagery that is comparable in these spectra by configuring the drone during acquisition to record at a similar zoom. Lastly, data may be in either photographic or video format, depending on the use case, thus requiring manual flying where an observer may be searching for the survey species (for example, puffins, whales), while also monitoring potential disturbance.

Data acquisition in Botswana - RPAS Power line inspection

By Thabo Semong
Botswana International University of Science and Technology (BIUST), Computing and Informatics Department, BIUST Drones

RPAS has emerged as a transformative tool for data acquisition in various fields. For the Botswana International University of Science and Technology (BIUST), the strategic adoption of RPAS significantly enhanced collaborations and community engagement. To harness this potential effectively, a systematic approach to data acquisition is essential. Pre-Flight Planning & Preparation is essential for all the missions. This is the most crucial phase, where an operation's success or failure is determined.

The Civil Aviation Authority of Botswana (CAAB) is mandated with securing the Botswana sky. They have a step-by-step framework to follow to execute successful and legally compliant data acquisition missions across Botswana. It is very critical to adhere to this protocol to ensure data integrity, regulatory compliance and safety with the CAAB, and the maximization of research and commercial outputs. Every designated pilot-in-command must hold a valid remote pilot license (RPL) from CAAB. For operations near airports, military zones, or protected areas, additional clearances may be required from relevant authorities.

Flight plans must be shared with the CAAB, especially for Beyond Visual Line of Sight (BVLOS) operations. In case of powerline inspections, public notification must be done since the drone will be flying in public areas. A thorough risk assessment and mitigation strategies before any mission is also important.

It is also important to choose the appropriate drone and select the sensor based on the objective of the mission. Botswana's conditions can change rapidly, hence monitoring of weather forecasts is critical in all missions, in order to avoid flying in high winds, rain, or extreme heat.

The Pilot-in-Command (PIC) is the final authority for the flight. The PIC takes off and initiates the pre-planned autonomous mission. The PIC monitors the drone's telemetry (battery, GPS signal, altitude, link strength), this is even more critical in cases where live feed is provided. Periodically check that the camera is triggering and capturing images or videos as expected.

By rigorously adhering to this structured workflow, the BIUST Drones Team position itself as a leader in safe, ethical, and scientifically rigorous data acquisition in Botswana. This protocol will not only enhance the quality of research and projects undertaken by BIUST but also ensure the long-term sustainability and positive reputation of the university's drone operations.

3. Data Processing Pipeline

The data processing pipeline documents all the steps to transform raw RPAS data into final research outputs. As part of the processing, it is essential to record all lineage information while protecting sensitive information.

Recommendations:

Data and metadata

- Document the processing workflow following the data processing checklist
- Separate raw, pre-processed and processed data and use version control
- Provide a quality statement giving insights about quality assessment and control (QA/QC) applied to your data

Privacy, legal and ethics considerations

- Blur or anonymise personal data if applicable
- Reduce the precision relative to sensitive location (e.g. of endangered species) if applicable

Resources and additional information

- RPAS data can be described under different levels of processing:
    - Raw data: also referred to as level 0 or unprocessed outputs directly exported from the instruments. This should also include any logs and navigation information.
    - Pre-processed data: also referred to as level 1 data. At this stage, calibration routines, georeferencing, and initial quality assurance and control (QA/QC) are applied to ensure data integrity and consistency.
    - Processed data and derived products: also referred to as level 2 and beyond. At this stage, the level 1 data have been corrected to create quantitative or thematic products (e.g., orthomosaics, digital surface models, or classified maps). Depending on the complexity and analysis of the outputs, the level can be adjusted from Level 2 to higher levels of processing.

While you process the data, it is best practice to make sure your raw data is stored as read-only to avoid unintentional changes to the data. Version control can be used to track the different steps and iterations a data set has undertaken as part of the data processing pipelines. It should also be used to track changes to the associated software or script used as part of the processing routine. These steps are key to avoiding data loss and helping you document the provenance of the data.

In addition, each level of RPAS data output must be accompanied by its respective metadata. Level 0 metadata should include the information recorded in the field as part of the data acquisition checklist. For Level 1 outputs, metadata must document all pre-processing steps. This should include the calibration results and parameters, quality assessment checks, and any other corrections applied, like GNSS processing information. Level 2 and higher-level products must document the lineage information, including reference to Level 0 and Level 1 metadata and any software or code with their associated version, the spatial resolution, coordinate reference systems, thematic classifications or derived variables, and reference data sources used for algorithm training and validation.

The list of required metadata are summarised in this RPAS data processing checklist:

Level 0: raw data

- Mission identifiers
- Sensor specifications
- Platform details
- Flight parameters (e.g., altitude, speed, GPS logs)
- Specific details regarding potential issues recorded on the field.
- Type of data, output format and description of variables

Level 1: pre-processed data, calibrated data

- Level 0 data information

- Calibration parameters and methodology
- QA/QC flags
- Georeferencing methods
- Corrections applied
- Pre-processing software and version
- Type of data, output format and description of derived variables

Level 2 and onwards: processed, derived data

- Level 0 and Level 1 data information
- Processing methodology and algorithms used
- Software versions
- Reference data used for algorithm training and/or validation
- Spatial resolution and coordinate reference systems
- Type of data, output format and description of derived variables

It is to be noted that the name of the levels might differ across disciplines but most workflows will consist of a raw/pre-processed/processed stages. Using descriptive terms instead of numbers to label levels is recommended to avoid misinterpretation (see NASA Planetary Data System archive example below).

➢ Calibration and processing of RPAS-mounted sensors is a critical step to ensure data accuracy and reliability across diverse scientific applications. Please refer to the previous section for additional guidance.

➢ To protect the privacy of individuals who might have been captured in RPAS imagery, it is essential to anonymise the data by blurring faces or other identifiable features. These practices help ensure compliance with data protection regulations like GDPR and ensure compliance with ethical standards in research. The following resource provides a guide to privacy and data protection risks that may arise when you operate RPAS:

  ○ European Aviation Safety Agency, EASA *Privacy Handbook for Drone Users,* https://www.easa.europa.eu/en/domains/civil-drones/privacy/privacy-handbook [last accessed 04/09/2025]

➢ To protect sensitive ecological information, particularly the locations of endangered species, it is advisable to reduce the spatial precision of RPAS data when applicable. Instead, generalising location data helps safeguard these areas, protect the wildlife from disturbance or exploitation, while still supporting scientific and conservation objectives. For example, reporting the presence of a specific species within a 10 km radius might be enough for a publication. For detailed guidance and best practices for generalising sensitive locations, refer to the following GBIF publication:

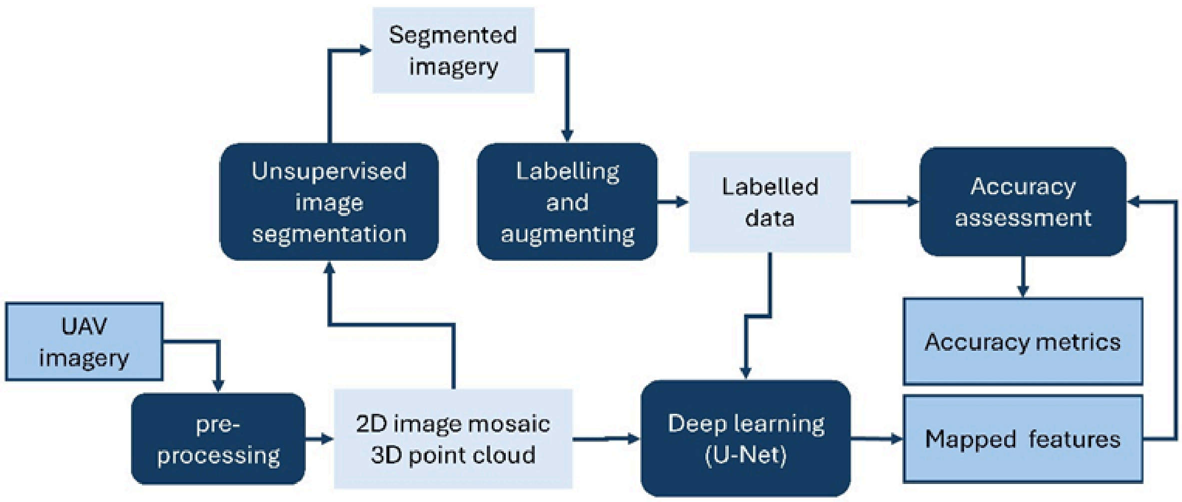
- Chapman AD (2020) Current Best Practices for Generalizing Sensitive Species Occurrence Data. Copenhagen: GBIF Secretariat. https://doi.org/10.15468/doc-5jp4-5g10.

