

Chunked, Compressed, & Cloud-native
N-dimensional arrays

CSCCE Open Source Tools Trials

Managing technical documentation using GitHub

Sanket Verma

Community Manager @ Zarr



@MSanKeys963

— — —

What I'll be talking?

- 👉 How to setup GitHub repositories for docs
- 👉 What is the Zarr Community?
- 👉 How we use Git & GitHub to manage the Zarr Specification?
- 👉 Peak into the ZEP Process

How to setup GitHub repos for docs?

ReadtheDocs Template

GitHub Repository



<https://github.com/readthedocs/tutorial-template>

readthedocs / tutorial-template Public template

Watch 15 Fork 1.1k Starred 257

Code Pull requests 2 Actions Security Insights

main 2 branches 0 tags Go to file Add file Code Use this template

humitos Merge pull request #53 from readthedocs/humitos/explicit-sphinx-conf ... f594c40 last month 5 commits

docs Use a YAML file and Python requirements 2 months ago

.gitignore Basic structure 2 years ago

.readthedocs.yaml Use a explicit `sphinx.configuration` last month

README.rst Basic structure 2 years ago

lumache.py Basic structure 2 years ago

pyproject.toml Basic structure 2 years ago

README.rst

Template for the Read the Docs tutorial

This GitHub template includes fictional Python library with some basic Sphinx docs.

Read the tutorial here:

<https://docs.readthedocs.io/en/stable/tutorial/>

About

Template for the Read the Docs tutorial

docs.readthedocs.io/en/stable/tutorial/

Readme

Activity

257 stars

15 watching

1.1k forks

Report repository

Contributors 3

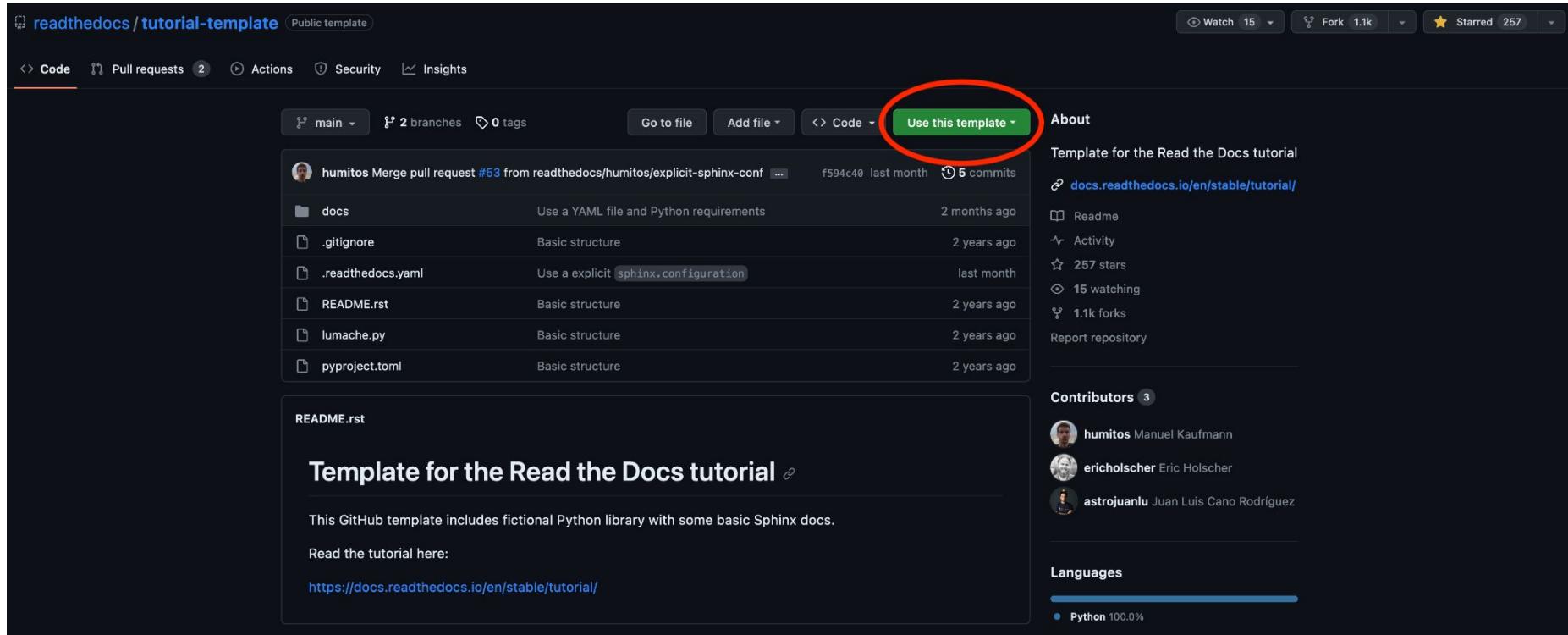
humitos Manuel Kauffmann

ericholscher Eric Holscher

astrojuanlu Juan Luis Cano Rodríguez

Languages

Python 100.0%



readthedocs / **tutorial-template** Public template

Watch 15 Fork 1.1k Starred 257

Code Pull requests 2 Actions Security Insights

main 2 branches 0 tags Go to file Add file Code Use this template ▾

humitos Merge pull request #53 from readthedocs/humitos/explicit-sphinx-conf ... f594c40

Create a new repository Open in a codespace

docs Use a YAML file and Python requirements 2 months ago

.gitignore Basic structure 2 years ago

.readthedocs.yaml Use a explicit `sphinx.configuration` last month

README.rst Basic structure 2 years ago

lumache.py Basic structure 2 years ago

pyproject.toml Basic structure 2 years ago

README.rst

Template for the Read the Docs tutorial

This GitHub template includes fictional Python library with some basic Sphinx docs.

Read the tutorial here:

<https://docs.readthedocs.io/en/stable/tutorial/>

About

Template for the Read the Docs tutorial

docs.readthedocs.io/en/stable/tutorial/

Readme

Activity

257 stars

15 watching

1.1k forks

Report repository

Contributors 3

humitos Manuel Kaufmann

ericholscher Eric Holscher

astrojuanlu Juan Luis Cano Rodríguez

Languages

Python 100.0%

Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository](#).

Required fields are marked with an asterisk ().*

Repository template



Start your repository with a template repository's contents.

Include all branches

Copy all branches from readthedocs/tutorial-template and not just the default branch.

Owner *



Repository name *

cscce_os_tools_trial

ccce_os_tools_trial is available.

