Implementing change in CF: An idea’s journey

Daniel Lee

EUMETSAT
Contents

• How are changes made to the CF Conventions?

• How is this implemented in GitHub?

• What improvements might we introduce?
Changes: Process spec

- Text form in **Rules**

- Fear not! This is actually pretty straightforward

- Exception: Errata
Implementation in GitHub

• This is a snapshot, see the repository’s CONTRIBUTING.md for current rules

• Issue is discussed in a single issue

• Text proposal is a pull request

• Moderator volunteers to guide discussion

• After content of pull request is agreed, it is merged
Implementation in GitHub: Example

- **Blue sky discussion on Discuss**
  - Volunteers swarm and begin work via issue, telcon, etc.

- **Proposal issue on Conventions** – all discussion goes here

- **Linked pull request proposing text** – review of changes takes place here (diffs and artefacts available)

- After discussion is concluded, a successful proposal is merged for the next release

- Throughout: Guidelines in CONTRIBUTING.md should help you!
What could be improved?

• Process improvements:
  1. Clarify when netCDF sample data is needed
  2. Work packaging a release could be “shifted left”
  3. Streamline the “3 week rules”
• Technical issues based off #267:
  1. Easy to lose track a proposal’s status
  2. Hard to recognize what changed and why
  3. Reduction of technical debt via use of aliases – also potential for general cleanup
Process improvement: 1. Sample data

• Currently sample netCDF files are needed in all cases. This is not always enforced.
  • Specify when it is (not) needed?
Process improvement: 2. Packaging earlier

• When a release is created, the changelog is updated and authors are added to the Conventions. Should this be done as part of the Pull Request?
Process improvement: 3. The long silence

• The Rules specify 2 periods of 3 weeks silence each before a merge
  • Waiting for further contributions
  • Waiting for objections following the moderator’s summary
• This inserts 2 time-triggered events into the workflow
• It might be better to request the summary and require silence thereafter only
Technical improvements: 1. Proposal status
Technical improvements: 1. Proposal status

• Issues are treated as “cards” that transition status along columns

• This reveals at a glance where an issue is located in its lifecycle without having to parse the comment thread

• Cards are “pulled” along horizontal axis
Technical improvements: 2. Following changes 😞

- We recommend that netCDF files that follow these conventions indicate:
  - the string value "CF-1.8" is a string that contains "CF-1.8".
  - The Conventions version number contained in that string can be used to find the web-based versions of this document from the links http://conventions.ucar.edu/netcdf/conventions.web page.
  - Subsequent versions of the CF Conventions will not make invalid a compliant usage of this or earlier versions of the CF terms and forms.

Identification of Conventions

We recommend that netCDF files that follow these conventions indicate this by setting the NUG defined global attribute Conventions to the string value "CF-1.8 ".

- It is possible for a netCDF file to adhere to more than one set of conventions, even when there is no inheritance relationship among the conventions. In this case, all common metadata is used with consistent names.

The NetCDF Data Model: Groups
Technical improvements: 2. Following changes 😊

- Line flow
- Trailing whitespace
Technical improvements: 3. Maintainable links

• Recent PR adjusted the same link. 60 lines touched.

• An alias would have reduced that by a lot

• We have a lot of links to maintain!
Way forward

• Big bang?

• Adjust when touched?

• Leave as-is?