

The Sentinels EOPF Toolkit: Driving Community Adoption of the Zarr data format for Copernicus Sentinel Data

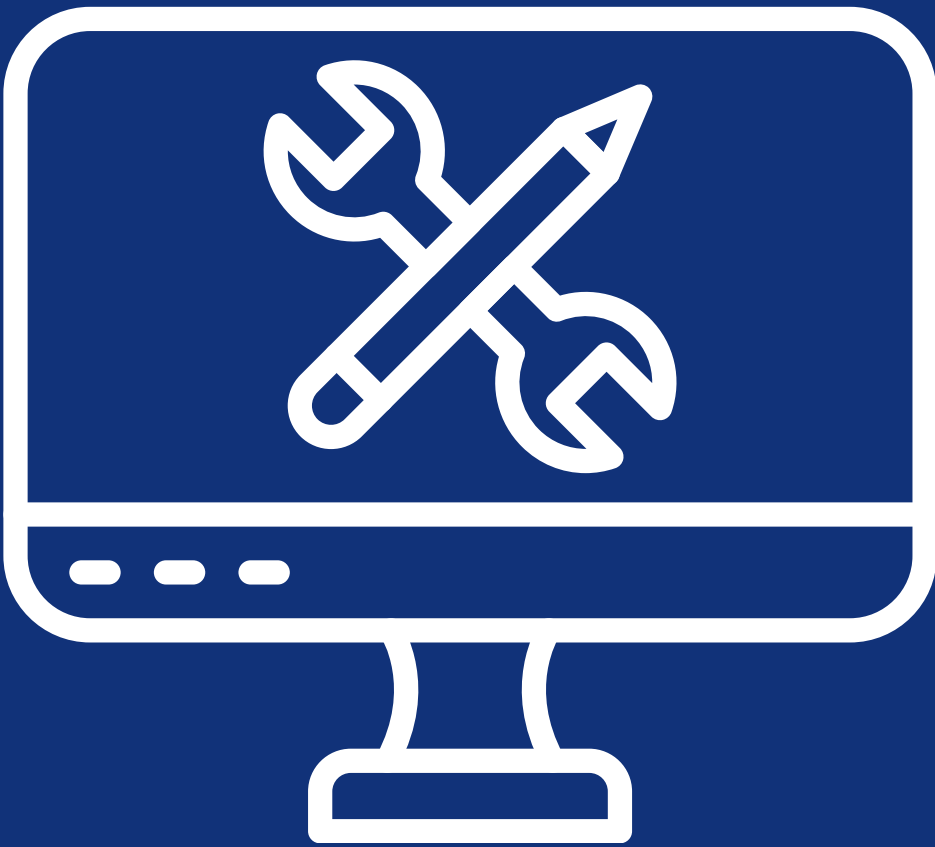
We are a community-driven toolkit that facilitates the adoption of the Zarr data format for Copernicus Sentinel data, targeting users who are new to cloud computing. The Sentinels EOPF Toolkit is developed by Development Seed, thriveGEO and Sparkgeo, with a group of champion users. Together, we are creating EOPF 101, a user-friendly resource consisting of documentation, Jupyter Notebooks and plug-ins that showcase the use of Zarr format Sentinel data for applications across multiple domains. In addition, community engagement activities such as a notebook competition and social media outreach will bring Sentinel users together and spark interaction with the new data format in a creative yet supportive environment.

Authors

Sabrina H. Szeto [1]
Dr. Julia Wagemann [1]
Emmanuel Mathot [2]
James Banting [3]

Affiliations

[1] thriveGEO
[2] Development Seed
[3] Sparkgeo



Introduction

Over the last few years, data standards like STAC and cloud-native data formats like Zarr and COGs have revolutionised how scientific communities work with large-scale geospatial data. The ESA Copernicus Earth Observation Processor Framework (EOPF) will be providing access to “live” sample data from the Copernicus Sentinel missions -1, -2 and -3 in the new Zarr data format through the EOPF Sentinel Zarr Samples Service.

The EOPF Toolkit engages the Sentinels user community to facilitate the adoption of the new Zarr data format, starting with the data published by the EOPF Sentinel Zarr Samples Service. The Toolkit team are test users of the Samples Service data. We also complement existing user engagement efforts by working with users who are new to cloud computing.

Guiding Questions & Answers

What common challenges do users face and how can we help them overcome them?

We are tailoring our resources to common questions that users are asking around data access and use

What resources would make it easier for Sentinel data users to use the new Zarr data format?

Easy-to-understand documentation, practical case studies, example notebook workflows, useful plugins

How can we foster a community of users who will actively contribute to the creation of this toolkit and support each other?

Engage Sentinel data users who are beginners at cloud computing through high quality educational resources, highlighting champion users and giving the user community the chance to try it out themselves via a notebook competition

EOPF 101



This online book is a community resource where you can learn how to use the EOPF Sentinel Zarr Samples Service by ESA. It is designed for Sentinel data users who are new to cloud computing.

- **Launch of first version** together with a notebook contest during Big Data in Space 2025
- **Development of thematic case studies** will occur during early 2026 together with champion users
- **Communications and outreach** through social media and conference presence

Book Chapters

Chapter 1 - Introduction to EOPF

Why cloud-optimized formats, why EOPF, why Zarr for EOPF, the EOPF data model

Chapter 2 - Introduction to Zarr

What is Zarr, what are benefits, how does it differ from Sentinel SAFE, Performance comparisons

Chapter 3 - Introduction to EOPF STAC

Introduction to STAC, how to access data from EOPF STAC catalog

Chapter 4 - Tools to work with Zarr

Different sections for each library and plugin

Chapter 5 - EOPF in Action

Extendable thematic case study library

Glossary

Plug-ins

We will develop a series of open-source libraries and plugins, including the following:

- **Explore Zarr in STAC:** Pystac and QGIS usage with the EOPF STAC catalogue
- **Stackstac**
- **R with Rarr**
- **GDAL** evolution of the current driver
- **Titiler-multidimensional:** Prepare a docker for starting a titiler tailored for EOPF Zarr

Case Studies

Together with a group of champion users, we will develop and publish technical and thematic case studies that include example Jupyter Notebook workflows for using Sentinels data in Zarr format.

Technical case studies (selected)

- Zarr with QGIS
- Zarr with R
- EOPF and STAC (xpystac)
- Multi-scale Zarr
- Dataset screening with rio-tiler and lonboard

Thematic case studies (selected)

- Monitoring coastal dynamics in cloud-prone regions using Sentinel-1
- African rangeland monitoring using Sentinel-2/ Sentinel-3 fusion
- Wildfire assessment with Sentinel-3 and Sentinel-2 data

Get involved



We welcome you to join in this community effort in the following ways:

- Follow us on GitHub (github.com/eopf-toolkit)
- Stay tuned to upcoming announcements
- Participate in the upcoming notebook competition



developmentSEED



thriveGEO



sparkgeo