Great repository names are short and memorable. Need inspiration? How about [fantastic-octo-funicular](#) ?

Description (optional)

This is a test repo for docs

Public

Anyone on the internet can see this repository. You choose who can commit.

Private

You choose who can see and commit to this repository.

You are creating a public repository in your personal account.

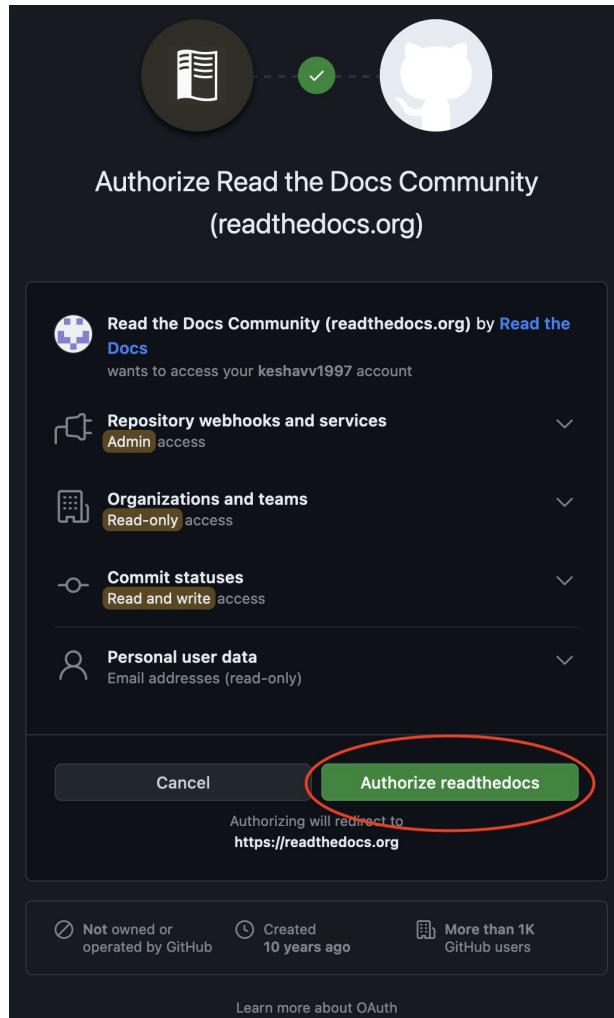
Create repository

Sign up

Using GitHub account



<https://readthedocs.org/accounts/signup/>



ReadtheDocs Dashboard



Ready to share your documentation?

You don't have any projects yet, but you can start building documentation by importing one. Not sure how to start documenting your project? Check out the [Tutorial](#) to learn how.

In a hurry? [Fork our template project](#) on Github, and [import it](#).

[Import a Project](#)

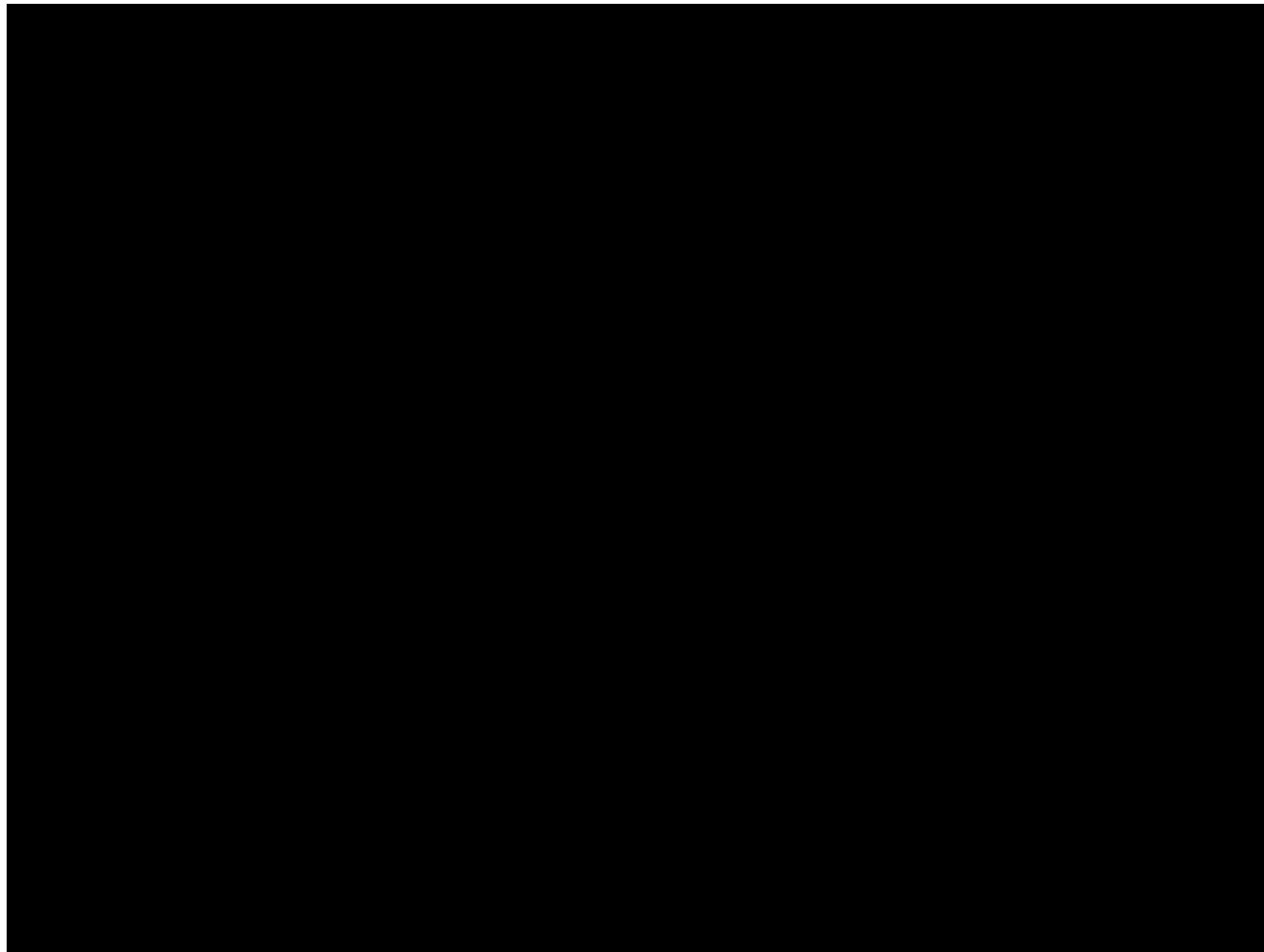
 [Search all docs](#)

Thanks!

