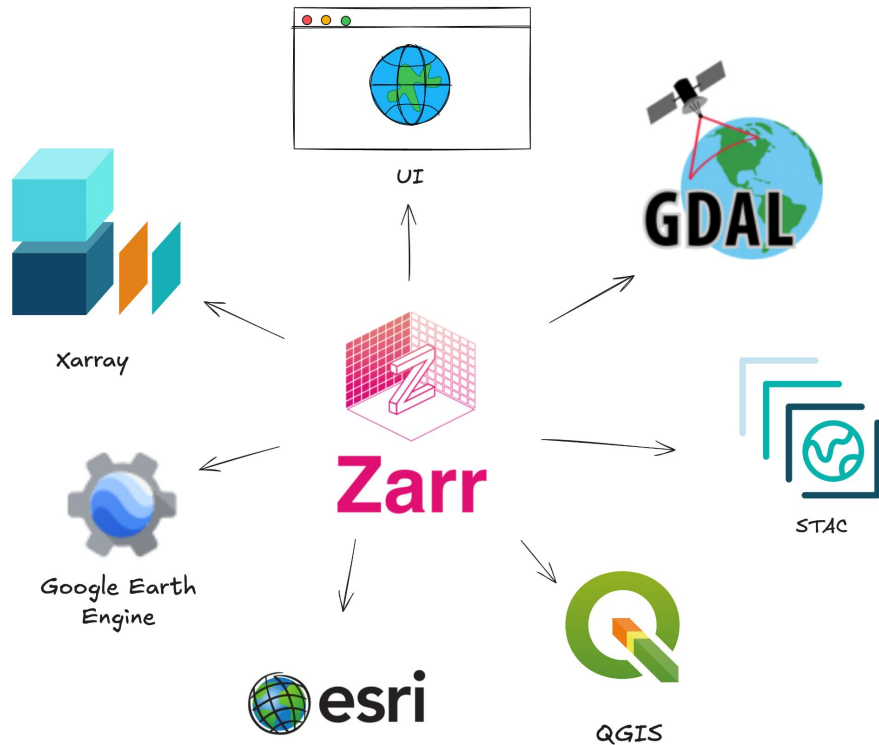


The GeoZarr Challenge

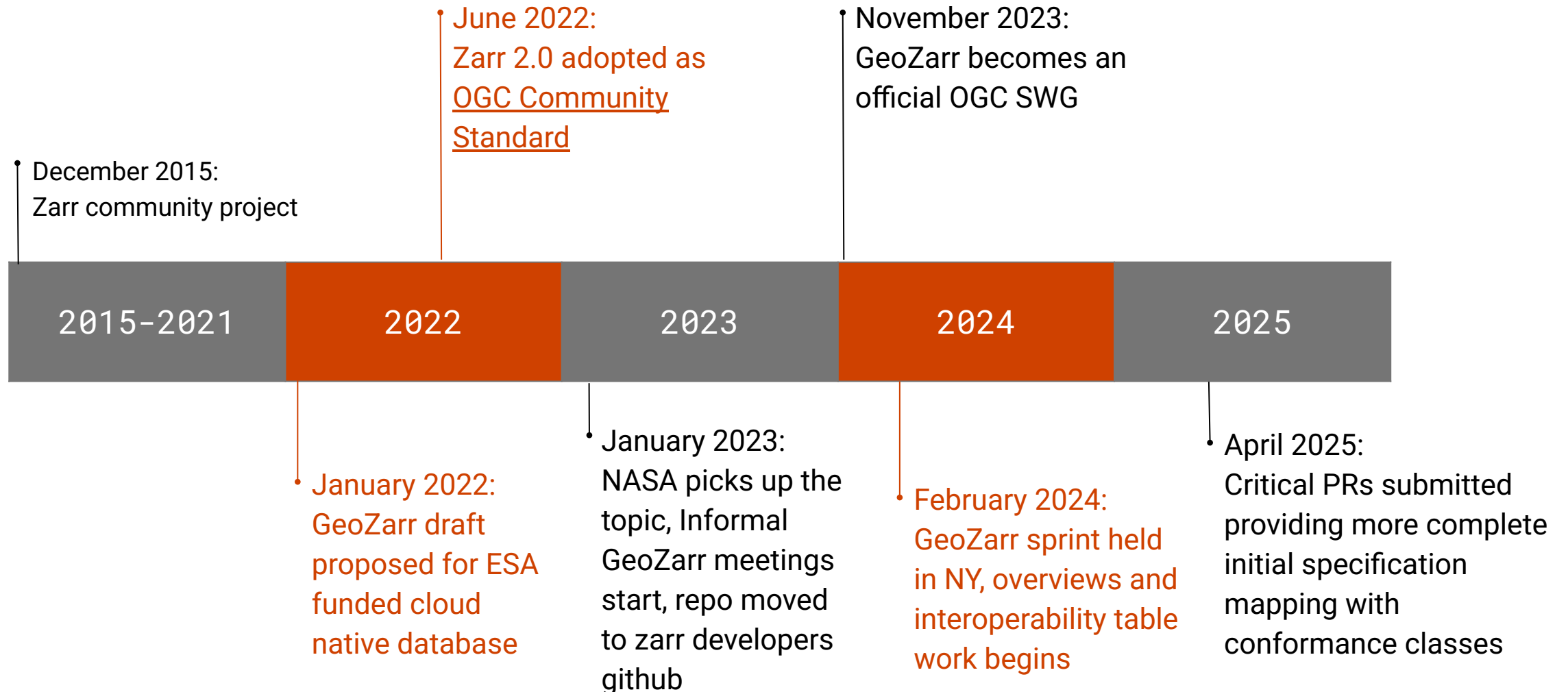
Standards building is hard, but we are in this together