Use case

Data processing pipeline at UKCEH

By France Gerard and Charles George at the UK Centre for Ecology and Hydrology (UK CEH)

The use case represents a workflow developed to identify and map individual Hawthorn shrubs in a rewilding site (Strawberry Hill, Bedfordshire, UK). A future operational version of this pilot would enable RPAS-based monitoring of woody (shrub) cover and biomass in restoration sites. The workflow is a typical example of RPAS 2D-image and 3D-point cloud use for object (shrub) mapping (detection and recognition) or semantic image/pointcloud segmentation using machine or deep learning.



The schematic above describes a typical workflow specifically designed to extract features/objects (e.g., individual shrubs) from a 2D RPAS image mosaic or 3D RPAS point cloud collected for a site. A similar workflow could be applied to identify features/objects from a 3D LiDAR point cloud, or when carrying out a semantic classification of the 2D mosaic or 3D point cloud. The main workflow steps are:

1. Pre-processing of individual image frame raw data, using Structure from Motion (SfM) software, into a geo-registered 2D image mosaic and a 3D point cloud.
2. Unsupervised 2D image segmentation (or 3D point cloud segmentation) to support the feature/object tracing and labelling, or alternatively, manual tracing and labelling to deliver a large set of 2D (or 3D) labelled data for training and validation.
3. Training of the classifier/model using the training data .

4. Applying the classifier/model (e.g., a U-Net) to the 2D image mosaic (or 3D point cloud) to produce the output map of features/objects.
5. Evaluating model performance using the independent validation data.

The key metadata to be recorded during the processing chain to ensure data re-use are (with the numbers reflecting the main workflow steps listed above and shown in schematic):

1. Pre-processed output files: include the complete (verbose) source data metadata; list the pre-processing software used, version and user-guided adjustments to parameters; output data extent or bounding box coordinates; location; projection; resolution (2D: pixel size; 3D:point density); number and wavelength of bands; unit and bitrate of image (point) values; file format.

2. Segmentation algorithm and version (software citation) used to support labelling plus list suggested below related to deep learning models.

3. Labelled datasets as output: class/object name, date of creation/version, link to repository, citation if available.

4. Classifier/model type/variant (software citation) and version plus list suggested below;

5. Labelled dataset source and version used for validation (link to repository); Independent validation metrics scores.

For the models used (AI/deep learning and others), based on a conversation captured on reddit (What metadata do you track when training models? : r/mlops) and Chatgbt suggestions, recommended essential Model Metadata are:

- Model Architecture & Weights: The core components of the model that define its structure and learned parameters.
- Data Details: Information about the datasets used for training, including their version and location.
- Training Hyperparameters: Settings used during training, such as learning rate, batch size, and optimizer choice (include list of hyper parameters that are not set).
- Performance Metrics: Recorded metrics from training, validation, and testing phases:cross validation metrics scores, testing metrics scores.
- Model Version & Lineage: A unique identifier for the model and its connection to previous versions and experiments.
- Environment & Dependencies: Details about the software libraries (e.g., TensorFlow, PyTorch), hardware (CPU/GPU models), and operating system used.
- Resource Utilization: model training and testing resources, run time resources used for training (CPU, Memory, GPU).
- User-Defined Tags: Customizable tags or descriptions to categorize the model's purpose, use case, and other relevant attributes.

The main challenges we currently face are related to (lack off) transparency and (lack off) interoperability of RPAS data pre-processing pipelines – Challenges (not all metadata challenges) related to processing RPAS image and liDAR data are listed here https://doi.org/10.1016/j.ecolind.2024.112970

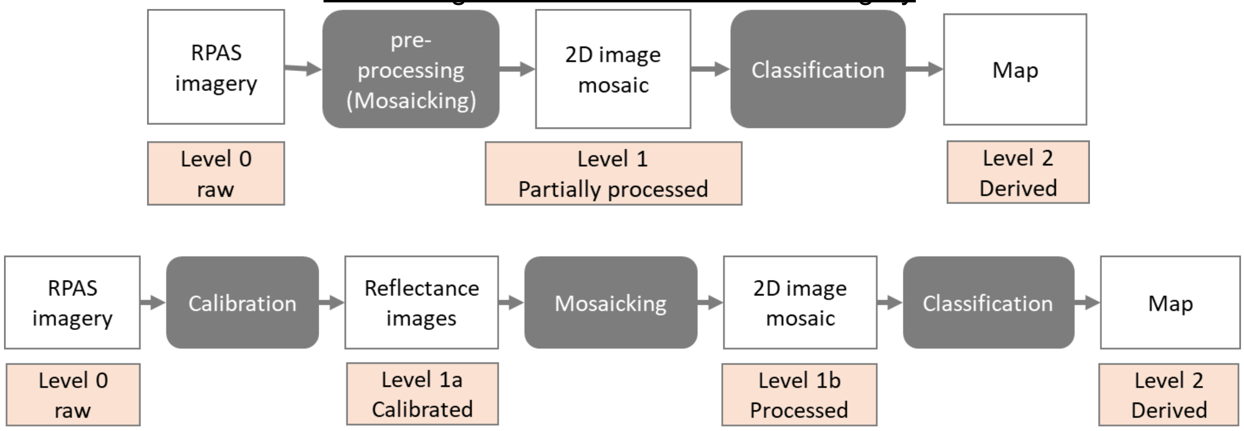
Following the example of the NASA Planetary Data System archive
By Jens Klump and France Gerard

The Geosciences Node of NASA's **Planetary Data System** (PDS) *archives and distributes digital data related to the study of the surfaces and interiors of terrestrial planetary bodies*, including Earth. Data handled in this archive is similar to data generated from RPAS campaigns. Archived data is labelled according to five data processing levels (https://pds.nasa.gov/datastandards/about/). Originally labelled as numbers (PDS3) users would use different sets of numbers, causing confusion. To avoid this confusion a revision (PDS4) replaced numbers with descriptive terms (see table below). We recommend a similar approach.

NASA's Planetary Data System data processing levels	
Processing Level	**Definition**
Telemetry	An encoded byte stream used to transfer data from one or more instruments to temporary storage where the raw instrument data will be extracted. **PDS does not archive telemetry data.**
Raw	Original data from an instrument. If compression, reformatting, packetization, or other translation has been applied to facilitate data transmission or storage, those processes will be reversed so that the archived data are in a PDS-approved archive format.
Partially Processed	Data that have been processed beyond the raw stage but which have not yet reached calibrated status.
Calibrated	Data converted to physical units, which makes values independent of the instrument.
Derived	Results that have been distilled from one or more calibrated data products (for example, maps, gravity or magnetic fields, or ring particle size distributions). Supplementary data used to interpret observational data, such as calibration tables or tables of viewing geometry, should also be classified as derived data if not easily matched to one of the other categories.

Below are example workflows for processing RPAS collected imagery into a classified map with corresponding processing levels labelled using numbers and descriptive terms similar to those in the Planetary Data System. It demonstrates how numbered levels could cause confusion and how descriptive terms are more intuitive and user friendly.