Your [support](#) of Read the Docs helps make the site better each and every month.

Learn More

Check out the [documentation](#) for [Read the Docs](#). It contains lots of information about how to get the most out of RTD.



[Overview](#)[Downloads](#)[Search](#)[Builds](#)[Versions](#)[Admin](#)

Your documentation is ready to use

Your documentation has been built. Ensure your documentation is kept up to date with every commit to your repository, by [setting up a webhook](#).

[View your documentation](#)

Versions

latest

Public

[Edit](#)

Build a version

latest

[Build version](#)**Repository**https://github.com/keshavv1997/cscce_os_tools_trials.git**Project Slug**

cscce-os-tools-trials

Last Built12 minutes ago passed**Maintainers****Badge**docs passing ⓘ**Tags**

Project has no tags. Add some in your [project settings](#).

Short URLs

cscce-os-tools-trials.readthedocs.io
cscce-os-tools-trials.rtfd.io

Projects >

cscce_os_tools_trials

[View Docs](#)[Overview](#)[Downloads](#)[Search](#)[Builds](#)[Versions](#)[Admin](#)

Recent Builds

Build Version:

latest

Passed version latest (html)

14 minutes ago

[Overview](#)[Downloads](#)[Search](#)[Builds](#)[Versions](#)[Admin](#)**Build #22177601**latest (24daf0d68aa030c35a1766b1a86caf04e95d0994)

Completed Oct. 10, 2023. 10:29 a.m.

Build took 59 seconds

[View docs](#)[View raw](#)**Build completed**

```
git clone --depth 1 https://github.com/keshavv1997/cscce_os_tools_trials.git .
git fetch origin --force --prune --prune-tags --depth 50 refs/heads/main:refs/remotes/origin/main
git checkout --force origin/main
git clean -d -f -f
cat .readthedocs.yaml
asdf global python 3.10.12
python -m virtualenv $READTHEDOCS_VIRTUALENV_PATH
python -m pip install --upgrade --no-cache-dir pip setuptools
python -m pip install --upgrade --no-cache-dir sphinx readthedocs-sphinx-ext
python -m pip install --exists-action=w --no-cache-dir -r docs/requirements.txt
cat docs/source/conf.py
python -m sphinx -T -E -b html -d _build/doctrees -D language=en . $READTHEDOCS_OUTPUT/html
```

Check the docs



<https://cscce-os-tools-trials.readthedocs.io/en/latest/>

Welcome to CSCCE OS Tools Trials' documentation!

The [Center for Scientific Collaboration and Community Engagement \(CSCCE\)](#) is a research and training center to support and study the emerging field of scientific community engagement.

This is test documentation created for the presentation.

Check the blog post here: <https://www.cscce.org/2023/09/21/github-and-bitergia-for-supporting-research-and-developer-communities-cscce-open-source-tools-trial-4/>

© Copyright 2021, Graziella. Revision 34b71d1a.

Built with [Sphinx](#) using a [theme](#) provided by [Read the Docs](#).