Metadata conventions are essential for interoperability and preservation



The community has been building GeoZarr since 2023, based on an ESA-sponsored draft

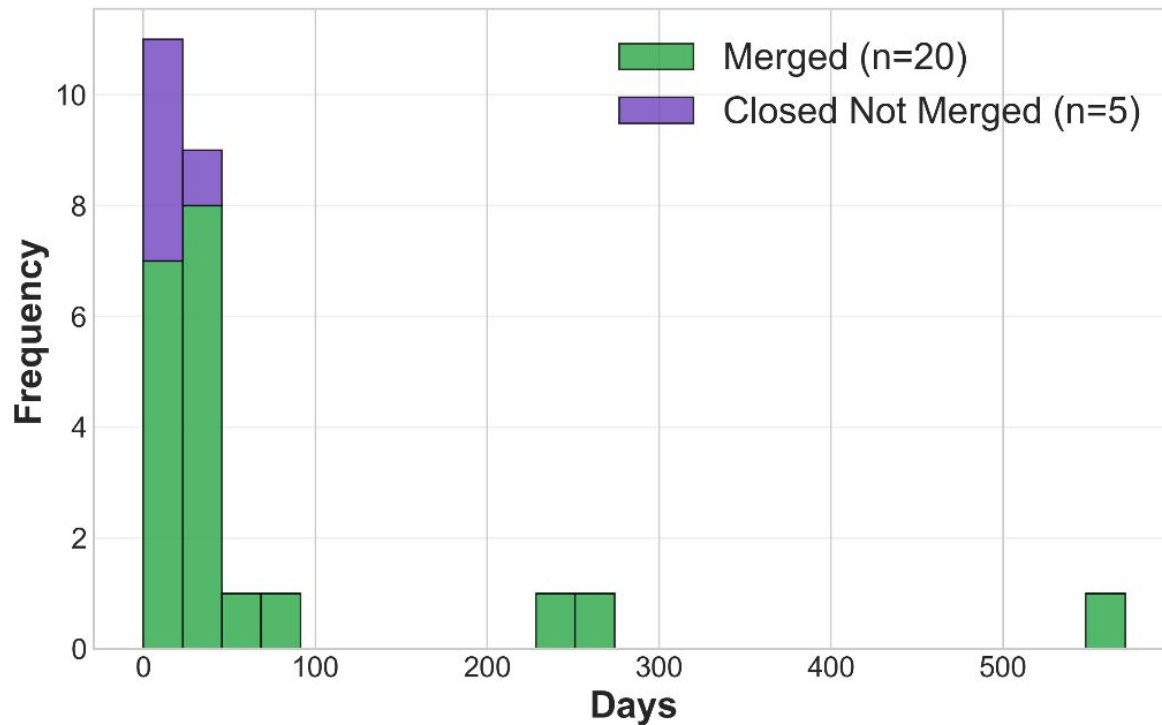




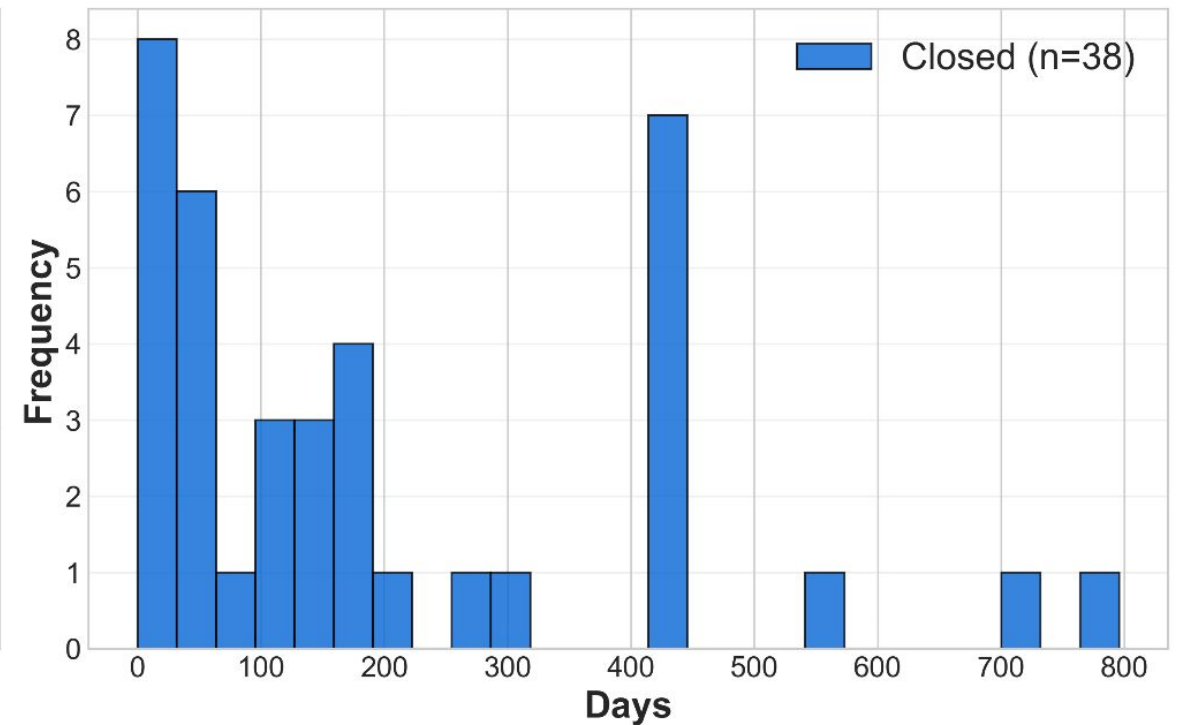
Metadata conventions are hard to build