Processing levels for RPAS collected imagery



Processing levels for RPAS collected Lidar data

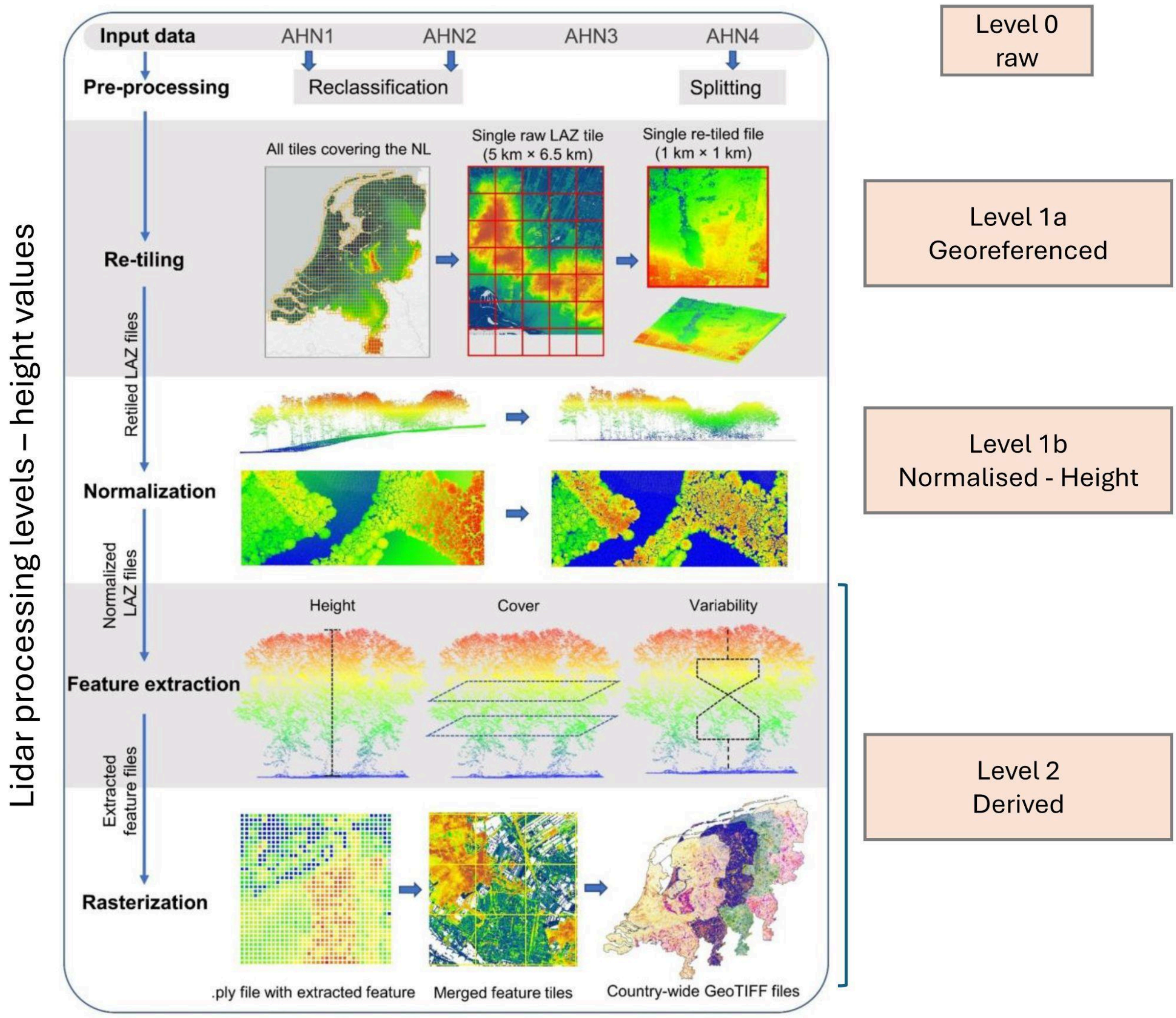
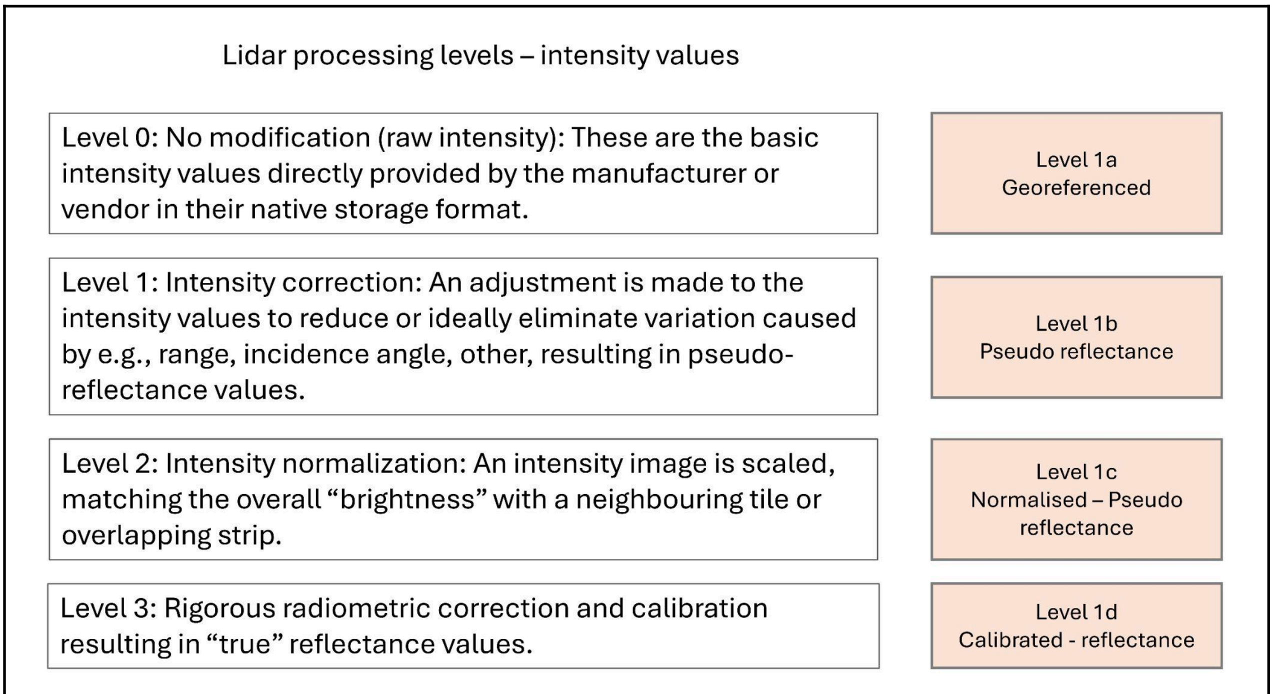


Image source: https://doi.org/10.3897/arphapreprints.e190520



Based on : Kashani AG, et al., A Review of LIDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. Sensors (Basel). 2015 Nov 6;15(11):28099-128. doi: 10.3390/s151128099.

4. Data Storage and Preservation

Data storage and preservation are important steps in ensuring that data is fit for the future. At this stage, it is essential to keep the data in an open format with all the required metadata to enhance reusability while providing secure storage to protect sensitive information.

Recommendations:

Data and metadata

- Use recognised metadata standards used in your disciplines and follow the recommendations for specific RPAS metadata
- Use open formats, where available.
- Adopt ontologies to enhance semantic clarity and interoperability.
- Use controlled vocabularies to describe Drone and related artefacts
- Utilise Persistent Identifiers for Instruments (PIDINST) for accurate referencing and traceability.

