A Comprehensive Machine-Readable Metadata Standard for Ocean Sensors

Eric Rehm[1], Brian King[2], Jean-Michel Leconte[3], Kim Martini[4]

[1]Sea-Bird Scientific, Bellevue, WA, USA , [2]National Oceanography Centre, Southampton, UK, [3]RBR Ltd, Ottawa, ON, Canada, [4]Tini Scientific, Seattle, WA, USA

OSM 2024, #ED44C-0223

What's the problem? (Argo static metadata example)

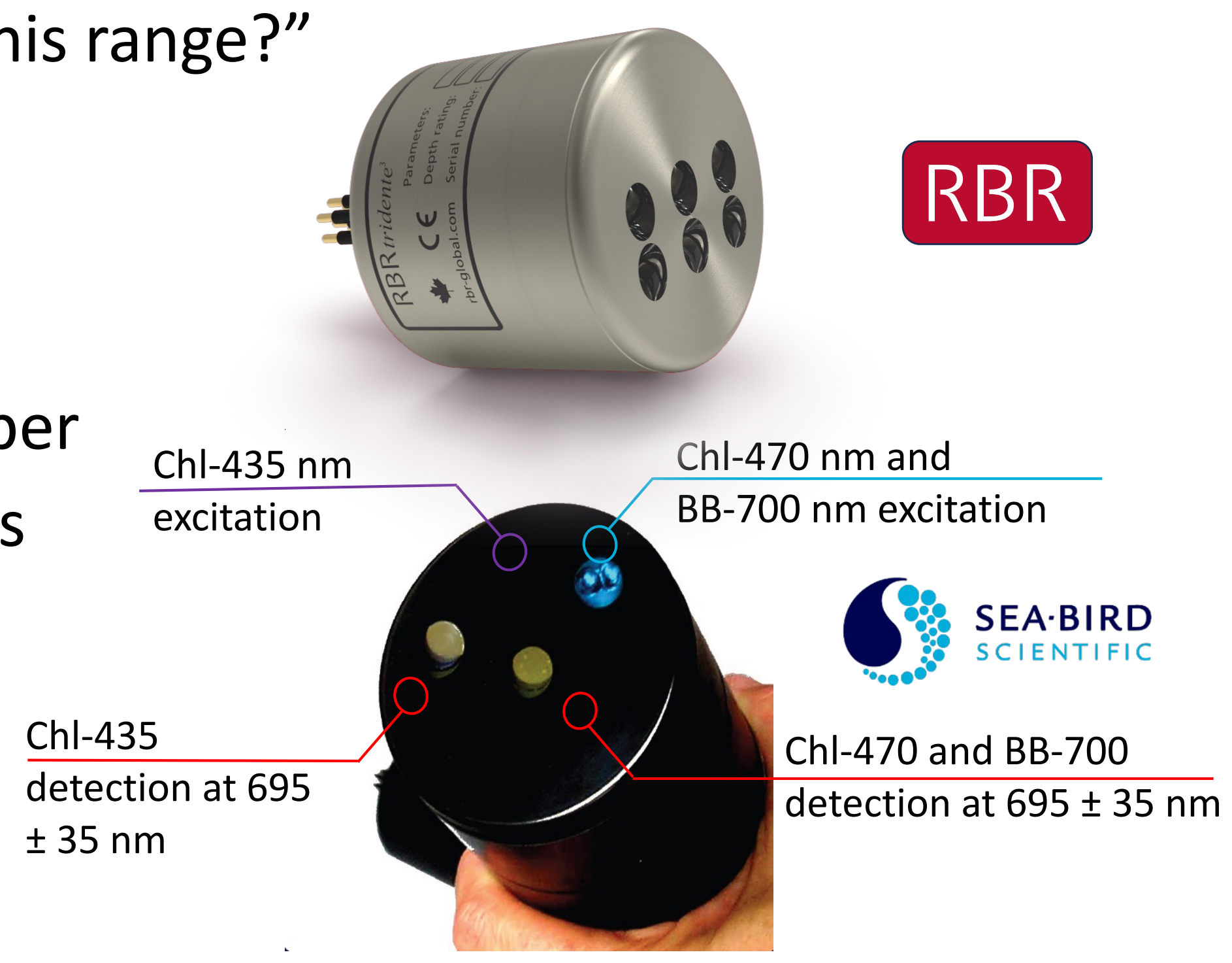
NUMBER OF ARGO FLOATS	NUMBER OF SENSORS / FLOAT
3894	3 (Core), 4 (O_2), 9 (Bio-Argo)
METADATA ENTRIES / SENSOR	NUMBER OF STATIC METADATA VALUES
15-20	200,000+