GeoZarr Spec repository

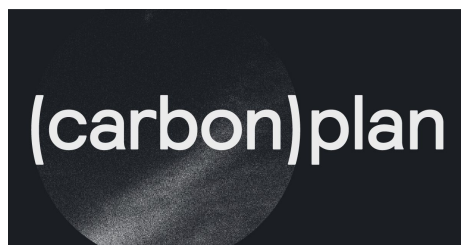
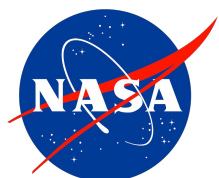
Time to close pull requests



Time to close issues



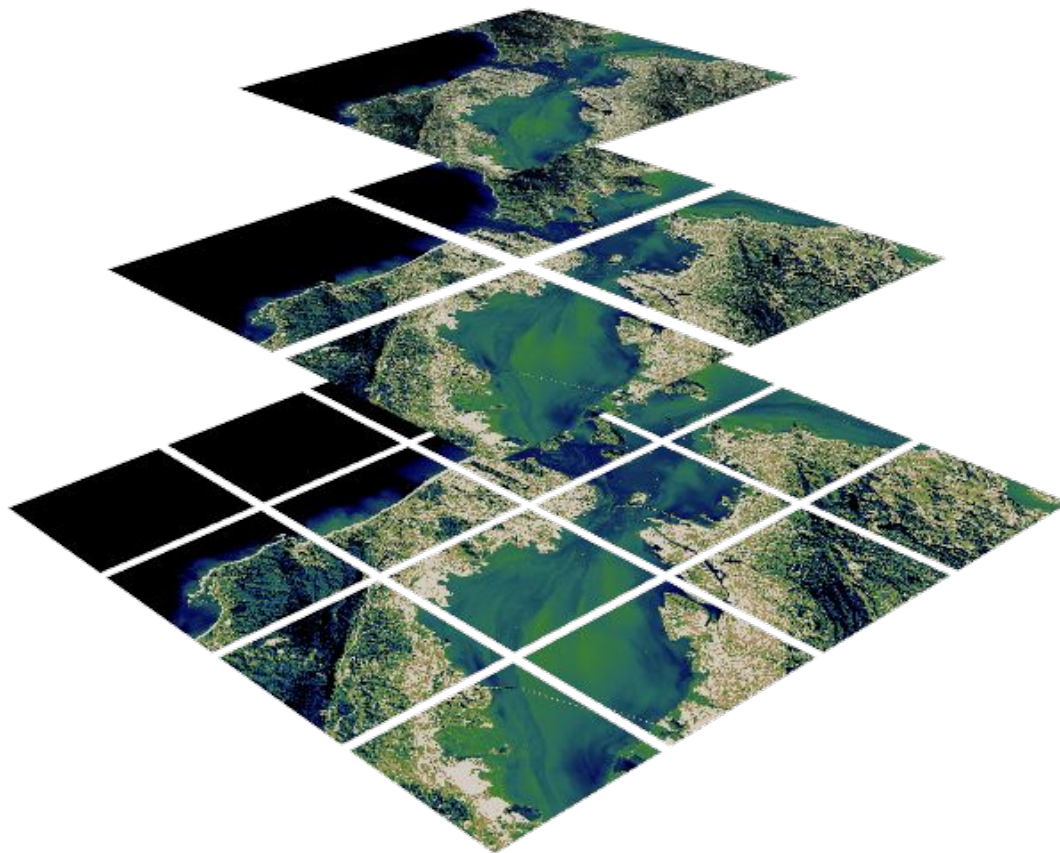
Metadata conventions are hard to build well