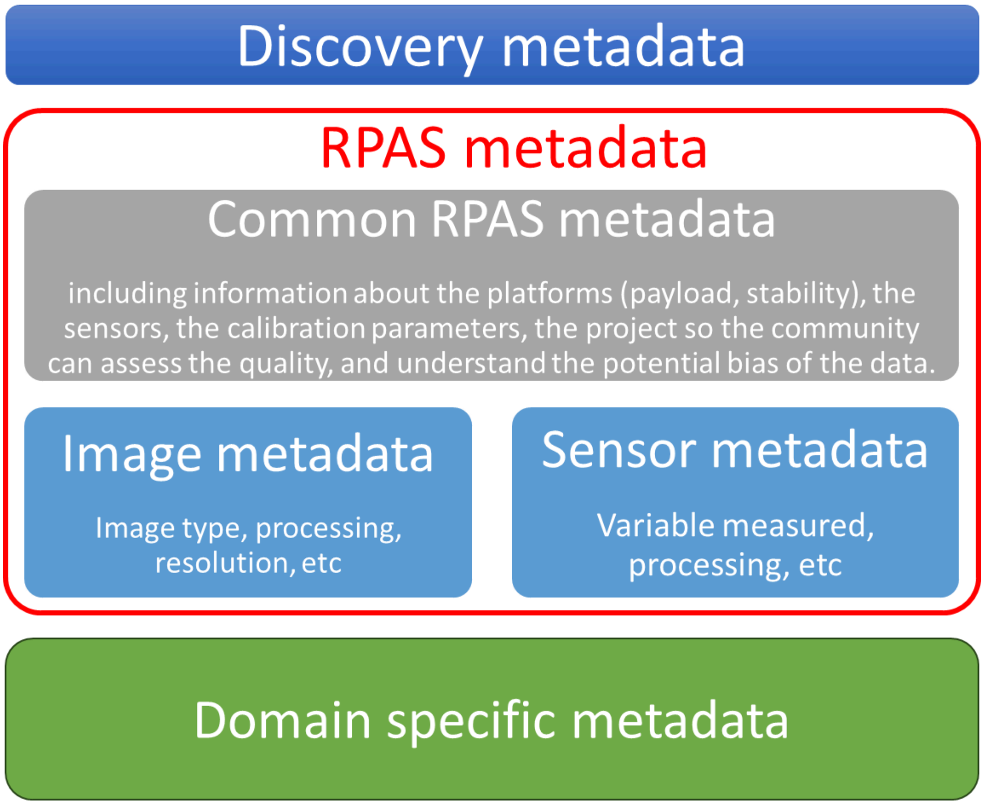
Privacy, legal and ethics considerations

- Establish secure storage with access restrictions in alignment with requirements

Resources and additional information

➢ Metadata are data that provide information about a specific dataset. They help describe the content, context, and give information on how to use the data, making it easier to find, access and reuse. In the context of RPAS, we can separate metadata into three main categories:

- **Discovery metadata** - These metadata refers to the metadata that enables users to find and identify a dataset. The core element of discovery metadata includes a title, the name of the authors, a date of creation and publication, an abstract, identifier (e.g. DOI) and the link to a license. Common discovery metadata schemas are Dublin Core and DataCite.
- **RPAS metadata** - These metadata are specific to an RPAS survey and include information about the platforms (payload, stability), the sensors, the calibration parameters, and the project, so the community can assess the quality and understand the potential bias of the data. The Minimum Information Framework developed by Barberi et al. (2023) as part of the RDA UAV Interest group provides a specific framework for the description of UAV metadata:
    - Barbieri, L., Wyngaard, J., Swanz, S., & Thomer, A. K. (2023). Making drone data FAIR through a community-developed information framework. Data Science Journal, 22, 1-1.
- **Domain-specific metadata** - These metadata have been developed to respond to the needs, terminology, and data types of a particular academic discipline. They will recommend the use of specific ontologies or require specific fields tailored to that discipline. In Earth and Environmental Sciences, a common standard is the ISO19115 standard and its associated XML implementation (ISO19139 standard). For atmospheric sciences the use of the CF metadata conventions https://cfconventions.org/ are widely used.
- **Controlled vocabularies** - these refers to a pre-defined list of standard list of terms to represent RPAS and related artefacts. For example, control vocabularies can be used to describe each drone platform and sensors mounted on platform for better annotation, SKOS model is widely used to represent RDF vocabulary.



The three types of metadata required to describe data derived from RPAS. After Fremand et al. (2023)

➢ When preparing data for long-term preservation, it's important to think about how it will be accessed and used in the future. One key distinction to make is between distribution and archival formats. Distribution formats are optimised for accessibility, ease of use, and compatibility with common tools making them ideal for sharing data with collaborators or the public. Archival formats, on the other hand, prioritise long-term preservation, ensuring that data remains readable and usable even as technologies evolve.

For long-term data archiving, it is best to use open data formats, as they are more likely to remain supported in the future and will thus help prevent the risk of data becoming obsolete. Open formats are publicly documented formats that are free to use and usually well supported by the community. Open formats can easily be read by a wide range of tools and platforms, promoting reuse. To promote long-term reuse of the dataset, it is also advised to use formats that support embedded metadata, such as netCDF. Indeed, it allows users to document the content of the file directly. In contrast, proprietary formats often require specific commercial software to read the data. If the associated company were to go bankrupt, those tools might not be easily available in the future.

Recommended formats include:

- **NetCDF, HDF5:** These formats are recommended for multidimensional environmental data. They store data in arrays and allow embedded metadata.

- **CSV:** Recommended for structured tabular data such as flight logs or sensor readings. Consider adding metadata either separately or within the file itself. The XCSV format can be considered as an example of an embedded-metadata CSV format: https://github.com/paul-breen/xcsv
- **ASCII/TXT**: Suitable for some sensor data. These formats can be used, for example, for output from proprietary systems.
- **GeoTIFF**: Suitable for georeferenced imagery and elevation models.
- **Shapefiles/GeoPackages:** Recommended for vector data such as flight paths or annotated features. KML and GeoJSON are also open formats that can be considered for sharing and visualization of vector data, although they may be less robust for complex datasets or long-term archival.
- **LAS/LAZ:** Suitable for LiDAR point cloud data.
- **Cloud Optimised Point Cloud (COPC)**: cloud-native format to represent laz file in a zip format.
- **JPEG, PNG, TIFF**: Suitable for images. These formats allow the addition of embedded metadata. The NCAS Image Metadata Standard provides a framework for describing images using embedded metadata:
  Hooper, D. A., Brooks, B. J., Garland, W. E., & Parton, G. A. (2022). The National Centre for Atmospheric Science Image Metadata Standard (NCAS-IMAGE) version 1.0 (1.0). Zenodo.
  https://doi.org/10.5281/zenodo.6368295
- **FFV1, MKV, MP4**: Suitable for video outputs. Videos can be very large in size, and still images may be easier to handle in practice. FFV1 and MKV are recommended for archival use. MP4 uses lossy compression, making it less suitable for preservation but useful for distribution.
- **ZARR**: A cloud-optimized format that is popular for large RPAS datasets. It is ideal for sharing and processing but not recommended for archiving.
- **Cloud Optimised Geotiff (CoG)**: Suitable to publish stitched images collected by drone.

➢ Ontologies can be defined as a set of concepts and categories in a subject area or domain that show their properties and the relations between them (Oxford dictionary). In the context of RPAS data management, ontologies can help you standardise the way you describe the relationship between the sensors, the platform used, and the output values obtained. It can also standardise the way you describe the context and subject of your study, improving interoperability between different datasets. The following ontologies can be used to describe RPAS-derived datasets:

  - The SOSA ontology (Sensor, Observation, Sample, and Actuator) provides a lightweight core for describing sensors, their observations, samples, and actuators. https://i-adopt.github.io/ontology/

- The I-ADOPT ontology framework facilitates interoperability between variable description models by providing standardised, machine-interpretable descriptions. https://www.w3.org/TR/vocab-ssn/
- The LANDRS ontology (Linked-Data API for Networked DRoneS) has been specifically designed in the context of small Unmanned Aircraft Systems research. It aligns with best practice ontologies like SOSA and NASA's SWEET. https://www.landrs.org/
- In Earth System, the EarthPortal also provides a list of existing ontologies: https://earthportal.eu/. The OntoPortal Alliance provides additional ontologies as well: https://ontoportal.org/

➢ In addition to using ontologies, you can also use PIDINST, the globally unique identifiers for instruments. PIDINST metadata (https://docs.pidinst.org/en/latest/) describes and gives a reference to the exact instrument used during a survey. By linking datasets to specific instruments via persistent identifiers, researchers can ensure transparency, support metadata enrichment, and facilitate long-term data stewardship across projects and institutions. PIDINST is still new, but you can read information from early adopters from:

- https://metadatagamechangers.com/blog/2024/9/12/instrumentsdatacite

➢ Where RPAS-derived data includes sensitive information, establishing secure storage with appropriate access restrictions is essential to ensure data integrity, confidentiality, and compliance with ethical and legal standards. Access controls should reflect the sensitivity of the data and be aligned with institutional data governance policies, ensuring that only authorised researchers and collaborators can access or modify datasets.