- Human transcription errors can mean bad data. A 0.5% error rate (optimistic) = 1000 metadata errors.
- Metadata standards enable search, discovery, and solutions to problems found in the field: "Which Argo floats have sensor serial numbers in this range?"

Metadata: Data about the data

To describe either of these 3-channel ocean sensors:

- Name, Maker, Model, Firmware version, Serial Number
- Excitation and Emission bandwidths and wavelengths
- Parameter units, accuracy, resolution
- Calibration coefficients
- Calibration date, calibration equation
- Links to other artifacts (calibration PDF)


RBR
Chl-435 nm excitation
Chl-470 nm and BB-700 nm excitation
SEA·BIRD SCIENTIFIC
Chl-435 detection at 695 ± 35 nm
Chl-470 and BB-700 detection at 695 ± 35 nm

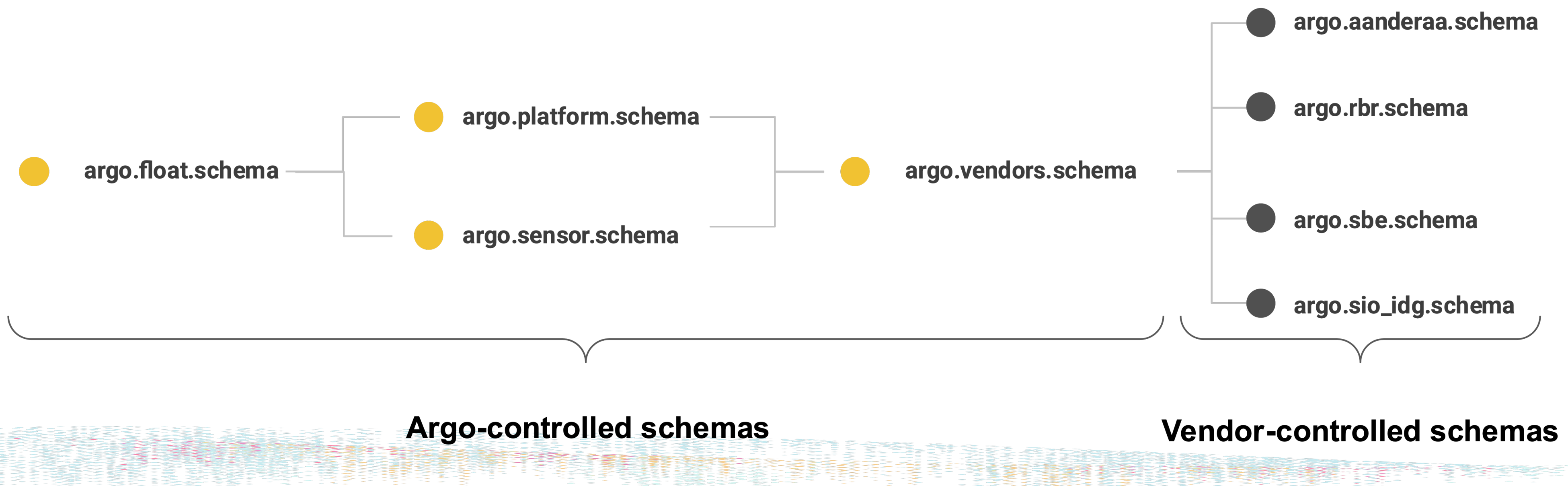
JSON Schema-driven standard sensor metadata

```
"SENSORS": {
  "type": "array",
  "items": {
  "type": "object",
  "properties": {
    "SENSOR": {
      "type" : "string",
      "format" : "uri",
      "pattern" : "^SDN:R25::",
      "description": "SENSOR string must be valid in current Argo reference table R25",
      "validation-uri" : "https://vocab.nerc.ac.uk/collection/R25/current/"
      },
    "SENSOR_MAKER": {
      "type" : "string",
      "format" : "uri",
      "pattern" : "^SDN:R26::",
      "description": "SENSOR_MAKER string must be valid in Argo reference table R26",
      "validation-uri": "https://vocab.nerc.ac.uk/collection/R26/current/"
      },
```