Zarr brings together many communities, leading to challenges melding approaches

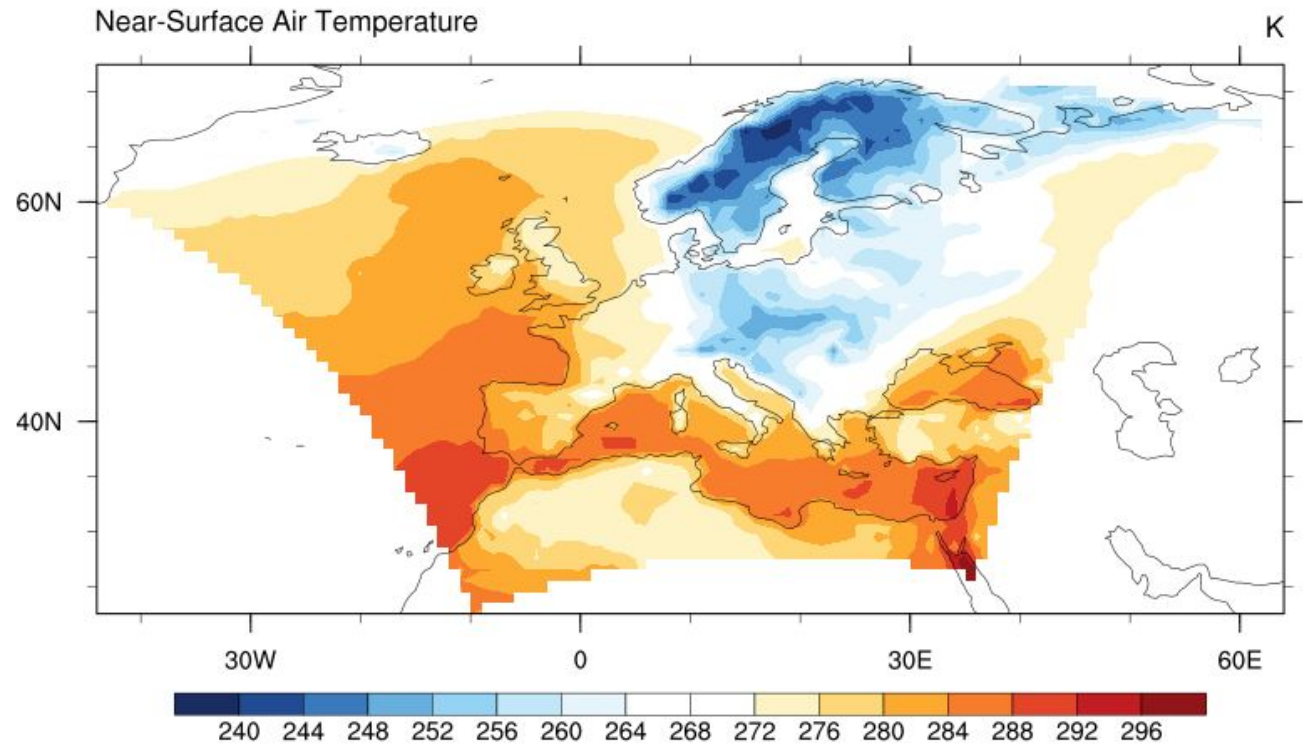


COG users rightly love implicit coordinates and overviews



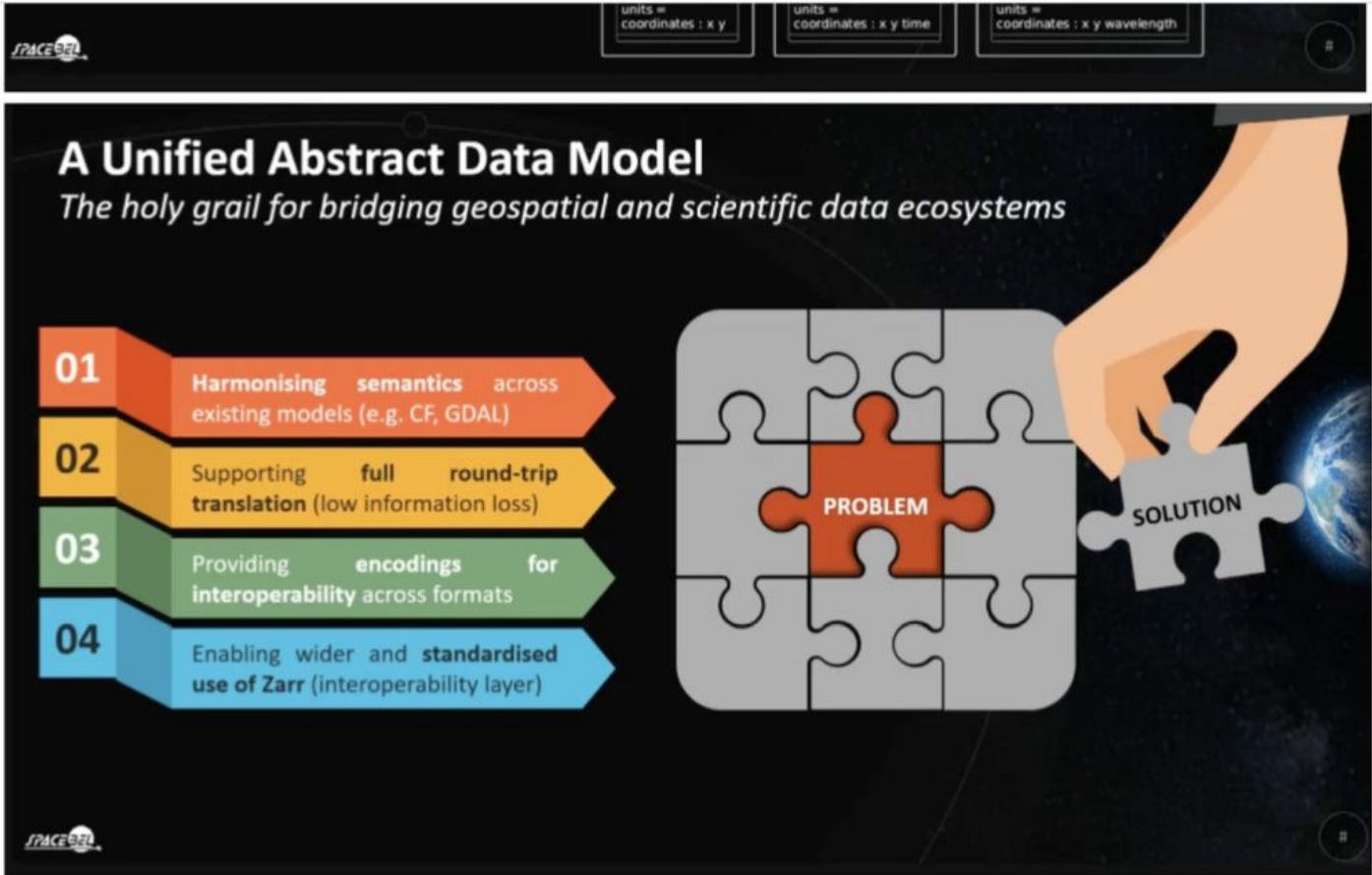


CF users rightly love comprehensive physical metadata, standardized vocabulary, and support for alternative grid structures



In April, Christophe proposed a composable system to meet all user needs

[Open](#) Call for Prototype/Implementation Owners for Different GeoZarr Profiles #63