Use case

The use of the SOSA ontology in the context of RPAS data (TERN)

By Siddeswara Guru
The University of Queensland

The Sensor, Observation, Sample, and Actuator (SOSA) ontology is a lightweight, general-purpose framework designed to model the interactions between sensors, observations, sampling activities, and actuators. It was developed jointly by the W3C and OGC to simplify and unify how sensor-related data is represented on the web. SOSA is a streamlined version of the earlier SSN (Semantic Sensor Network) ontology, focusing on usability and modularity. It supports integration with other standards like PROV-O, QUDT, and Darwin Core, making it highly interoperable.

In the context of RPAS data, SOSA provides a semantic structure for describing drone, related artefacts, samples and observations related to samples. This includes:

- RPAS described as a platform
- sampler mounted on RPAS platforms (sensors mounted on RPAS are modelled as a

sosa:sampler because they are sampling a feature (taking images).

- Deployment information and calibration related information
- Observations made during flight (e.g., geochemical sampling, vegetation surveys)
- Sampling procedures (e.g., soil or air sample collection)
- Actuators, if the RPAS performs actions like releasing payloads or triggering instruments

SOSA is also foundational to the TERN Ontology, which supports plot-based ecological data integration. This ontology extends SOSA to describe site visits, ecological attributes, and survey metadata, and is implemented in platforms like EcoPlots.

RPAS missions generate diverse datasets—multispectral imagery, LiDAR, thermal data, and telemetry—that must be harmonised with ground-based ecological observations. Without a shared semantic framework, integrating these datasets across platforms and institutions is difficult.

TERN extends the SOSA ontology to create the TERN Ontology, which semantically describes RPAS-derived observations alongside traditional ecological data:

- RPAS Platform: Modelled as sosa:Platform, representing the drone and its configuration. Different platforms are represented as controlled list and represented as platform type property. For example UAV, satellite remote sensing, ecological site, flux tower are platform type.
- Sampler: Instruments onboard (e.g., multispectral camera, LiDAR) are represented as SOSA:Sampler because the device is used to sample the feature
- Observations: Data derived from the sample images (e.g., NDVI, canopy height) are represent as sosa:Result with observable property (canopy height) phenomenon time using a procedure (methods used to derive result).
- Sampling: Ground truthing or soil sampling activity triggered by RPAS is represented using sosa:Sampling performed by a sosa:Sampler where the result is a sosa:sample. The sample can be a FeatureOfInterest for further observation or sampling.
- Site Visits: TERN Ontology adds tern:Site and tern:SiteVisit to describe contextual metadata that may be related to a campaign that happens in a particular area of interest (i.e., site) and artefacts related to site visit is captured in tern:SiteVisit

The model is deployed in EcoPlots, TERN's ecological data integration platform, enabling harmonisation of RPAS and ground-based data. Controlled vocabularies are published as linked data with persistent URIs, improving interoperability and reuse.

5. Data Sharing and Reuse

To foster reuse of the RPAS-derived data, it is advised that the data be published through a trusted repository. To protect sensitive information, it is essential to define a license and access in accordance with the requirements of privacy.

Recommendations:

Data and metadata

- Share your data through a trusted data repository using open formats and all the necessary metadata required for reuse

Privacy, legal and ethics considerations

- Adhere to privacy laws and ethical guidelines, particularly regarding sensitive data by:
  - Following the principle ‘as open as possible, as closed as necessary’
  - Defining a license in line with the requirement of privacy

Resources and additional information

➢ A trusted data repository is a facility that ensures long-term preservation of research data. In some cases, research repositories will be certified (e.g. with CoreTrustSeal) to promote their good data management practices and trustworthiness. As part of the deposit process, a data manager will check the data and ensure they are published with the required metadata to foster reuse. A persistent identifier, such as a Digital Object Identifier (DOI), will be minted, allowing easy citation of your research output. For optimal data quality and usability, researchers are encouraged to deposit their datasets in discipline-specific repositories. These repositories offer domain-relevant expertise and curation, ensuring quality, discoverability and reuse of the data by others in the field.

  - The Registry of Research Data Repositories https://re3data.org lists and describes research data repositories and can be sorted by disciplines and filtered by attributes, e.g., certification status.
  - Some repositories also follow the TRUST principles. Standing for Trust, Responsibility, User-focus, Sustainability and Technology, the TRUST principles have been developed by the RDA TRUST Interest group, which has created a list of TRUSTed repositories: https://archive.rd-alliance.org/rda-community-effort-trust-principles-digital-repositories
  - To allow long-term preservation, an open format and extensive metadata to foster reuse are advised. Section 4 provides additional resources and recommendations on formats.

➢ To support open science while respecting ethical and security constraints, it is best practice to follow the principle of being “as open as possible, as closed as necessary”. As a general rule, all data should be open, meaning that the data are freely accessible online and ready to be reused by others. To clarify how the data can be used, you should apply a data license supporting reusability. A data license is a legal arrangement between the data creator and the user. Without one, data may be considered copyrighted in some jurisdictions, restricting further reuse of the data. To foster openness, it is thus essential to choose a license allowing free reuse of the data. Open licenses include open government licences or Creative Commons (CC) Licenses. A recommended licence is CC-BY 4.0 (Attribution) licence, which allows users to use the data for any purpose as long as the source of the data is properly attributed.

  - To learn more about Creative Commons Licenses: https://creativecommons.org/

Some RPAS-derived data may contain some sensitive information related to privacy or the location of endangered species. In those cases, section 3 provides additional resources on how best to protect the privacy of individuals or sensitive location information. Guidelines include recommendations around blurring faces or other identifiable features or generalising location information. In specific cases where the data cannot be open to the public even with additional processing, you must establish secure storage systems with restrictive access controls. A special data use agreement might be necessary to share your data. Examples for sensitive RPAS data are identifiable persons where consent had not been obtained, or other privacy or commercial considerations. Where CARE principles are concerned, early discussion with stakeholders and communities is advised. To learn more about open science, please refer to:

- The European Research Executive Agency website, which provides a definition of open science in the European context: https://rea.ec.europa.eu/open-science_en

Use case

Publication of RPAS data at the UK Polar Data Centre *By Mari Whitelaw and Alice Frémand* *UK Polar Data Centre, British Antarctic Survey*
At the UK Polar Data Centre, we manage data from various disciplines reflecting the multidisciplinarity of polar science. The way we manage and publish data is often different depending on the discipline. We discuss here the publication of two very different RPAS-derived datasets. **Gravity data collected in West Antarctica** In this example, we present how we published gravity data collected over Marguerite Bay in Antarctica using a Windracers Ultra UAV as part of their “SWARM survey”. The data were collected by the aerogeophysics team of the British Antarctic Survey (BAS) in February 2024. The team has been collecting airborne data for decades, mainly using aircraft. This was the first gravity survey using a large RPAS to collect gravity data at BAS. Strong from their experience of using airborne platforms, it was a smooth data management shift for the scientific team. Most of the challenges were technical, but as similar sensors were used, the output data were almost identical to what they used to collect using aircraft. For publication, the same processes were undertaken as the one used from the other airborne gravity dataset. The protocol was previously discussed in this data paper. The RPAS-derived data consists of tabular data published in CSV format. To refer to the metadata, the citation has been added to the CSV file directly. As the data were similar in form to their previous gravity datasets, it was possible to use the same routines to integrate them to our geophysics portal: the Polar Airborne Geophysics Data Portal. The data can be accessed via: Jordan, T., Robinson, C., Reed, T., & Toomey, R. (2024). *Airborne gravity data over Marguerite Bay collected with a Windracers Ultra UAV (2023/24 season)* (Version 1.0) [Data set]. NERC EDS UK Polar Data Centre. https://doi.org/10.5285/3a9c8604-2bca-48c1-a40c-48a873076581

Fur seal colony survey data collected in South Georgia

This example highlights a relatively new use case for UK PDC, albeit one that is expected to become much more prevalent. Members of the BAS Mapping and Geographic Information Centre (MAGIC) team collected aerial imagery from multiple wildlife colonies on the Sub-Antarctic island of South Georgia. One survey season produced multiple datasets, of which 29 have so far been published. The aerial surveys were conducted as part of the Darwin Plus (DPLUS) 109 project: Initiating Monitoring Support for the SGSSI-MPA Research and Monitoring Plan for the purpose of establishing a baseline population estimate of key indicator species around South Georgia.

