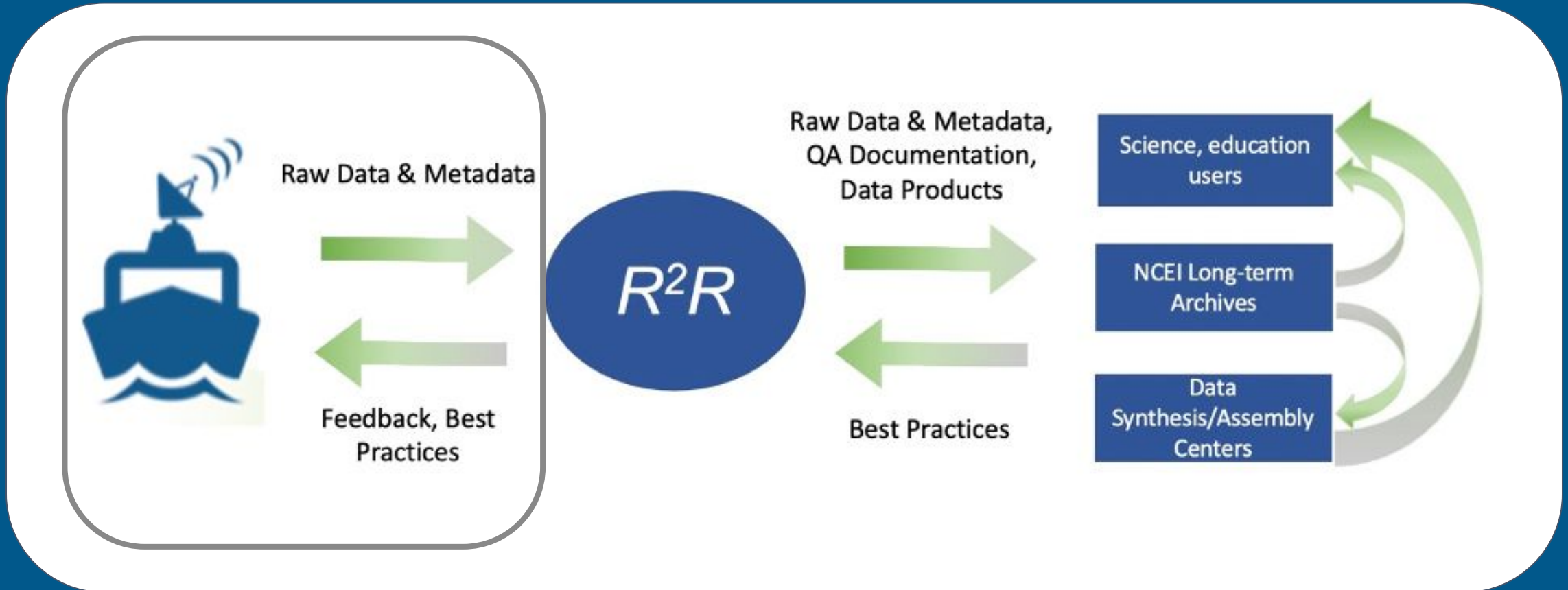


Towards Known Data Formats

Dru Clark (SIO) R2R

R2R Workflow



Device Data Format Descriptions

- Necessary for scientists to re-use these data in the future, and for NCEI to accept data for archive
- Describes the structure of the file such as the name and units of each column
- Allows for the end user to know how to use a file
(Consider if you had to write a data parser)

Current Status

- Not all formats are described
- File formats change over time
- Reviewing methods to keep this up to date



Current Data Formats

Well-described

- Self describing file formats NetCDF (multibeam data)
- Output from systems like SCS, Coriolix and OpenVDM
- Known data types - segy
- Proprietary formats - ctd

Format descriptions not routinely available for others



Some Self-describing Approaches

OSU Coriolix

- Headers in the file describe the data columns

NOAA SCS

- xml configuration file

SIO Met

- pdf file

Others we don't know about?



Questions

- Do you know how your vessel stores format info?
- What do you use as format documentation?
 - code?

