

Data Deliverables Summary for ROV Jason

Overview

Introduction

The data package received at the conclusion of a *Jason* cruise is a collection of real-time automated vehicle and navigation logs, raw data collections, user event logs, vehicle video, video stills, and post-processed products. This document is intended to describe the contents and organization of the package so groups can process and use it more efficiently after the cruise.

Data Deliverables

General organization of the data package

The *Jason* data package will be delivered on one or more portable USB hard drives to the Chief Scientist for the cruise. This is a first-generation copy of the complete data package, provided for transport back to the Chief Scientist's home institution. Long-term and online preservation of the data package will be managed by WHOI in accordance with the [NDSF data policy](#).

When the data package is delivered at the end of the cruise, the chief scientist will be asked to:

1. Review a form that describes the package inventory and acknowledges delivery of the package.
2. Assign embargo duration on the various data components ([NSF policy](#) states a maximum of 2 years).
3. Name institutions for whom acknowledgement is due when cruise data is used for outreach or commercial purposes.

The content on the hard drives will resemble the following, though variations may occur depending on *Jason* configuration and specific sensors used during the cruise:

Vehicle Data Deliverables

Vehicle Data Hard Drive:

Cruise ID

|-----[Documentation](#)

|----Sealog
|----H264Recordings
|----HighlightsHD
|----HighlightsUHD
|----SciCamStills
|----VidGrabs
|----Multibeam
|----Nav
|----Renav
|----Vehicle

1. Documentation

Location: /CruiseID/Documentation/

Description: This directory contains all documentation and metadata necessary for interpretation of the contents of the complete *Jason* data package. The contents are described below.

Contents:

- This summary
- Vehicle metadata spreadsheet of cruise and lowering statistics and logs
- Automated dive and cruise summary reports
- Vehicle data formats definitions file: JASON2_ROV_DataFormats_vX.X.xls
- Applicable sensor calibration records and certificates
- Pre-dive vehicle configuration photographs
- Copy of the NDSF Data Archive Policy
- Copy of the data archive acknowledgement and assignment form

2. Sealog

Location: /CruiseID/Sealog/

Description: Sealog is an augmented event logger. During the cruise, watchstanders use a web interface to make event entries using either free text or hot buttons, while sensor data and video framegrabs are co-registered at the time of each event. Periodic autosnaps of data and framegrabs trigger to fill in the narrative between manual user events. The result is a web browsable review of the lowering, available onboard during the cruise and ashore following the cruise.

The lowering narrative is augmented in two ways in post-processing. In one process, native Sulis/SciCam still image captures are made into an event and added to the Sealog timeline. Additionally, a renavigation process is applied to improve position estimates. Both real time and renavigated positions are displayed in the final Sealog product, with the positional difference between the two (in meters) shown. See Section 5.2 for a detailed description of the renavigation process.

The Sealog content included in the cruise data package serves a dual purpose. Firstly, the .csv file provided furnishes a comprehensive table of events incorporating sensor measurements alongside the titles of the co-registered framegrabs located in the “images” subdirectory. Furthermore, these and the supplementary files and formatting are provided to facilitate the transfer of the package to a shoreside server.

Shoreside access: After the cruise, the full narration will be hosted on a shoreside server, which can be reached at <https://sealog.whoi.edu/sealog-jason/> . Viewing of this shoreside server content can be password protected for up to two years.

3. Video and Imaging System Products

3.1. H264 Recordings

Location: /CruiseID/H264Recordings/J2-####/

Description: Three 1080i video streams ("pilot", "brow", and "science") are captured by duplicate direct-to-hard drive recorders and sorted by lowering in this directory. The presented files are MPEG Transport Stream (.ts) compressed using the h.264 codec, which provides excellent viewing quality but is unsuitable for production video editing. Filenames include the camera name and start timestamp and are clipped at a definable time interval, usually 15 minutes. In addition to the video files, metadata that is broadcast in real-time on the Jason network is captured in order to produce subtitle files (.srt format).

In post-processing, video and subtitle files are merged into a Matroska container file (.mkv) to yield a video with optional overlays that are best viewed using software such as *VLC Media Player*. An additional subtitle file is also generated which utilizes the renavigated vehicle positions. See Section 3.6 for a detailed description of the renavigation process. When viewing these files, there are five .srt files to choose from in playback, each of which will produce a line of text overlain on the video with the following information:

Subtitle file 1: UTC time, real time latitude, real time longitude, heading, depth.

Subtitle file 2: vehicle ID, cruise ID, lowering ID, date

Subtitle file 3: localX, localY, roll, pitch, altitude

Subtitle file 4: origin latitude, origin longitude, UTM zone

Subtitle file 5: UTC time, renav latitude, renav longitude, real time heading, real time depth.

Contents:

- Metadata: Text (.txt) files, named in the same yyyyymmddhhmmss.ssssss format as the available video files, capturing the metadata used in subtitle files in timestamped stanzas
- Subtitles: Subtitle (.srt) files containing the information types outlined in the subtitle file types list above.