These datasets presented numerous challenges:

- Each flight produced numerous outputs which involved different levels of processing. Each dataset contains up to three levels; raw data (gnss files, log files and image files), post flight output files (csrs ppp processing output files, ppk processed image files) and final outputs (digital surface model files, orthorectified mosaic files). Data derived from the image, in this case colony counts, will be published as separate datasets.
- The datasets are relatively large, raising issues of data storage. In addition, numerous large image files proved challenging to check.
- Quality control. Large numbers of large files prove challenging to check. The workflows are not yet in place and so a proportion of the data was checked for quality.
- UAV log files are commonly stored in proprietary formats. There appears to be no appetite within the industry at present to allow log files to be easily converted into open formats, making compliance with FAIR data principles difficult. For these datasets, the log files have been included in their proprietary format, as we took the view that having the data there is better than not being included.

Example citation: Fenney, N., Hollyman, P., Coleman, J., Fox, A., Trathan, P., & Collins, M. (2024). Aerial survey of the Hound Bay, South Georgia, fur seal colony during the 2021/22 season, flight DPLUS_109_12 (Version 1.0) [Data set]. NERC EDS UK Polar Data Centre. https://doi.org/10.5285/978f67ee-8d89-444b-8ea2-e2871a233d90

Summary and conclusion

RPAS are increasingly used in research across many disciplines. However, the management of the associated data can be challenging due to a lack of standards, the complexity of some data outputs and their associated sizes. To facilitate the use of RPAS-derived data in research and promote safe and ethical survey operations, the Research Data Alliance (RDA) Small Uncrewed Aircraft and Autonomous Platform Data Working Group has developed specific recommendations. In 5 steps, they advise to:

- Plan RPAS surveys early to get all necessary legal and privacy documents, including consents and data strategy, to avoid delays
- Fly the RPAS safely and record all necessary information following ethical and legal regulations to support later reuse of the data
- Document the processing pipeline and protect sensitive information through specific anonymisation or location generalisation techniques
- Store the data with all the required metadata to support long-term preservation

- Share the data through a trusted repository with a license to comply with the principle “as open as possible, as closed as necessary”.

These recommendations are a first step to support researchers, data practitioners and RPAS data users in managing their data from mission planning to publication. However, we acknowledge that some topics need to be further investigated. Those include:

- The development of training opportunities for scientists to learn more about data management and specifically related to UAV data. Some training material and a working example could be developed based on these recommendations.
- Further investigations towards data storage and NetZero considerations.
- Challenges around proprietary log files and their associated costs.
- The development of recommendations for enhanced data and metadata interoperability with computational infrastructure to allow cloud processing for online data analysis, visualisation, and other applications.
- Further guidance on describing complex processing pipelines in order to facilitate reproducibility.

References

‘AirHub - How to Use Standard Operating Procedures in Drone Operations’. Accessed 3 October 2025. https://www.airhub.app/resources/news/drone-standard-operating-procedures-in-high-stakes-operation.

Anderson, Philip, Luke Bateson, Barbara J. Brooks, et al. *Net Zero Aerial Capability (NZArC) Scoping Report*. UKRI Natural Environment Research Council, 2024. https://www.bas.ac.uk/wp-content/uploads/2024/10/NZArC_Scoping_Report_2024.pdf.

Aragon, Bruno, Kasper Johansen, Stephen Parkes, et al. ‘A Calibration Procedure for Field and UAV-Based Uncooled Thermal Infrared Instruments’. *Sensors* 20, no. 11 (2020): 3316. https://doi.org/10.3390/s20113316.

Barbieri, Lindsay, Jane Wyngaard, Sarah Swanz, and Andrea K. Thomer. *Making Drone Data FAIR Through a Community-Developed Information Framework | Data Science Journal*. 25 January 2023. https://doi.org/10.5334/dsj-2023-001.

Bierlich, K. C., Josh Hewitt, Clara N. Bird, et al. ‘A Workflow of Open-Source Tools for Drone-Based Photogrammetry of Marine Megafauna’. *PeerJ* 13 (2025): e19768. https://doi.org/10.7717/peerj.19768.

Carroll, Stephanie, Ibrahim Garba, Oscar Figueroa-Rodríguez, et al. ‘The CARE Principles for Indigenous Data Governance’. *Data Science Journal* 19 (2020). https://www.research.ed.ac.uk/en/publications/the-care-principles-for-indigenous-data-governance.

Chapman, Arthur D. *Current Best Practices for Generalizing Sensitive Species Occurrence Data. Version 1.* 2020. https://doi.org/10.15468/doc-5jp4-5g10.

CoreTrustSeal. 'CoreTrustSeal'. Accessed 3 October 2025. https://www.coretrustseal.org/.

Cramer, M., H. J. Przybilla, and A. Zurhorst. 'UAV CAMERAS: OVERVIEW AND GEOMETRIC CALIBRATION BENCHMARK'. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-2-W6 (August 2017): 85–92. https://doi.org/10.5194/isprs-archives-XLII-2-W6-85-2017.

Creative Commons. 'Homepage'. Accessed 3 October 2025. https://creativecommons.org/.

Cunliffe, Andrew M., Karen Anderson, Leon DeBell, and James P. Duffy. 'A UK Civil Aviation Authority (CAA)-Approved Operations Manual for Safe Deployment of Lightweight Drones in Research'. *International Journal of Remote Sensing* 38, nos. 8–10 (2017): 2737–44. https://doi.org/10.1080/01431161.2017.1286059.

'Current State Regulations'. Accessed 3 October 2025. https://www2023.icao.int/safety/UA/UASToolkit/Pages/State-Regulations.aspx.

Daniel Kissling, W., Yifang Shi, Jinhu Wang, et al. 'Towards Consistently Measuring and Monitoring Habitat Condition with Airborne Laser Scanning and Unmanned Aerial Vehicles'. *Ecological Indicators* 169 (December 2024): 112970. https://doi.org/10.1016/j.ecolind.2024.112970.

Daniels, Louis, Eline Eeckhout, Jana Wieme, Yves Dejaegher, Kris Audenaert, and Wouter H. Maes. 'Identifying the Optimal Radiometric Calibration Method for UAV-Based Multispectral Imaging'. *Remote Sensing* 15, no. 11 (2023): 2909. https://doi.org/10.3390/rs15112909.

DataCite Schema. 'DataCite Schema'. Website. Accessed 3 October 2025. https://schema.datacite.org/.

DCMI. 'Home'. 30 April 2025. https://www.dublincore.org/.

'DMP Tool'. Accessed 3 October 2025. https://dmptool.org/.

'DMPonline'. Accessed 3 October 2025. https://dmponline.dcc.ac.uk/.

'Drones & Air Mobility | EASA'. 2 October 2025. https://www.easa.europa.eu/en/domains/civil-drones.