Zarr Community

ZARR OSS



AND SPECIFICATION

USERS & DEVS



CONTRIBUTORS & MAINTAINERS

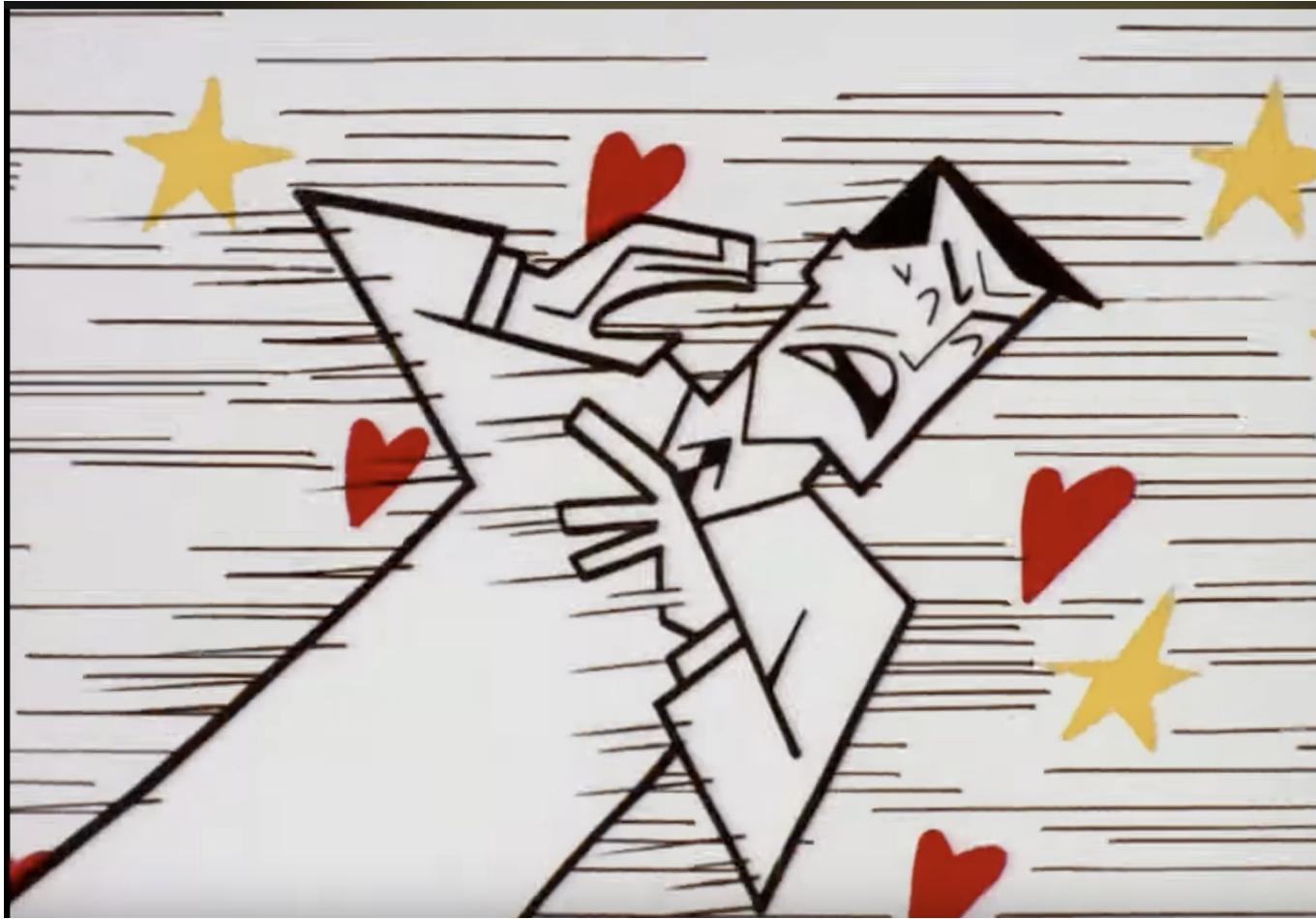


GOVERNANCE AND COUNCIL

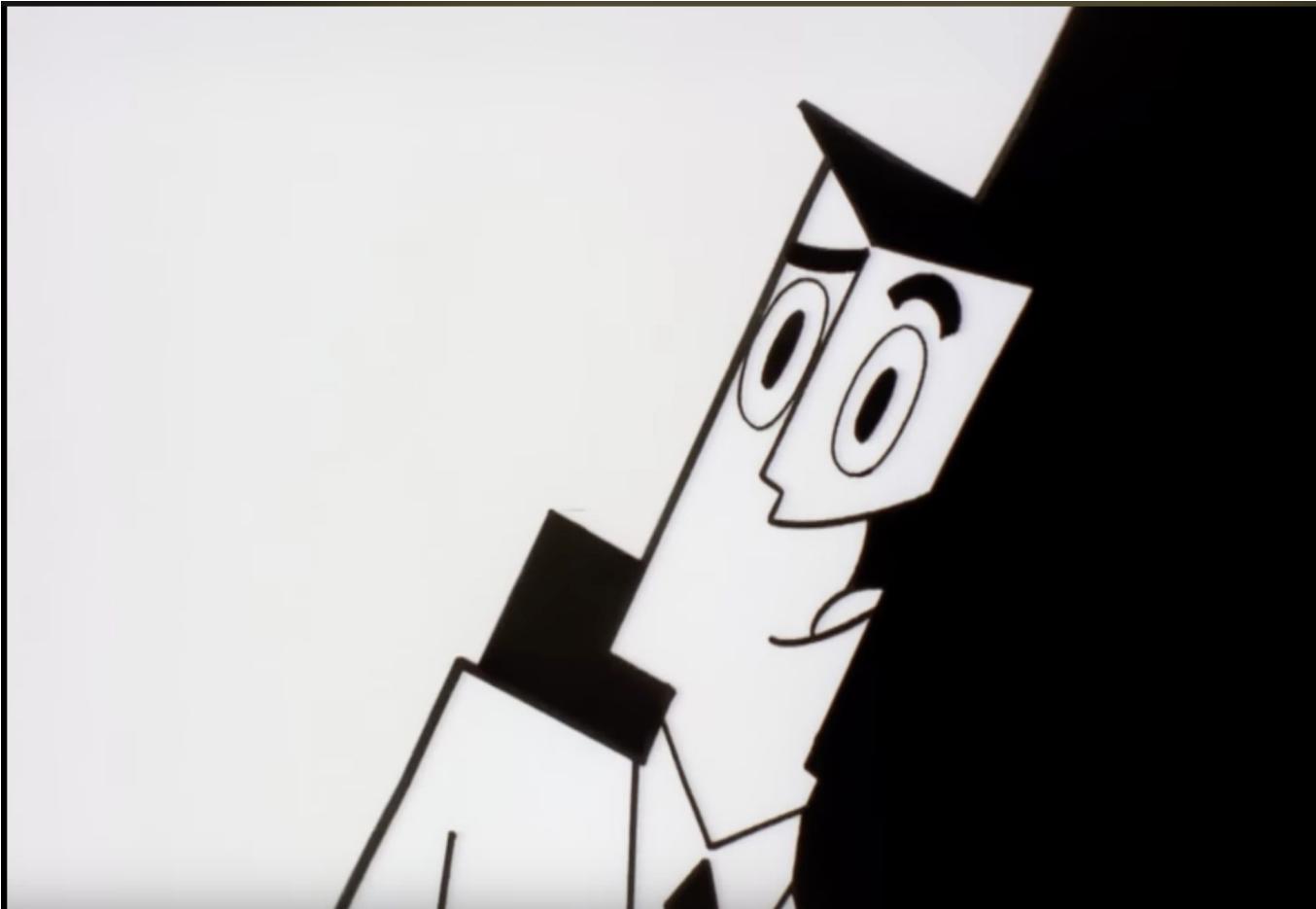


EASY TO USE

HACKABLE AND AGNOSTIC









How we're using Git & GitHub?

GitHub offers a range of features!

- ~ Issues to raise a concern/feature/discuss/track progress
- ~ Pull Requests (PR) to submit changes to code/documentation
- ~ Project boards to organise workflow, create roadmaps etc.
- ~ GH Discussions for polls, QnA, general discussions in threaded style

Search or jump to... Issues Pull requests Issues Codespaces Marketplace Explore

zarr-developers/zarr-specs Public

Issues 72 Pull requests 6 Discussions Actions Projects 1 Wiki Security Insights

Code Issues Pull requests Discussions Actions Projects Wiki Security Insights

main 5 branches 0 tags

Go to file Add file Code

joshmoore Merge pull request #261 from MSanKeys963/update_V3_status ... 97453eb on Aug 24 437 commits

docs .gitignore .readthedocs.yaml LICENSE.txt README.md

Update V3 status and minor fixes. 4 years ago

File system store spec (#47) 5 months ago

Update .readthedocs.yaml 4 years ago

Add CC-BY 3.0 license file (#36) 2 months ago

Update V3 status and minor fixes. 2 months ago

README.md

Zarr Specification

Zarr core protocol for storage and retrieval of N-dimensional typed arrays



Zarr

For the v1 and v2 specs, please see <https://github.com/zarr-developers/zarr-python/tree/main/docs/spec>.