- Videos: Original video recording files written to hard drive. These are MPEG Transport Stream (.ts) files compressed (output rate 6000 kbps) using the h.264 codec. Image resolution is 1920x1080 pixels.
- MKV files: The Matroska container video files named in the format *Cam_yyyymmddhhmmss.ssssss.mkv and merged with optional metadata subtitle overlays, produced in post-processing.

3.2. High and Ultra High Definition Video Highlights

Location: /CruiseID/HighlightsHD/J2-#### & /CruiseID/HighlightsUHD/J2-####

Description: For those moments that may be processed for production or outreach, select activities of the dive can optionally be captured from the watchstander's choice of high definition (1080i) pilot, brow, or ultra-high definition (2160p/4K) science camera streams. These clips are minimally compressed using the Apple ProRes422 family of codecs (2.7GB/3 minutes for high definition, 15GB/3 minutes ultra definition) and contain embedded time code. The resulting file is a .mov file playable by QuickTime and editable using NLE software such as Final Cut or Premiere.

A summary log of these clips for the cruise can be found as part of the full-cruise metadata spreadsheet located in the Documentation directory, in addition to the raw log files found in this directory.

Contents:

- MOV files: Video clips renamed to include lowering ID, resolution, start time, and stop time and titled in the format J2-####-definition-yyymmddhhmmss-yyymmddhhmmss.mov
- ExtractedTimeMetadata: A compilation of files in the format *_tcmd.txt, sourced from a command line query to extract metadata from each clip. This information is used in building filename and time code information for each clip.
- Batch_list.csv and .txt: Information from extracted time metadata, organized in a format to support the clip renaming process to timestamped file names.
- Log_J2-####_hilites.csv: A log file documenting the translation of filenames from the originally recorded clips to the final product filenames.

3.3. Science Camera Stills

Location: /CruiseID/SciCamStills/J2-####/

Description: This camera may act as both a 4k UHD video camera and as a still image camera at 5968x3352. It is configured to take photographs under the control of the science party watch leader with the press of a button. The action of capturing this photo does not affect the 4K video stream or any downstream products. Alternate still camera configurations may be discussed as part of the pre-cruise planning process. Photos are retained on an internal memory card and are offloaded once the vehicle is on deck. Images are renamed according to the capture timestamp

imbedded in the EXIF field of each image. For those captures that fell within the renavigation time frame, a table relating the still image name with the associated navigation information (.ppfx file) accompanies the image directory.

Contents:

- JPG files: Images renamed from the originals according to the timestamp embedded in the EXIF fields of the individual photos, and modified in their IPTC fields to include creator and copyright information.
- PPFX files: For those captures that coincided with the renavigation timeframe, a table that combines still image name with navigation info accompanies the image directory.

3.4. Video Grabs

Location: /CruiseID/VidGrabs/J2-####/

Description: A framegrab system captures from a selectable variety of sources, primarily the Insite Mini-Zeus cameras known as SciCam, BrowCam, and PilotCam. The system works natively on 4k streams, grabbing from the SciCam stream at 3840x2160 pixels. It scales images from 1080i streams, saving the images at 2740x1540 pixels. Color space is RGB, bit depth is 8 bits.

Contents:

- TIF files: Captured images with copyright information added to the IPTC field and renamed in the format (cam)_yyyymmddhhmmss.sss.tif
- JPG files: Captured images if .jpg was selected as the original file type. Images have had copyright information added to the IPTC field and are renamed in the format yyyymmdd_hhmmss_(cam).jpg.
- JPG directory: Thumbnail images in .jpg format corresponding to any .tif files located in the directory above.

4. Multibeam

Location: /CruiseID/Multibeam/J2-####/

Description: *Jason* can carry a 200-400 kHz Kongsberg 2040 multibeam sonar when required. Use of the multibeam is arranged in the pre-cruise planning process. Raw data is provided in the proprietary Kongsberg .kmall format. The *Jason* at-sea data engineer may produce a quick-look gridded product that is based on a first-cut renavigation and multibeam soundings that have been processed using automated scripts. Any further processing is conducted by the science party and is not included as a standard data product.

5. Navigation

5.1. Navigation Auxiliary Information

Location: /CruiseID/Nav/

Description: The Nav directory contains files auxiliary to the primary raw and processed navigation data presented in the Vehicle and Renav directories, respectively. This navigation information may include navigation interface underlay maps, targets, and initialization files, as well as configuration and exported files from Jason's USBL navigation system. As many lowerings occur within the same geographic area, contents of this directory are not sorted and contain information used across all lowerings for the cruise.

5.1.1 NavG

Location: /CruiseID/Nav/NavG/

Description: Contents of the NavG directory are sourced from the control van's navigation computer running the in-house NavG software interface.