Edney, A., Tom Hart, M. Jessopp, et al. 'Best Practices for Using Drones in Seabird Monitoring and Research'. *Marine Ornithology* 51, no. 2 (2023). https://doi.org/10.5038/2074-1235.51.2.1544.

Fenney, Nathan, Philip Hollyman, Jamie Coleman, Adrian Fox, Philip Trathan, and Martin Collins. 'Aerial Survey of the Hound Bay, South Georgia, Fur Seal Colony during the 2021/22 Season, Flight DPLUS_109_12'. Version 1.0. With Philip Hollyman, Nathan Fenney, Jamie Coleman, et al. NERC EDS UK Polar Data Centre, 24 October 2024. Image/jpeg,image/tif,image/dng,text/plain,application/pdf,application/RINEX,bb3,application/CSRS,text/csv, 1065 files, 14.80 GB. https://doi.org/10.5285/978F67EE-8D89-444B-8EA2-E2871A233D90.

Fremand, A., Klump, J., Manthorpe, S., Whitelaw, M., Gerard, F., Garland, W., Semong, T., George, C., & Ricketts, H. 'RDA Recommendations on Remotely Piloted Aerial System data management: Log sheet and acquisition checklist'. *Zenodo.* (2026) https://doi.org/10.5281/zenodo.18999774

Fremand, Alice. *Towards a Data Commons: Imagery and Derived Data from Autonomous and Remotely Piloted Aerial Vehicles*. November 2023. https://nora.nerc.ac.uk/id/eprint/536398/.

Fremand, Alice. *UAV Data Management Handbook*. November 2023. https://nora.nerc.ac.uk/id/eprint/536392/.

Frémand, Alice C., Julien A. Bodart, Tom A. Jordan, et al. 'British Antarctic Survey's Aerogeophysical Data: Releasing 25 Years of Airborne Gravity, Magnetic, and Radar Datasets over Antarctica'. *Earth System Science Data* 14, no. 7 (2022): 3379–410. https://doi.org/10.5194/essd-14-3379-2022.

'Full Article: A UK Civil Aviation Authority (CAA)-Approved Operations Manual for Safe Deployment of Lightweight Drones in Research'. Accessed 3 October 2025. https://www.tandfonline.com/doi/full/10.1080/01431161.2017.1286059.

GFZ German Research Centre For Geosciences, Humboldt-Universität Zu Berlin, Germany Karlsruhe Institute Of Technology (KIT), et al. 'Registry of Research Data Repositories'. Preprint, DataCite, 2013. https://doi.org/10.17616/R3D.

Global Indigenous Data Alliance. 'CARE Principles'. Accessed 3 October 2025. https://www.gida-global.org/care.

Graybeal), OntoPortal Alliance (Clement Jonquet, John. 'OntoPortal'. OntoPortal. Accessed 3 October 2025. https://ontoportal.org/.

'Guidelines for Calibration of Uncrewed Aircraft Systems Imagery | U.S. Geological Survey'. Accessed 3 October 2025. https://www.usgs.gov/publications/guidelines-calibration-uncrewed-aircraft-systems-imagery.

Habermann, Ted. 'Instruments@DataCite'. Metadata Game Changers, 13 September 2024. https://metadatagamechangers.com/blog/2024/9/12/instrumentsdatacite.

Harris, Colin M., Heike Herata, and Fritz Hertel. 'Environmental Guidelines for Operation of Remotely Piloted Aircraft Systems (RPAS): Experience from Antarctica'. *Biological Conservation* 236 (2019): 521–31. https://doi.org/10.1016/j.biocon.2019.05.019.

Hodgson, Jarrod C., and Lian Pin Koh. 'Best Practice for Minimising Unmanned Aerial Vehicle Disturbance to Wildlife in Biological Field Research'. *Current Biology* 26, no. 10 (2016): R404–5. https://doi.org/10.1016/j.cub.2016.04.001.

Hooper, David A., Barbara J. Brooks, Wendy E. Garland, and Graham A. Parton. *The National Centre for Atmospheric Science Image Metadata Standard (NCAS-IMAGE) Version 1.0*. 7 October 2022. https://doi.org/10.5281/zenodo.6368295.

Jordan, Tom, Carl Robinson, Tom Reed, and Rebecca Toomey. ‘Airborne Gravity Data over Marguerite Bay Collected with a Windracers Ultra UAV (2023/24 Season)’. Version 1.0. With Tom Jordan, Carl Robinson, Tom Reed, Rebecca Toomey, and UK Polar Data Centre. NERC EDS UK Polar Data Centre, 6 November 2024. Text/csv, 1 file, 15 MB. https://doi.org/10.5285/3A9C8604-2BCA-48C1-A40C-48A873076581.

Kissling, W. Daniel, Jinhu Wang, and Yifang Shi. ‘D4.10 Demonstrator Pipeline for Habitat Condition Metric Extraction and Parallel and Distributed Computing in a Cloud Environment’. *ARPHA Preprints* 7 (March 2026): e190520. https://doi.org/10.3897/arphapreprints.e190520.

Krahl, Rolf, Louise Darroch, Robert Huber, et al. *Metadata Schema for the Persistent Identification of Instruments*. Version 1.0b1. 2021. https://doi.org/10.15497/RDA00070.

LANDRS. ‘Syna Theme’. 7 September 2017. https://www.landrs.org/landrs-toolkit.github.io/.

Magagna, Barbara, Gwenaelle Moncoiffe, Maria Stoica, et al. *The I-ADOPT Interoperability Framework: A Proposal for FAIRer Observable Property Descriptions*. 4 March 2021. https://doi.org/10.5194/egusphere-egu21-13155.

Mark Schildhauer, Matthew B. Jones. *OBOE: The Extensible Observation Ontology, Version 1.1*. KNB Data Repository, released 2016. https://doi.org/10.5063/F11C1TTM.

‘Open Science - European Commission’. Accessed 3 October 2025. https://rea.ec.europa.eu/open-science_en.

‘Privacy Handbook | EASA’. Accessed 3 October 2025. https://www.easa.europa.eu/en/domains/civil-drones/privacy/privacy-handbook.

RDA. ‘The TRUST Principles: An RDA Community Effort’. 18 May 2020. https://archive.rd-alliance.org/trust-principles-rda-community-effort-0.

Stocker, Markus, Louise Darroch, Rolf Krahl, et al. ‘Persistent Identification of Instruments’. *Data Science Journal* 19 (May 2020): 18. https://doi.org/10.5334/dsj-2020-018.

‘Tips for Responsible Drone Use | U.S. Fish & Wildlife Service’. Accessed 3 October 2025. https://www.fws.gov/refuge/oregon-islands/tips-responsible-drone-use.

Tmušić, Goran, Salvatore Manfreda, Helge Aasen, et al. ‘Current Practices in UAS-Based Environmental Monitoring’. *Remote Sensing* 12, no. 6 (2020): 1001. https://doi.org/10.3390/rs12061001.

‘UAS Standard Operating Procedures’. National Severe Storms Laboratory National Weather Service National Oceanic and Atmospheric Administration, 2023. https://inside.nssl.noaa.gov/uas/wp-content/uploads/sites/35/2023/07/NOAA_UASD-Standard-Operating-Procedures-NSSL_071923.pdf.

‘UAS Toolkit Home’. Accessed 3 October 2025. https://www.icao.int/uas-toolkit-home.

‘Viewing Marine Life | NOAA Fisheries’. Accessed 3 October 2025. https://www.fisheries.noaa.gov/insight/viewing-marine-life.

Wang, Ning, Markus Christen, and Matthew Hunt. ‘Ethical Considerations Associated with “Humanitarian Drones”: A Scoping Literature Review’. *Science and Engineering Ethics* 27, no. 4 (2021): 51. https://doi.org/10.1007/s11948-021-00327-4.

‘Welcome to EarthPortal | Data Terra EarthPortal’. Accessed 3 October 2025. https://earthportal.eu/.