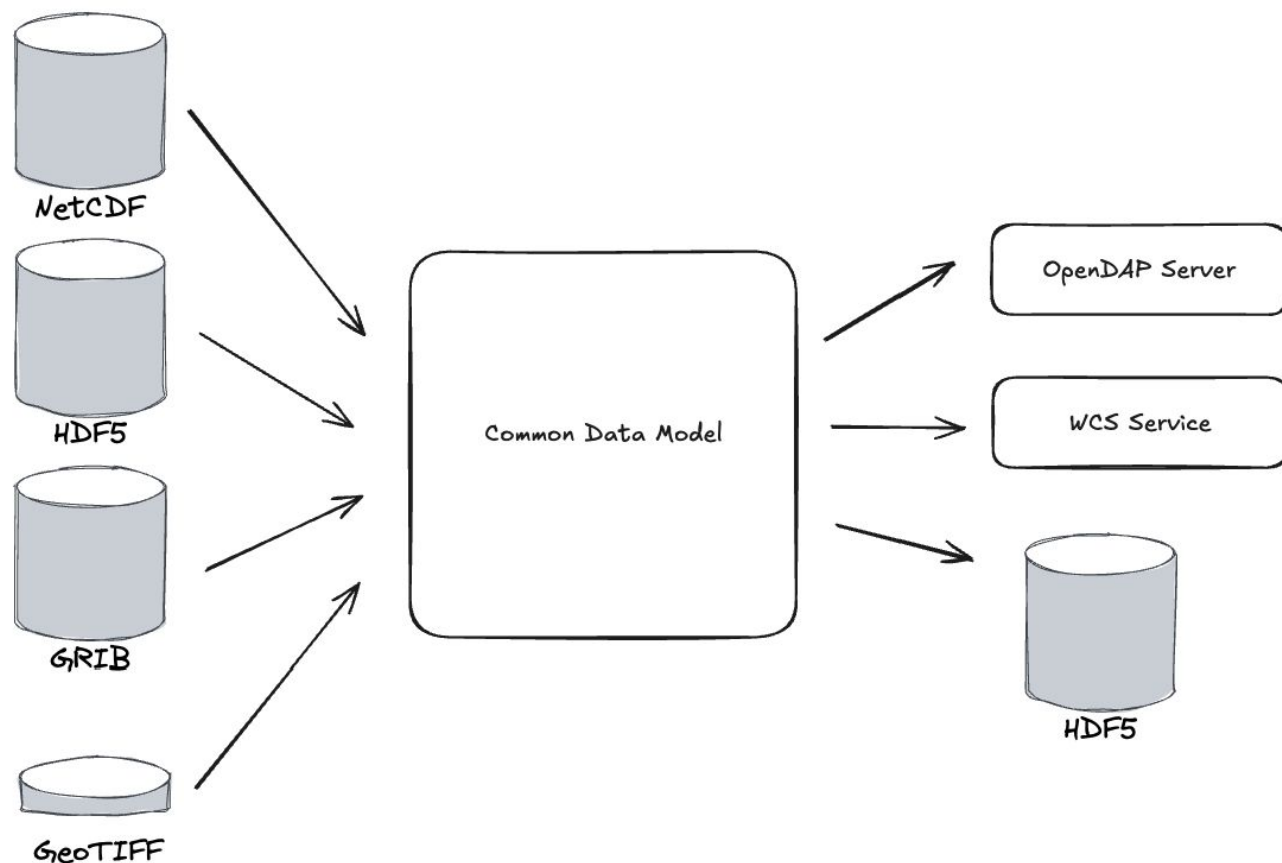


A Unified Abstract Data Model
The holy grail for bridging geospatial and scientific data ecosystems

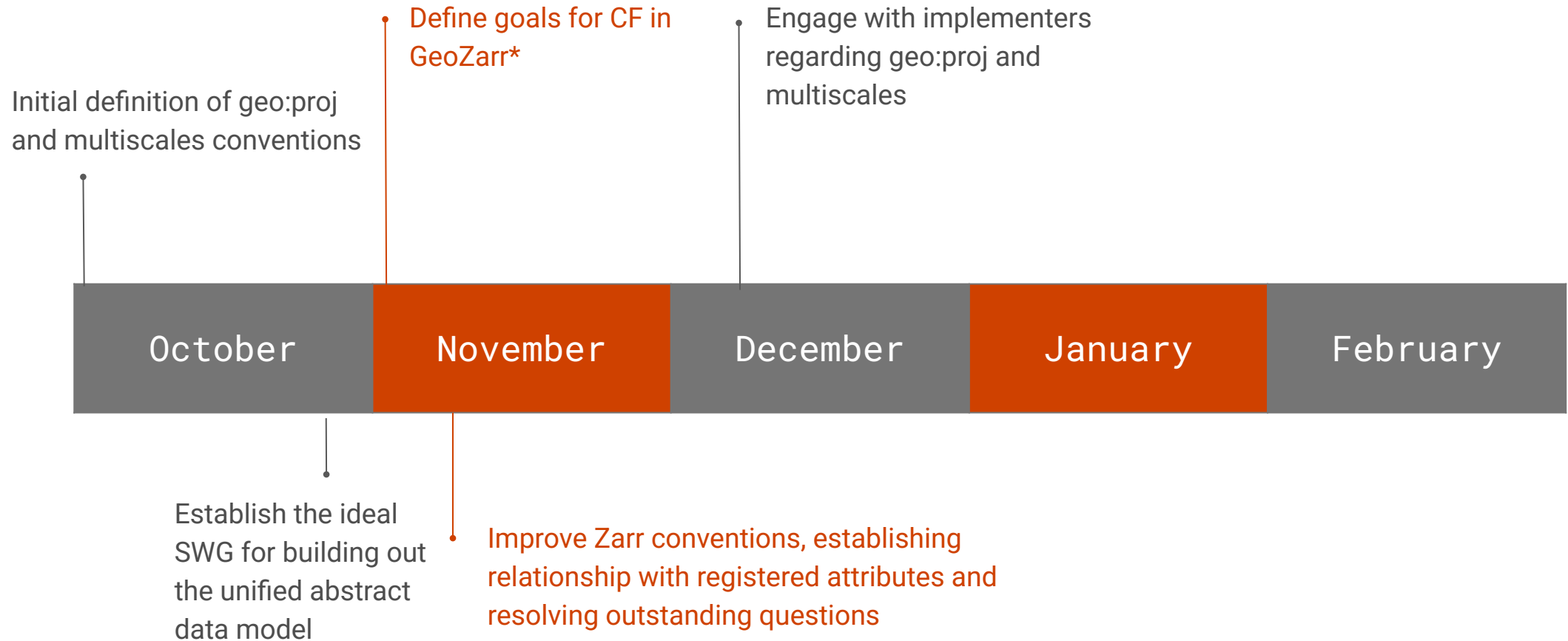
- 01** Harmonising semantics across existing models (e.g. CF, GDAL)
- 02** Supporting full round-trip translation (low information loss)
- 03** Providing encodings for interoperability across formats
- 04** Enabling wider and standardised use of Zarr (interoperability layer)