The rendered docs of the main branch are available at <https://zarr-specs.readthedocs.io>

Edit Pins Unwatch 29 Fork 23 Starred 65

About

Zarr core protocol for storage and retrieval of N-dimensional typed arrays

zarr-specs.readthedocs.io

storage arrays zarr

Readme View license Code of conduct Activity 65 stars 29 watching 23 forks

Report repository

Releases

No releases published

Packages

No packages published

Contributors 16



Issue Tracker



zarr-developers / zarr-specs Public

Code Issues 72 Pull requests 6 Discussions Actions Projects 1 Wiki Security Insights

ZEP0002 Review
#254 opened on Jul 27 by normanrz
Open 20

Filters is:issue is:open Labels 17 Milestones 0 New issue

Author	Label	Projects	Milestones	Assignee	Sort
72 Open	66 Closed				
Byte ranges need to permit reading from the end #266 opened 2 weeks ago by LDeakin					
ZEP0001 status clarification, not "Final"? #265 opened 3 weeks ago by mkitti					
ZEP0002 Review #254 opened on Jul 27 by normanrz 20					
Adding sparse array support #245 opened on May 25 by elyall 12					
Position-dependent fill value #243 opened on May 19 by jbms					
Information available to sharding codec #241 opened on May 16 by cbarnes 2					
Can fallback data types have fallbacks? #239 opened on May 10 by cbarnes 2					
Extension for attribute datatypes? protocol-extension 6					
sharding as a codec rather than array storage transformer #220 opened on Mar 9 by jbms 4					

Information available to sharding codec #241

[New issue](#)[Open](#)

clbarnes opened this issue on May 16 · 2 comments



clbarnes commented on May 16 · edited

Contributor

...

Currently the sharding codec spec reads that "each integer must be divisible by the chunk_shape of the array as defined in the chunk_grid [Array metadata](#)". However, array->array codecs can change the shape of the requested chunk (e.g. the transpose codec permuting dimensions); in this case, the sharding codec's chunk shape should correspond to the permuted shape rather than the array's shape.

The sharding codec will also need to duplicate empty-chunk logic from the array class, which also means it will need access to the array's fill value. However, as array->array codecs could feasibly change values in the array (could they also change the data type?), the fill value may also need these transformations applied to it.

This probably all falls under the `codec.compute_encoded_representation_type` method in the spec.



normanrz commented on May 17

Member

...

In zarrita I pass a subset of the array metadata to the codecs: <https://github.com/scalableminds/zarrita/blob/v3/zarrita/array.py#L101-L110>



jbms commented on May 18

Member

...

As far as the issue of the chunk shape being affected by prior codecs (also applies if there is nested sharding), you are correct. I tried to fix the wording to address that issue in this (otherwise unrelated) PR <https://github.com/zarr-developers/zarr-specs/pull/237/files>

Assignees



Labels



None yet

Projects



None yet

Milestone



No milestone

Development

No branches or pull requests

Customize

[Unsubscribe](#)

You're receiving notifications because you're watching this repository.

3 participants

[Convert to discussion](#)

Pull Requests



Screenshot of a GitHub repository page showing the Pull Requests tab. The repository is `zarr-developers/zarr-specs` (Public). The Pull Requests section displays 6 Open pull requests and 116 Closed pull requests. The Open pull requests are listed below:

PR #	Title	Author	Opened On	Review Status	Comments
#264	Require transpose codec <code>order</code> option to be specified as an explicit permutation	jbms	last month	Review required	3
#263	Rename endian codec to bytes	normanrz	Aug 21	Approved	18
#262	ZEP0004 Review - Zarr Conventions	MSanKeys963	Aug 18	Draft	2
#257	Add bfloat16 data type	jbms	Aug 5	Review required	1
#256	Add zstd codec	jbms	Aug 1	Review required	4
#205	Extension proposal for zarr accumulation	hailiangzhang	Feb 15	Review required	10

Code Issues 72 Pull requests 6 Discussions Actions Projects 1 Wiki Security Insights

ZEP0004 Review - Zarr Conventions #262

Edit < Code

Draft MSanKeys963 wants to merge 1 commit into zarr-developers:main from MSanKeys963:ZEP0004



Conversation 2

Commits 1

Checks 0

Files changed 3

+29 -0



MSanKeys963 commented on Aug 18

Member ...

Hi everyone! 🐦

I did some preliminary work for [ZEP0004](#) review, as mentioned [here](#).

@rabenat, please have a look and let us know your thoughts. Thanks!



Add conventions for ZEP0004 review ...

ab8d4a1



rabenat commented on Aug 18

Member ...

Thanks so much [@MSanKeys963](#) for getting this started. It's a perfect place to start. Here's what I will try to do over the next few days

- Port over some of the text from ZEP4 to the specs website
- Create a convention template
- Create at least one full example convention

In the meantime, we can use this PR to continue the discussion started in <https://github.com/zarr-developers/zeps/pull/28/files> where ZEP4 was first proposed.

@ivirshup I know you have lots of ideas here, and you have been very patient as this ZEP has moved forward very slowly. 🙏 I'd love to hear more about your use cases for conventions and data and the other projects you're involved with.



1

Reviewers

Suggestions



Request

At least 1 approving review is required to merge this pull request.

