Compliance Checker

Presented by Luke Campbell
Background
The standards in practice within IOOS and Communities

- **The netCDF User's Guide (NUG)**
- **Cooperative Ocean/Atmosphere Research Data Service (COARDS)**
- **netCDF Attribute Convention for Dataset Discover (ACDD)**
- **NetCDF Climate and Forecast (CF) Metadata Conventions**
- **NCEI NetCDF Templates**
Background

- Many Standards
- Each overlap in some degree
- They don't always intersect nicely
The goal of the IOOS Compliance Checker is to provide data providers with confidence that their data is compliant and consistent with the community.
Compliance Checker: Core Standards
Compliance Checker: Core Standards

Original Goal was for ACDD attributes.

From ACDD we expanded into CF and IOOS Vocabulary Metadata

Our latest release has ACDD 1.3 and support for plugins

- CF 1.6
- ACDD 1.1
- ACDD 1.3
## Corrective Actions

<table>
<thead>
<tr>
<th>Name</th>
<th>Priority</th>
<th>Corrective action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventions</td>
<td>3</td>
<td>Attr Conventions does not contain 'ACDD-1.3'</td>
</tr>
<tr>
<td>acknowledgment/acknowledgement</td>
<td>3</td>
<td>Neither 'acknowledgment' nor 'acknowledgement' attributes present</td>
</tr>
<tr>
<td>creator_email</td>
<td>2</td>
<td>Attr creator_email is empty or completely whitespace</td>
</tr>
<tr>
<td>creator_institution</td>
<td>1</td>
<td>Attr creator_institution not present</td>
</tr>
<tr>
<td>creator_name</td>
<td>2</td>
<td>Attr creator_name is empty or completely whitespace</td>
</tr>
<tr>
<td>creator_type</td>
<td>1</td>
<td>Attr creator_type not present</td>
</tr>
<tr>
<td>date_metadata_modified</td>
<td>1</td>
<td>Attr date_metadata_modified not present</td>
</tr>
</tbody>
</table>
Plugin-based Framework
Plugin-based Framework

Introducing a plugin-based framework

- Provides support for more checkers without overwhelming the main project
- Allows groups and developers to build their own compliance checkers based on the plugin framework
Plugin-based Framework

• Our first plugin that we published online was the GliderDAC checker

• Why we needed a checker just for GliderDAC

• Integrating it with the web-checker
Plugin-based Framework

• More plugins on the way:
  • A plugin for the NCEI Templates. Version 1.1 checker is almost ready
  • NCEI Templates checker 2.0 will follow shortly
  • A plugin for s-grid started in February
• The framework’s approach:
  
  • Given a dataset, return a result
  • Each result has a severity, LOW, MED, HIGH
  • Each result has a score and an "out of" value
  • Each result without a perfect score has help messages for how to fix

• This approach enables virtually any kind of metadata checking for netCDF datasets.
More than a command line tool
More than a command line tool

After a year of development we saw the immediate limitations of having something that only works as a command line tool:

- Challenging to install
- Output was hard to use in other programs
- Corrective actions messages weren't intuitive
  - "What do you mean "cf_role exists!?"
More than a command line tool

Improvements as a library

- Switching to Anaconda
- Windows support and UDUnits.py
- Unicode Troubles
- Improving Messages
- Adding new output capabilities:
  - JSON
  - HTML
  - Plaintext
More than a command line tool

• Creating a web interface
• Online web checker
Online Web Checker
Online Web Checker

• No python required

• Just an internet connection and a dataset

• Supports uploads and links to OPeNDAP

• Reports are in a printable format
Demonstration