The diagram illustrates a puzzle where a hand is placing a piece labeled **SOLUTION** into a larger puzzle. One piece in the center of the puzzle is labeled **PROBLEM**. The puzzle is set against a background of a globe and a starry sky.

The unified abstract data model is broader than GeoZarr



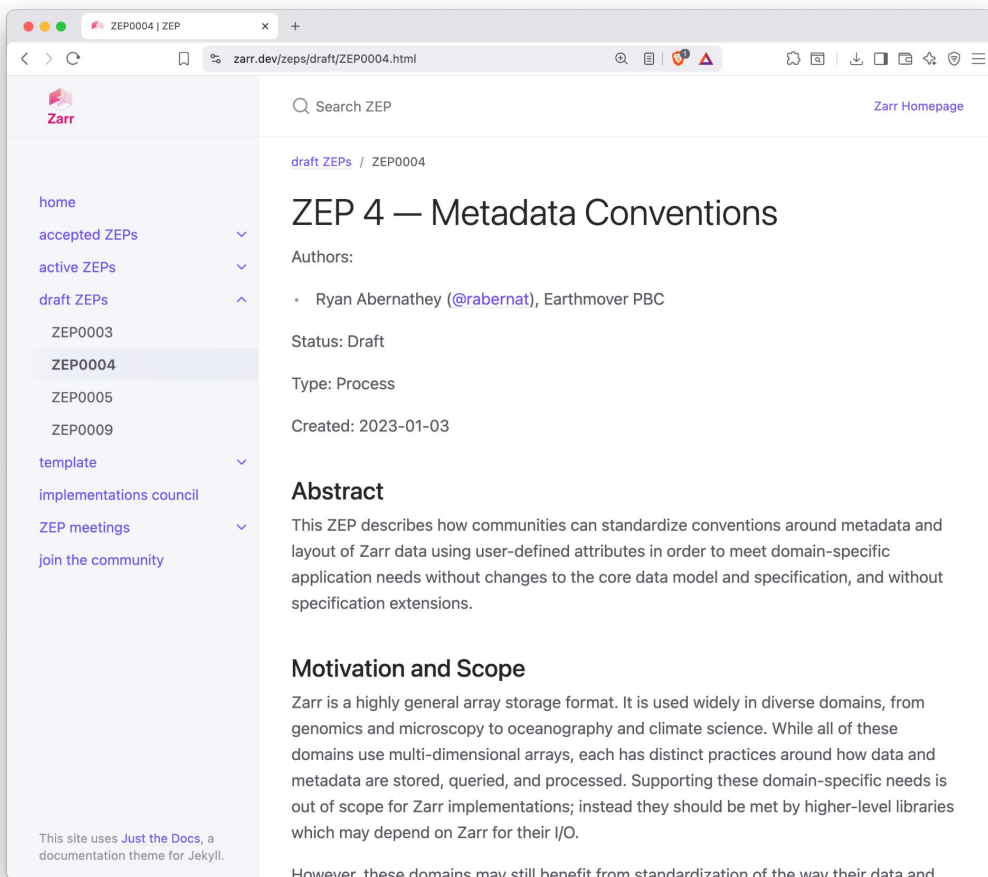
Roadmap*: GeoZarr V1 RC with geo:proj, multiscales, and CF