Assignees

No one—assign yourself

Labels

None yet

Projects

None yet

Milestone

No milestone

Development

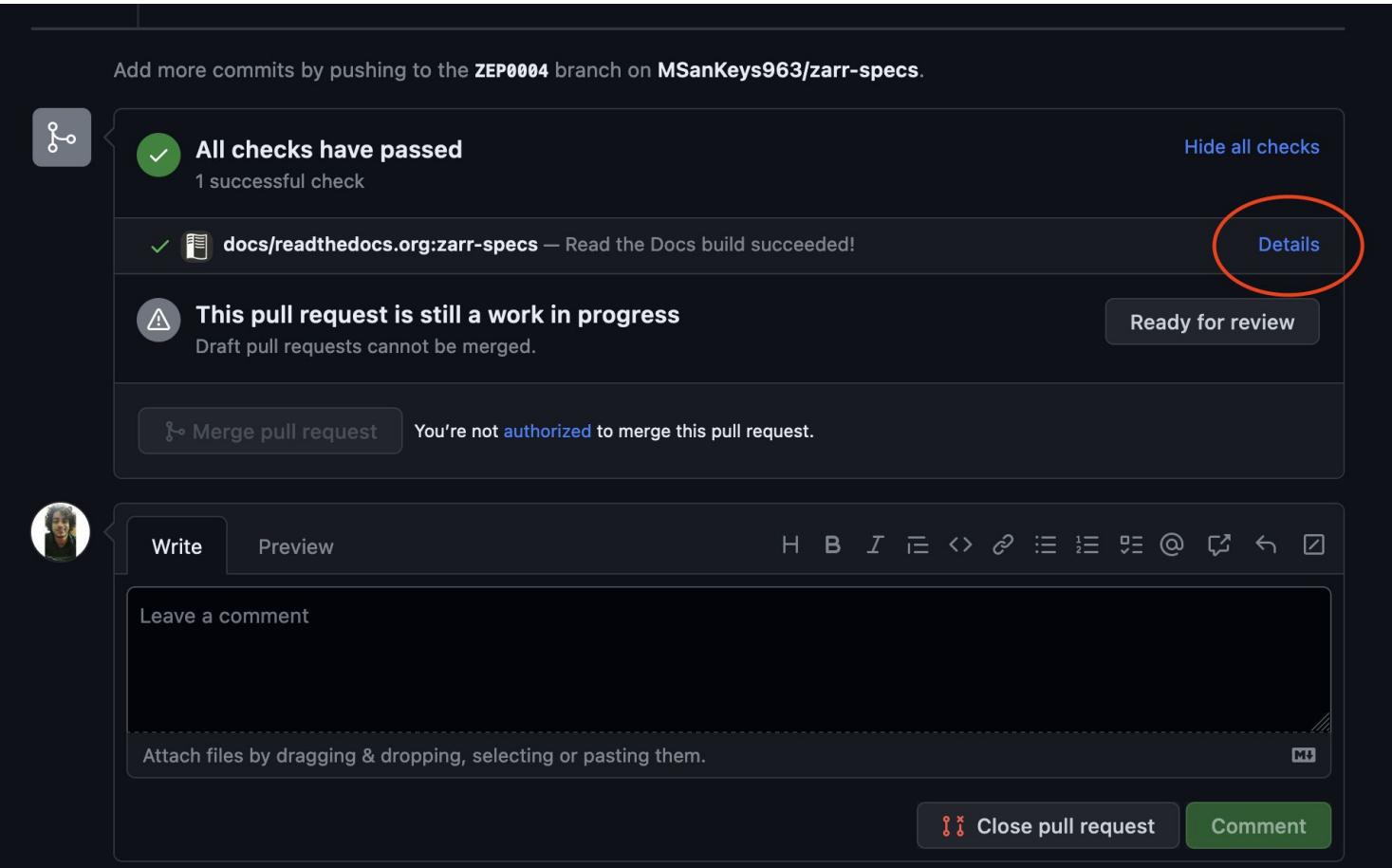
Successfully merging this pull request may close these issues.

None yet

Notifications

Customize

Add more commits by pushing to the [ZEP0004](#) branch on [MSanKeys963/zarr-specs](#).



⚠ Warning

This page was created from a pull request (#262).

Specifications

v3

Core

Data Types

Codecs

Stores

Array Storage Transformers

Conventions

Conventions

v2

Zarr spec v2

v1

Zarr spec v1

Fixes formatting for the "Core data types" table #166

[Code](#)

Merged

joshmoore merged 1 commit into [main](#) from [core-data-types-table-formatting](#) on Nov 17, 2022

Conversation 0

Commits 1

Checks 0

Files changed 1



jstriebel commented on Nov 16, 2022

Member

...

Fixes formatting for the "Core data types" table, which was not rendered previously due to the incorrect number of columns in the last row. (Follow-up from [#155](#))



Fixes formatting for the "Core data types" table, which was not rendered previously due to the incorrect number of columns in the last row. (Follow-up from [#155](#))

Verified



0ff20fe

jstriebel requested review from **joshmoore** and **MSanKeys963** last year

jstriebel self-assigned this on Nov 16, 2022



MSanKeys963 approved these changes on Nov 16, 2022

jstriebel requested review from **rabernat** and **jakirkham** last year



joshmoore approved these changes on Nov 17, 2022

[View reviewed changes](#)

[View reviewed changes](#)

joshmoore merged commit [2af973a](#) into [main](#) on Nov 17, 2022
1 check passed

[View details](#)

[Revert](#)

+11 -12

Reviewers

MSanKeys963
 joshmoore
 rabernat
 jakirkham



Assignees

jstriebel



Labels

None yet



Projects

ZEP1
Status: Done



Milestone

No milestone



Development

Successfully merging this pull request may close these issues.

Project Board



⌚ Zarri-Python - Benchmarking and Performance

Project Board + New View

Filter by keyword or by field

Discard Save

Done 7 This has been completed

- perfcapture #14 Simple API for specifying each benchmark workload
- zarr-benchmark #1 Pick benchmarking tool
- perfcapture #6 Implement parameterised `Workload.run` & `Dataset.prepare` methods
- perfcapture #11 Maybe rename `Dataset.prepare` to `Dataset.create`
- perfcapture #5 Increment the Development status in `setup.cfg`
- zarr-benchmark #5 Implement dataset & workload based on presentation from meeting on Thu 28th Sept.
- zarr-benchmark #4 Copy files from a Python project skeleton (to enable CI etc.)

In Progress 2 This is actively being worked on

- perfcapture #15 What to measure during benchmarking?
- zarr-benchmark #6 Run Intel VTune against `SlowMemcpyWorkload`

Todo 10 This item hasn't been started

- perfcapture #4 Measure & record performance while benchmarks run!
- perfcapture #12 Consider defining workloads & benchmarks in yaml to be consistent with `zarr_implementations`
- perfcapture #7 Document how to add a new workload and dataset
- perfcapture #10 Automated integration test which runs the cli against all examples
- perfcapture #9 Document terminology
- perfcapture #1 Automate building & publishing of docs
- perfcapture #8 Consider different names for project (instead of `perfcapture`)?

