

# R2R: Cruise Directory Structure

Robert Arko  
Lamont-Doherty Earth Observatory

# Goal

Develop a fleet-standard cruise directory structure to:

- Differentiate routine underway, national facility, and science party data sets
- Differentiate data files from documentation
- Identify the instrument system that produces each data set

# Constraints

- Cruise ID is unique (within R2R system)
- Standard vocabulary of instrument type, make, and model
- Preserve “all” of the “original” data
- Preserve version of data that the science party took home
- Never break up or rename data files

# Issues

## 1. Compression

Files or directories?

Algorithm? (gzip, ZIP, etc)

Is it even needed?

# Issues

## 2. Checksums

MD5 most common but SHA1 is “better”

single top-level manifest?

# Issues

## 3. Filesystem type

NTFS most flexible? (Windows, Linux, MacFUSE)

# Issues

## 4. Filename/directory structure

spaces and/or non-ASCII characters

capitalization?

filename = {cruise\_}device{\_datatype}\_time{.suffix}

data files hourly? daily?

raw vs. processed in different directories?

dedicated “docs” directory

# Issues

## 5. “Public folder”

sensitive/non-science photos, emails, etc.





Thank you.