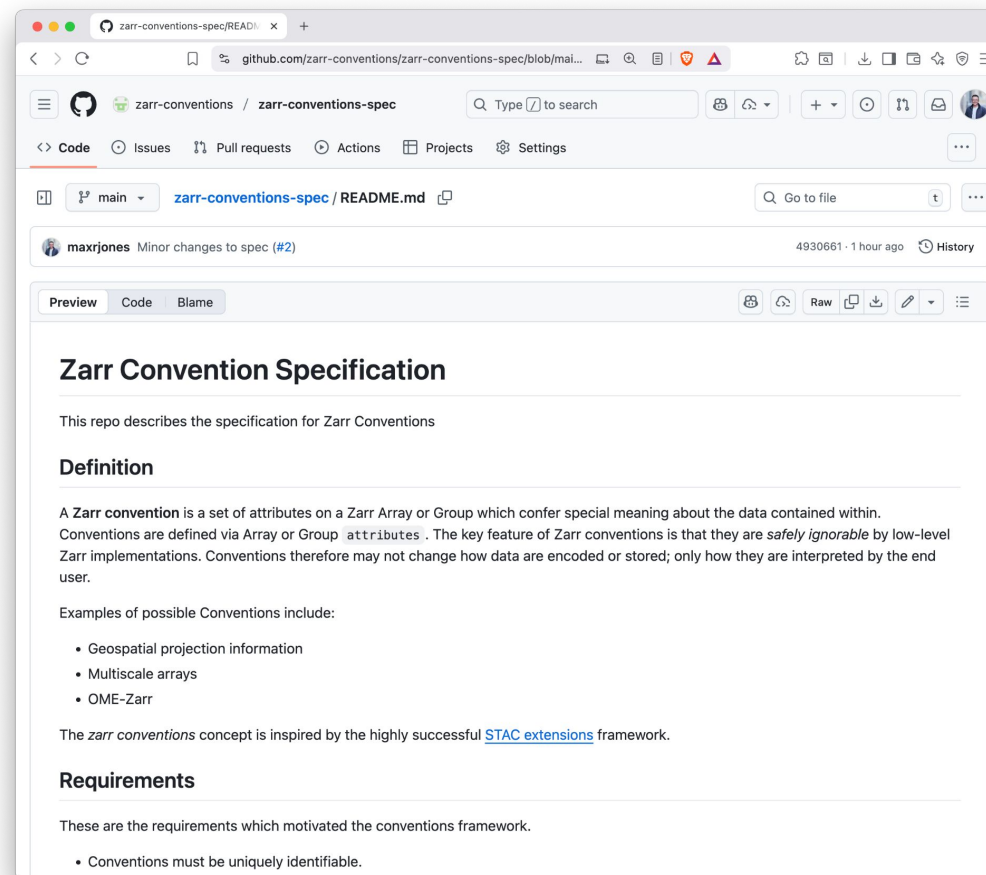
*documentation of existing practices, or building an optimal solution?

*suggestions building on [Christophe's draft](#)

GeoZarr should focus on the encoding portion of CDM



The screenshot shows a web browser displaying a draft page for ZEP 4 — Metadata Conventions on the Zarr website. The page has a sidebar with navigation links: home, accepted ZEPs, active ZEPs, draft ZEPs (selected), ZEP0003, ZEP0004, ZEP0005, ZEP0009, template, implementations council, ZEP meetings, and join the community. The main content area includes a search bar, a link to Zarr Homepage, and a list of draft ZEPs. The selected ZEP 4 page shows the title "ZEP 4 — Metadata Conventions", authors (Ryan Abernathey (@rabernat), Earthmover PBC), status (Draft), type (Process), and creation date (2023-01-03). The abstract states: "This ZEP describes how communities can standardize conventions around metadata and layout of Zarr data using user-defined attributes in order to meet domain-specific application needs without changes to the core data model and specification, and without specification extensions." The motivation and scope section explains that Zarr is a highly general array storage format used in diverse domains, and that domain-specific needs are out of scope for Zarr implementations, instead they should be met by higher-level libraries which may depend on Zarr for their I/O. However, these domains may still benefit from standardization of the way their data and



The screenshot shows a web browser displaying the GitHub repository page for zarr-conventions-spec. The page includes a search bar, a list of repository files (main, zarr-conventions-spec / README.md), and a commit history table. The commit history table shows a commit by maxrjones titled "Minor changes to spec (#2)" with a timestamp of 4930661 · 1 hour ago. The commit message is "Minor changes to spec (#2)". The commit details section shows the commit message and a list of files changed. The README file is selected, showing the title "Zarr Convention Specification" and the content: "This repo describes the specification for Zarr Conventions". The definition section states: "A Zarr convention is a set of attributes on a Zarr Array or Group which confer special meaning about the data contained within. Conventions are defined via Array or Group attributes. The key feature of Zarr conventions is that they are safely ignorable by low-level Zarr implementations. Conventions therefore may not change how data are encoded or stored; only how they are interpreted by the end user." Examples of possible Conventions include: Geospatial projection information, Multiscale arrays, and OME-Zarr. The README also mentions that the zarr conventions concept is inspired by the highly successful STAC extensions framework. The requirements section states: "These are the requirements which motivated the conventions framework." and lists: "Conventions must be uniquely identifiable."