Schema features that support machine-based validation:

- **Semi-structured:** organization, no enforced property order
- **Typed fields:** "array", "string", "uri", regular expressions for pattern validation
- **Controlled vocabulary**: URI's reference NERC Vocabulary Server entries: "SDN:R25::"
- **Current schema conforms to Argo User Guide, V3.41.1,** http://dx.doi.org/10.13155/29825
- **GitHub Repository & Issue Tracker:** https://github.com/euroargodev/sensor_metadata_json

Argo Float/Sensor metadata schema: controlled but flexible


argo.float.schema

argo.platform.schema

argo.sensor.schema

argo.vendors.schema

argo.aanderaa.schema

argo.rbr.schema

argo.sbe.schema

argo.sio_idg.schema

Argo-controlled schemas

Vendor-controlled schemas

JSON libraries with schema validators exist for most languages

C/C++/C#, FORTRAN, Java/JavaScript, MATLAB, Perl, PHP, Python, Ruby, Visual Basic, etc.

Validators point to errors: On instance['SENSORS'][1]: 'SENSOR_MODEL' is a required property

Programmatically creating or modifying a JSON description is easy in any language:

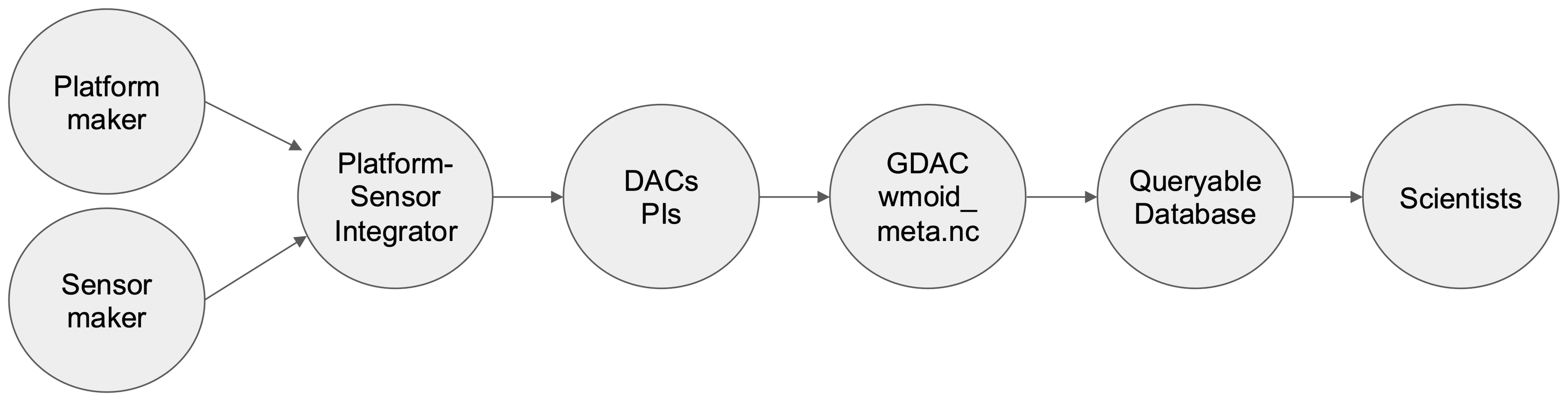
```
data['SENSORS'][0]["SENSOR_MODEL"] = TRIOS_model_map[instrument_type]
data['SENSORS'][0]["SENSOR_MODEL_FIRMWARE"] = fw_version
data['SENSORS'][0]["SENSOR_SERIAL_NO"] = instrument_sn
data['SENSORS'][0]["sensor_vendorinfo"]["TRIOS_RAMSESType"] = instrument_type
```

```
Json("SENSORS")(2)("SENSOR") = "SDN:R25::" & Sensor
Json("SENSORS")(2)("SENSOR_MODEL") = "SDN:R27::" & Map.Model
Json("SENSORS")(2)("SENSOR_SERIAL_NO") = SN
Json("PARAMETERS")(2)("PREDEPLOYMENT_CALIB_DATE") = CalDate
```