‘Welcome to the Planetary Data System’. Accessed 11 March 2026. https://pds.nasa.gov/.

West, Jonathan P., and James S. Bowman. ‘The Domestic Use of Drones: An Ethical Analysis of Surveillance Issues’. *Public Administration Review* 76, no. 4 (2016): 649–59. https://doi.org/10.1111/puar.12506.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, et al. ‘The FAIR Guiding Principles for Scientific Data Management and Stewardship’. *Scientific Data* 3, no. 1 (2016): 1. https://doi.org/10.1038/sdata.2016.18.

Wyngaard, Jane, Lindsay Barbieri, Andrea Thomer, et al. ‘Emergent Challenges for Science sUAS Data Management: Fairness through Community Engagement and Best Practices Development’. *Remote Sensing* 11, no. 15 (2019): 15. https://doi.org/10.3390/rs11151797.

‘Your Tasks: Data Management Plan | RDMkit’. Accessed 3 October 2025. https://rdmkit.elixir-europe.org/data_management_plan.

Appendix 1

This work has been primarily based on the following literature:

Previous work led by the RDA Interest Group:

- sUAV IG has conducted survey previously (see Wyngaard, J., Barbieri, L., Thomer, A., Adams, J., Sullivan, D., Crosby, C., ... & Bell, T. (2019). Emergent challenges for science sUAS data management: fairness through community engagement and best practices development. Remote Sensing, 11(15), 1797. https://doi.org/10.3390/rs11151797)
- Making Drone Data FAIR Through a Community-Developed Information Framework - https://datascience.codata.org/articles/10.5334/dsj-2023-001

Work led by members of the community:

- Fremand, Alice. 2023 UAV data management handbook. UK Polar Data Centre, British Antarctic Survey, 13pp. https://nora.nerc.ac.uk/id/eprint/536392/
- Fremand, Alice. 2023 Towards a data commons: Imagery and derived data from autonomous and remotely piloted aerial vehicles. UK Polar Data Centre, British Antarctic Survey, 24pp. https://nora.nerc.ac.uk/id/eprint/536398/
- Philip Anderson (SAMS), Luke Bateson (BGS), Barbara Brooks (NCAS), Richard Dale (SAMS), Nick Everard (UKCEH), Charlotte Francoz (NOC), Charles George (UKCEH), France Gerard (UKCEH), Tom Jordan (BAS), Zixia Liu (NCEO), Aser Mata (PML), Hugo Ricketts (NCAS), Carl Robinson (BAS), Pilvi Saarikoski (BAS), Beatrix Schlarb-Ridley (BAS), Kay Smith (BGS), James Strong (NOC), Martin Wooster (NCEO), 2024. A Review of the Role of Uncrewed Aerial Systems in the Decarbonisation Strategy of NERC Aerial Activities. https://www.bas.ac.uk/wp-content/uploads/2024/10/NZArC_Scoping_Report_2024.pdf

In addition, members of the working group have participated in various engagement activities to collect use cases. For each use case, the following information was asked:

Title: title of your example

Contributor: give a contact detail here with an email address.

URL link: if it exists, provide here a URL link.

Brief summary: give a summary explaining what your use case is about. Is it a dataset? A repository? A project? This summary can be as long or as short as you want.

Data life cycle stage: What process does this example relate to? Is it linked to data collection? Data processing? Data integration? Interoperability?

Target communities: Give here who this project targets or the field of studies

Outcome: Give here any outcome of the project if applicable

Challenges: Give here any challenges you have or are still encountering with your example.

Related projects: Give here any related project that might be interesting for the WG.

The questionnaire was purposely free-form to allow users to give as many details as they wanted. Additional feedback has been compiled from various engagement activities led either online or in person.

List of engagement activities:

Name	Type of engagement	What
Monthly meeting	Internal	Working sessions (online)
RDA P22	External - International	Oral session (online)
RDA P24	External - International	Oral Session (online)
EGU 2025	External - International	Poster (in-person)
NZArC	External - UK	Presentation (online and in-person)
RDA P25	External - International	Oral session (in-person)

Appendix 2: Log sheet and RPAS acquisition checklist

On the field, it is important to record key metadata to make sure you can process the data you have collected using your Remotely Piloted Aerial System (RPAS) and they can be reused in the future. This document comprises a checklist and logsheets you can use to best plan your survey. An editable version can be found here: Fremand, A., Klump, J., Manthorpe, S., Whitelaw, M., Gerard, F., Garland, W., Semong, T., George, C., & Ricketts, H. (2026). RDA Recommendations on Remotely Piloted Aerial System data management: Log sheet and acquisition checklist. Zenodo. https://doi.org/10.5281/zenodo.18999774

How to use this document:

- Download the document
- Modify it to take into account your specific requirements
- Bring the templates with you on the field
- Use the template on the field to document the survey

Data collection checklist

Here is a list of metadata information and activity you might want to record as part of your survey:

Summary data collection document

- Project
- Survey name
- Start date/time of first flight of the survey
- End date/time of last flight of the survey
- Geographic extent of the survey
- Total numbers of flights
- Platform make and model
- Mission planning software
- List of sensors make and model and possible configurations. For each sensor, metadata might include make, model, sampling rate, resolution, measurement capability and accuracy, their location on the platform such as mount and orientation, firmware and version
- Calibration information
- General comments

Log information for each flight

- Flight ID
- Date
- Start time flight
- End time flight
- Latitude start
- Latitude end
- Longitude start
- Longitude end
- Operator name
- Weather information (visibility/cloud cover)
- Platform
- List of sensors
- Logs of events including time, comments, cause, effects of events
- Specific configuration notes (are all the sensors on for the flight?, location of the sensors on the platform)
- Calibration done/ calibration flight
- Make sure it is recording data and GNSS information

UAV summary data collection

Project	
Survey name	
Date and time of first flight	
Date and time of last flight	
Total number of flights	

Geographic extent

Latitude min	
Latitude max	
Longitude min	
Longitude max	

Insert map here

Summary of the survey

...
...
...
...
...
...

Description of the platform:
Name: ....
Make and model:...
Payload:..
Mission planning software:...

Description of sensors:

Sensor 1

- Name:……………………………………………………………………………………………
- Make and model:………………………………………………………………………
- Sampling rate:……………………………………………………………………………
- Resolution/accuracy:…………………………………………………………………
- Location on drone:……………………………………………………………………
- Mount and orientation:…… (here several configurations can be mentioned) …………
- Firmware and version:…………………………………………………………………
- Calibration:………………………………………………………………………………

Sensor 2

- Name:……………………………………………………………………………………………
- Make and model:………………………………………………………………………
- Sampling rate:……………………………………………………………………………
- Resolution/accuracy:…………………………………………………………………
- Location on drone:……………………………………………………………………
- Mount and orientation:…… (here several configurations can be mentioned) …………
- Firmware and version:…………………………………………………………………
- Calibration:………………………………………………………………………………

Sensor 3

- Name:……………………………………………………………………………………………
- Make and model:………………………………………………………………………
- Sampling rate:……………………………………………………………………………
- Resolution/accuracy:…………………………………………………………………
- Location on drone:……………………………………………………………………
- Mount and orientation:…… (here several configurations can be mentioned) …………
- Firmware and version:…………………………………………………………………
- Calibration:………………………………………………………………………………

Calibration information:

………

………

………

………

UAV flight survey log sheet

Flight ID:	**Site:**
Pilot:	**Observer:**
art:	**d:**
Time start:	**Time end:**
Latitude start:	**Latitude end:**
Longitude start:	**Longitude end:**
Wind speed/direction:	**Visibility(>5km, moderate, poor, <5km):**
Precipitation: Y/N	**Cloud cover:**

GNSS ON	YES/NO
Calibration information	
Recording data	YES/NO

List of sensors in use:

□ **Sensor 1** □ **Sensor 2** □ **Sensor 3** □ **Sensor 4**

Configuration notes: ....

..

Time	Comment / Cause / Effect

Time	Comment / Cause / Effect
Notes :	