In review 0 This item is in review

- perfcapture #2

Unassigned 0 This item is not assigned yet

- perfcapture #3

In discuss 0 This item is being discussed

- perfcapture #13

+ Add item

Git lets you keep track of versioning and changes via commits 

Commits

main

- o Commits on Aug 24, 2023
 - Merge pull request #261 from MSanKeys963/update_V3_status ...
joshmoore committed on Aug 24 
Verified 97453eb ⚡ ↗ ↘
- o Commits on Aug 11, 2023
 - Update V3 status and minor fixes.
MSanKeys963 committed on Aug 11 ✓
d3da2b9 ⚡ ↗ ↘
- o Commits on Aug 10, 2023
 - Merge pull request #260 from zarr-developers/mark-zep1-accepted ...
joshmoore committed on Aug 10
Verified 8741619 ⚡ ↗ ↘
 - Mark ZEP 1 as accepted.
jstriebel committed on Aug 10 ✓
cccfb0d ⚡ ↗ ↘
- o Commits on Jul 27, 2023
 - Merge pull request #248 from jbms/remove-fallback-data-type ...
jbms committed on Jul 27
Verified f7c5534 ⚡ ↗ ↘
 - Remove fallback data types from specification ...
jbms committed on Jul 27 ✓
cc1d163 ⚡ ↗ ↘
 - Merge pull request #249 from jbms/remove-default-binary-representation ...
jbms committed on Jul 27
Verified f7f7a2e ⚡ ↗ ↘
 - Merge pull request #255 from normanrz/patch-2 ...
jbms committed on Jul 27
Verified f2db45d ⚡ ↗ ↘

sts 6 Discussions

Commits

main ▾

- Commits on Aug 24, 2023
- Commits on Aug 11, 2023

[zarr-developers/zarr-specs](#) on Aug 11

↳ **Update V3 status and minor fixes. #261**

Hi everyone. 🙌 Building this on [#260](#). The changes in this PR are as follows: After...

`zarr-developers:m... ← MSanKeys963:updat...`

 You commented on and opened this pull request

 Approved

Merge pull request [#261](#) from MSanKeys963/update_V3_status ...

 **joshmoore** committed on Aug 24

Update V3 status and minor fixes.

 **MSanKeys963** committed on Aug 11 ✓

Update V3 status and minor fixes. #261

Edit [Code](#)

Merged

joshmoore merged 1 commit into [zarr-developers:main](#) from [MSanKeys963:update_V3_status](#) on Aug 24

Conversation 4

Commits 1

Checks 0

Files changed 7

+19 -25



MSanKeys963 commented on Aug 11

Member [...](#)

Hi everyone. 🌟

Building this on #260.

The changes in this PR are as follows:

- After going through the recent changes introduced by [Mark ZEP 1 as accepted](#). #260, I realised the date for ZEP0001 acceptance was mentioned as 13th May. The actual date for acceptance is 15th May, when the last vote came in. Check [here](#) and [here](#).
- Also, we're mentioning [Zarr core development team](#) in the specification documents. This shows no team members, and this link throws a 404 to anyone outside our GitHub org. I suggest we remove the hyperlink until we make the teams public, which is visible to everyone (especially users outside the org).

Please have a look and let me know your thoughts.

CC: [@zarr-developers/zep1-editors](#) [@jhamman](#)



1

Update V3 status and minor fixes.

✓ d3da2b9



joshmoore commented on Aug 11

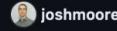
Member [...](#)

LGTM. Did you look into redirecting from the "draft" URL? (Sorry if that was handled elsewhere?)



2

Reviewers



joshmoore



Assignees

No one—assign yourself



Labels

None yet



Projects

None yet



Milestone

No milestone



Development

Successfully merging this pull request may close these issues.

None yet

Notifications

Customize

Unsubscribe

You're receiving notifications because you're watching this repository.

- o- Commits on Apr 27, 2019
 - allow zero d arrays**
alimanfoo committed on Apr 27, 2019 Verified ↗ a704813 🔗
 - o- Commits on Apr 25, 2019
 - minor spacing**
alimanfoo committed on Apr 25, 2019 Verified ↗ 58830bb 🔗
 - suggestions for further structure**
alimanfoo committed on Apr 25, 2019 Verified ↗ 31b937b 🔗
 - minor style edit**
alimanfoo committed on Apr 25, 2019 Verified ↗ e989269 🔗
 - added content on grid, memory layout, codecs, metadata, stores**
alimanfoo committed on Apr 25, 2019 Verified ↗ afc208a 🔗
 - o- Commits on Apr 24, 2019
 - fix confusing text re paths**
alimanfoo committed on Apr 24, 2019 Verified ↗ 831a3c1 🔗
 - fix title**
alimanfoo committed on Apr 24, 2019 Verified ↗ 4d27946 🔗
 - dimension lengths**
alimanfoo committed on Apr 24, 2019 Verified ↗ 7f54bf9 🔗
 - added some content on data types and chunks**
alimanfoo committed on Apr 24, 2019 Verified ↗ c9df31c 🔗
 - initial work on conceptual model**
alimanfoo committed on Apr 24, 2019 Verified ↗ e73f1a6 🔗
 - create skeleton for v3.0 core protocol spec**
alimanfoo committed on Apr 24, 2019 Verified ↗ 5a3421f 🔗
 - o- Commits on Jan 15, 2019
 - Start off the repo** ...
jakirkham committed on Jan 15, 2019 ↗ 00a05f0 🔗
- Newer Older