Visual Basic

Metadata Value Chain: From Makers to Scientists

It can be a long journey from Sensor Makers to Scientists.


Platform maker

Sensor maker

Platform-Sensor Integrator

DACs PIs

GDAC wmoid_ meta.nc

Queryable Database

Scientists

We encourage machine-to-machine JSON metadata delivery via a REST API. A fake example:
https://api.example.com/v1/metadata?serial_number=205908&format=argometadatajson

Schema-compliant JSON sensor description

```
"SENSORS": [
  {
    "SENSOR": "SDN:R25::CTD_PRES",
    "SENSOR_MAKER": "SDN:R26::RBR",
    "SENSOR_MODEL": "SDN:R27::RBR_PRES_A",
    "SENSOR_SERIAL_NO": "205908",
  },
```

```
  "predeployment_vendorinfo": {
    "vendor_schema" : "RBR",
    "version" : "0.1",
    "certificate" : "https://oem-lookup.rbr-
    global.com/api/v1/instruments/205908/chann
    els/8/certificate"
    }
```

```
"PARAMETERS": [
  {
    "PARAMETER": "SDN:R03::PRES",
    "PARAMETER_SENSOR": "SDN:R25::CTD_PRES",
    "PARAMETER_UNITS": "dbar",
    "PARAMETER_ACCURACY": "1",
    "PARAMETER_RESOLUTION": "0.02",
    "PREDEPLOYMENT_CALIB_EQUATION": "Pcorr = X0+(Pmeas-X0-X1*(Tpres-X5)-X2*(Tpres-X5)^2-
    X3*(Tpres-X5)^3)/(1+X4*(Tpres-X5)); Pmeas = C0+C1*VR+C2*VR^2+C3*VR^3;",
    "PREDEPLOYMENT_CALIB_COEFFICIENT_LIST": {
      "C0": "-55.76767", "C1": "4.0054912E3", "C2": "-66.53745", "C3": "6.39357",
      "X0": "10.0361",   "X1": "184.89905E-3","X2": "330.90703E-6",
      "X3": "-999.3263E-9", "X4": "-86.53429E-6", "X5": "21.942171"
    },
```

(This example JSON sensor description can be retrieved online. Visit oem-lookup.rbr-global.com)

Conclusions

Starting in 2020 with a web API implemented by RBR Ltd., academic and industry members of the Argo Data Management Team have developed a comprehensive sensor and float metadata standard that enables machine-to-machine sensor metadata communication, thereby reducing human transcription errors and enabling more comprehensive query and search for oceanographic platform and sensor data. Sensor makers can easily augment existing calibration sheet production by generating JSON metadata from their calibration databases and workbooks.

The developers of this schema encourage other sensor/platform vendors and observing systems, (vessel fleets, glider fleets, and ocean observatories/IOOS/OOI/etc.) to consider this standard.