Contents:

- **Targets:** These are CSV files with either .csv or .tgt file extensions containing labeled latitude and longitude vehicle position targets displayed in the surface navigation user interface during a lowering. These targets include both those created during the pre-planning process, and those created manually by the navigator during the course of the lowering. These files may alternatively be found in /CruiseID/Vehicle/Rawdata/Navest/.
- **Origins:** A .txt or series of .txt files containing the latitude and longitude of the navigational origin point(s) in use during the cruise. Vehicle control systems use simple equidistant coordinates with a fixed scaling between meters displacement from the origin, which has several advantages for realtime control of a vehicle. Unlike UTM grid coordinates, the x and y axes intersect at right angles and align with true east and north respectively at the origin. Origin information is required for any translations between vehicle x and y positions and latitude and longitude.
- **Maps:** The map subdirectory may contain georeferenced .grd files providing a bathymetric underlay used in the vehicle's navigation user interface, or simple image files in .png, .tif, or .bmp format along with a corresponding text file geolocating the boundaries of the image file.
- **Inifiles:** Initialization configuration files used by the vehicle's navigation software, specifying information such as origin, vehicle and map display preferences, and navigation source messages.

5.1.2 USBL

Location: /CruiseID/Nav/USBL/

Description: USBL (Ultra-Short BaseLine) navigation data is sourced directly from the Sonardyne Ranger2 USBL system located either in the Jason control station or on the host ship. The entirety of the raw system log files for the cruise are exported in the manufacturer's fixed .csv file format. This section may also include exported configuration and calibration records from the Ranger2 software.

5.2 Renavigated Vehicle Position History

Location: /CruiseID/Renav/

Description: Following a lowering, the at-sea data engineer executes dedicated post-processing scripts known as renavigation. This performs an initial 'grooming' of the navigation data and generates an output file that merges inertial and USBL navigation data. Real-time position estimates are improved as a result of the post-processing. The primary product of interest produced by this process is a 1Hz ASCII .ppi file containing post-processed navigation data and a selection of other sensor data from the lowering.

6. Vehicle

Location: /CruiseID/Vehicle/

Description: Logged vehicle data is collected in two top-level directories, Rawdata and Procddata. It is expected that most users will use the material from the Procddata hierarchy. The Rawdata hierarchy is provided to support any obligations to submit raw cruise data to a national data center.

6.1 Raw Vehicle Sensor Data

Location: /CruiseID/Vehicle/Rawdata/

Description: Directory Rawdata contains sensor data as it was collected in real time, captured as hourly files over a calendar day (UTC). Under Rawdata are the logfiles collected by *navest*, the core navigation software package used by Jason, and by control computers. It will also contain navigation sensor information as it was collected through the vehicle's *Robot Operating System* software.

- Reference Documentation:
See /CruiseID/Documentation/JASON2_ROV_Data_FormatsvX.X.xls

Contents:

- YYYYMMDD: Raw vehicle sensor data logged as hourly yyyyymmdd_hhmm.XXX files according to sensor type and sorted over a calendar day (UTC).
- Roslog: System message log files generated in the execution of *Robot Operating System* processes.

- Rosbag: Bag files used in ROS for recording and replaying the data utilized in the vehicle's *Robot Operating System*.
- Navest: Raw navigation data as captured in sensors' native format and native sampling rate by the vehicle's *navest* navigation server and logged as hourly yyyyymmdd_hhmm.DAT ASCII files for the duration of the cruise.

6.2 Processed Vehicle Sensor Data

Location: /CruiseID/Vehicle/Procdata/J2-####/

Description: After a lowering, files from the Rawdata directory have been concatenated by record type and restructured in the following ways:

1. Hourly files copied so that type-by-day also exists as type-by-cruise.
2. Hourly files copied so that type-by-day also exists as type-by-lowering.
3. Hourly files concatenated so that type-by-hour also exists as type-by-lowering.
4. The processed .csv filetype created by reformatting type-by-lowering files and adding appropriate column headers.
5. Navest and ROS logs copied so that they exist by-lowering.

Contents:

- J2-####_(recordtype).raw: Vehicle sensor data sorted by record type as it exists for the lowering, derived from the raw yyyyymmdd.(recordtype) files captured in the Rawdata directory and sorted by lowering.
- J2-####_(recordtype).csv: Files which mirror the corresponding .raw files above but have been reformatted to comma separated value records to better facilitate data ingestion.
- Rosbag: Original bag files originated from *Robot Operating System* and sorted to include only those from the designated lowering.
- Navest: Navigation sensor data for the lowering captured by the vehicle's *navest* navigation server in *.DAT ASCII files, along with a collection of Matlab processing and output files resulting from the renavigation process discussed in section 5.2. A copy of the resulting .ppi file is also located in the directory.
- (Recordtype): Directories containing the relevant hourly yyyyymmdd_hhmm.(recordtype) sensor files covering the timespan of the lowering of interest.

NDSF contacts

- Jason Vehicle Manager
 - Matt Heintz, mheintz@whoi.edu, 508-289-3426
- Jason Data Engineer
 - Scott McCue, smccue@whoi.edu, 508-289-3462
- NDSF Associate Director for Data and Science Ops

- Christina Haskins, haskins@whoi.edu, 508-289-3920
- NDSF Director
 - Andrew Bowen, abowen@whoi.edu, 508-289-2643
- NDSF Chief Scientist
 - Dr. Anna Michel, amichel@whoi.edu