Zarr Specification

V1 → V2 → V3

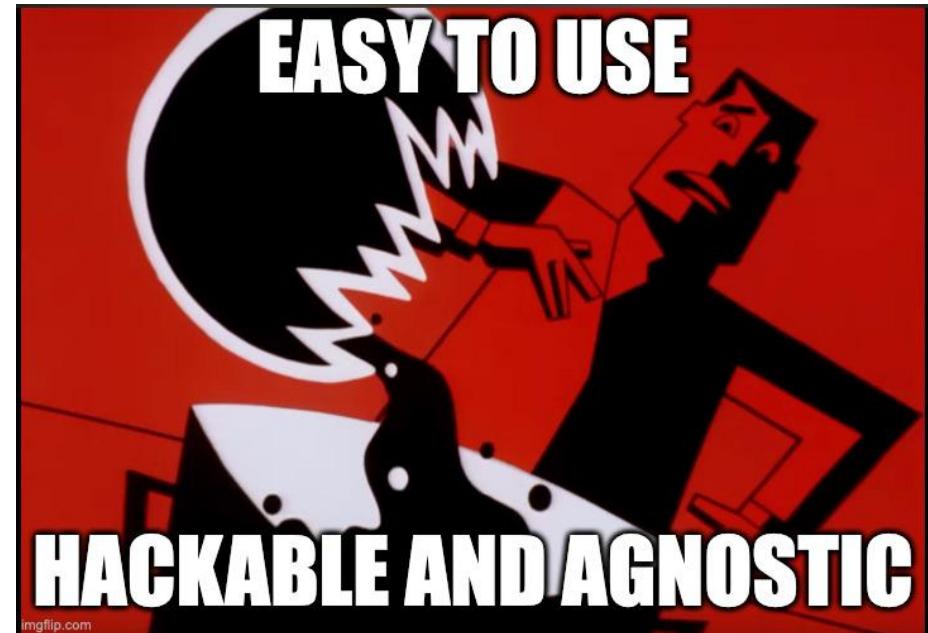


<https://zarr-specs.readthedocs.io/>

Peak into the ZEP Process

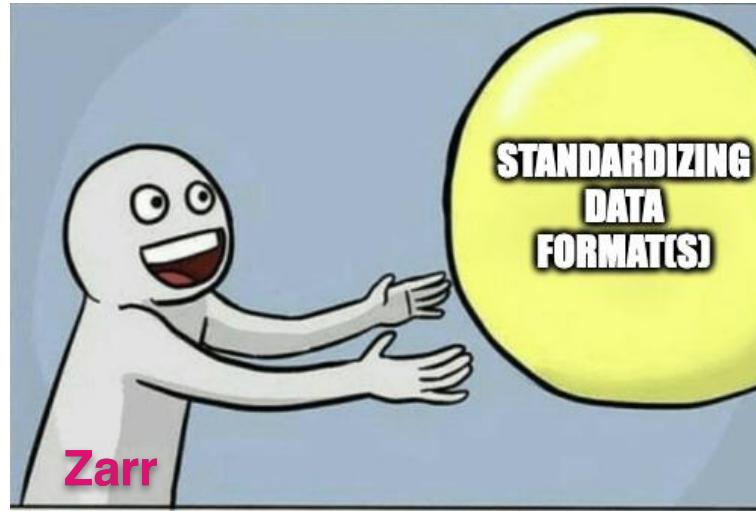


Dialing back to...



We have a large and diverse active community!

But... 



We needed a
structured way to
solicit and ***process***
the feedback!

zarr.dev/zeps



Zarr

Search ZEP

Zarr Homepage

active ZEPs / ZEP0000

home

active ZEPs

ZEP0000

draft ZEPs

template

implementations council

ZEP meetings

join the community

ZEP 0 — Purpose and process

Author: Sanket Verma (@MSanKeys963), Zarr

Email address: svsanketverma5@gmail.com

Status: Active

Type: Process

Created: 2022-14-03

Discussion: <https://github.com/zarr-developers/governance/pull/16>

What is ZEP?

ZEP stands for Zarr Enhancement Proposal. A ZEP is a design document providing information to the Zarr community, describing a modification or enhancement of the Zarr specification, a new feature for its processes or environment. The ZEP should provide specific proposed changes to the Zarr specification and a narrative rationale for the specification changes.

We intend ZEPs to be the primary mechanism for evolving the spec, collecting community input on major issues and documenting the design decision that has gone into Zarr. In addition, the ZEP author is responsible for building consensus within the community and documenting dissenting opinions.

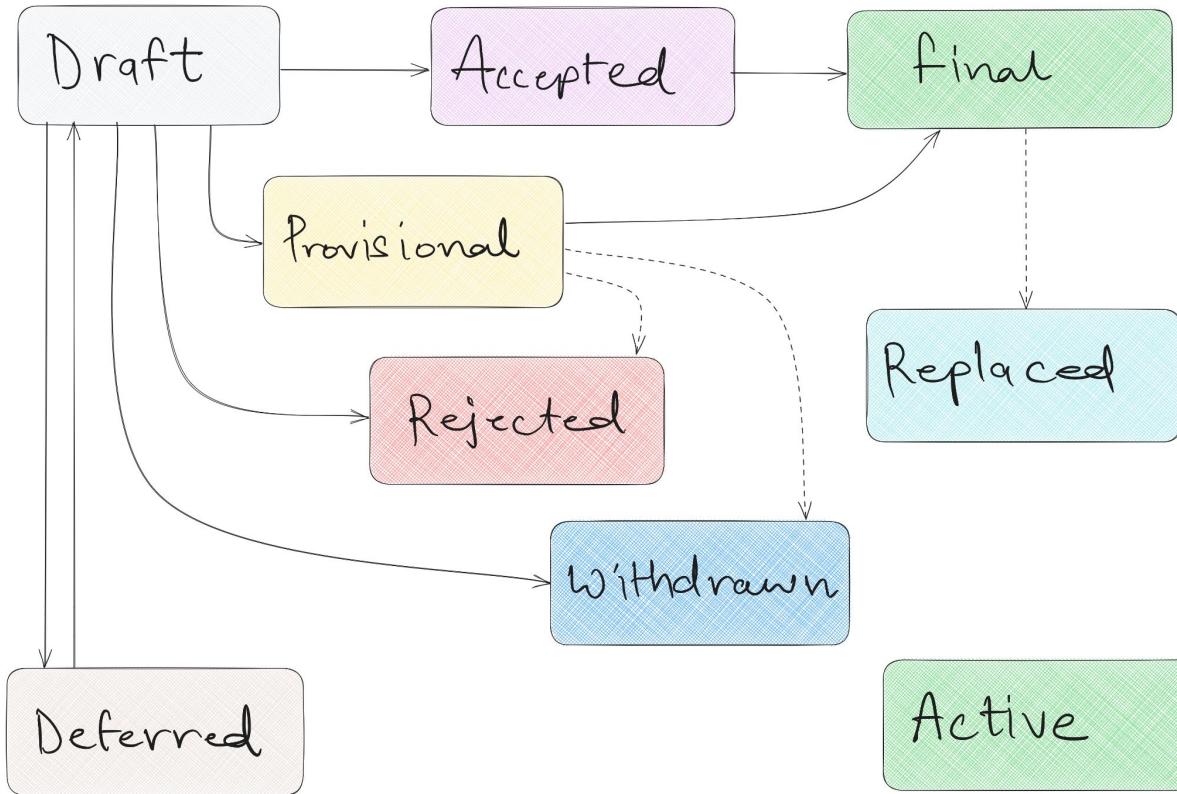
Because the ZEPs are maintained as text files in a versioned repository, their revision history is the historical record of the feature proposal.

WHERE:

- Developers refer to contributors and maintainers of the project
- User(s) refers to an individual or group of individuals or the broader community using the project in any way.

This site uses [Just the Docs](#), a documentation theme for Jekyll.

ZEP Flowchart ↗↖↔



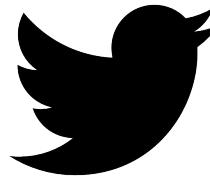
Braindump



ZEP Inception Blog post



Thank you! 🙏



@zarr_dev