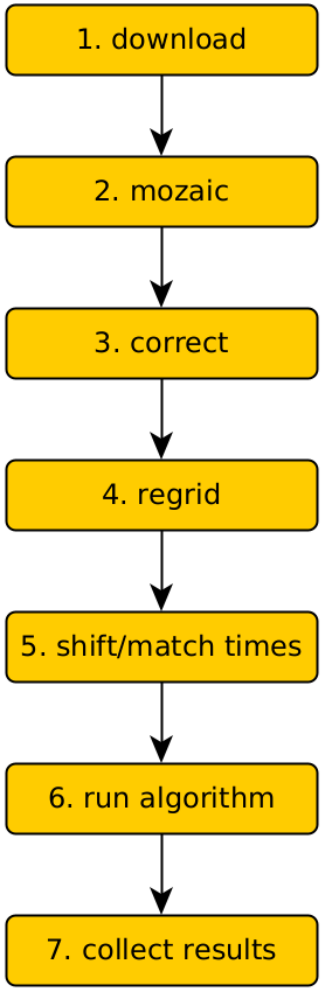
# openEO: open science for Earth observation analytics

Edzer Pebesma, https://mastodon.social/@edzer

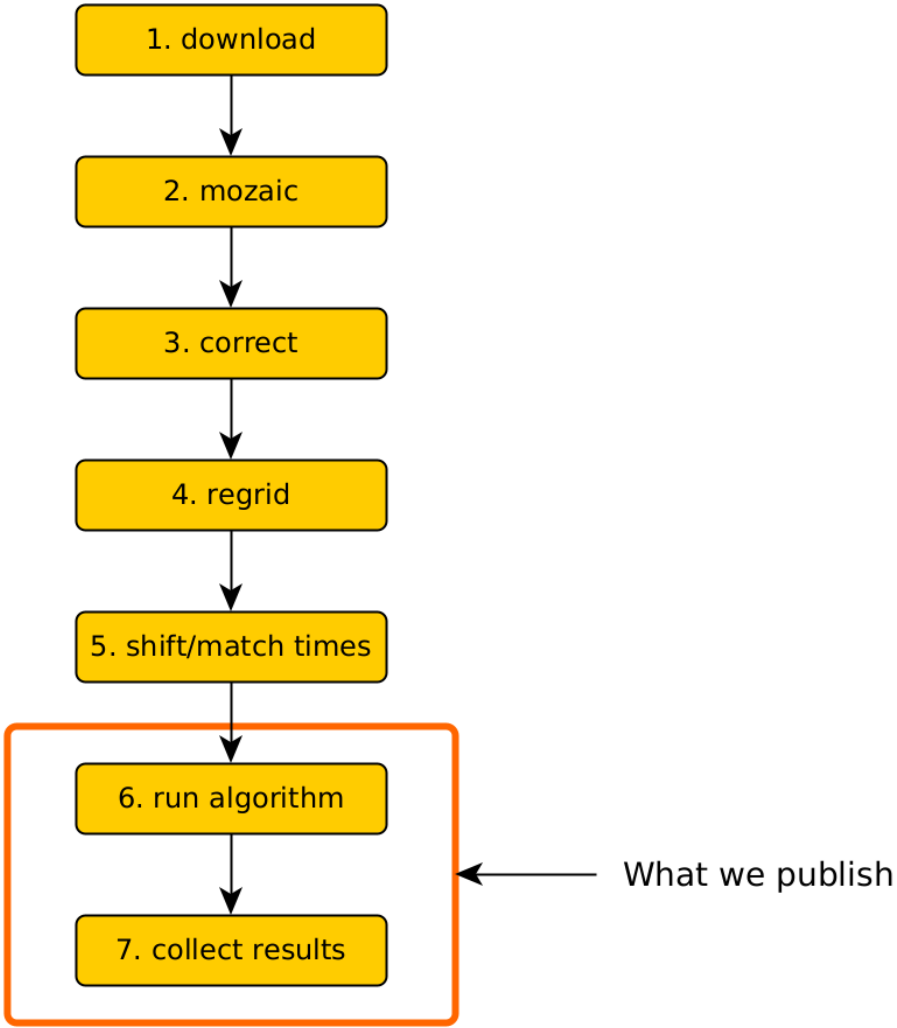
ifgi
Institute for Geoinformatics
University of Münster

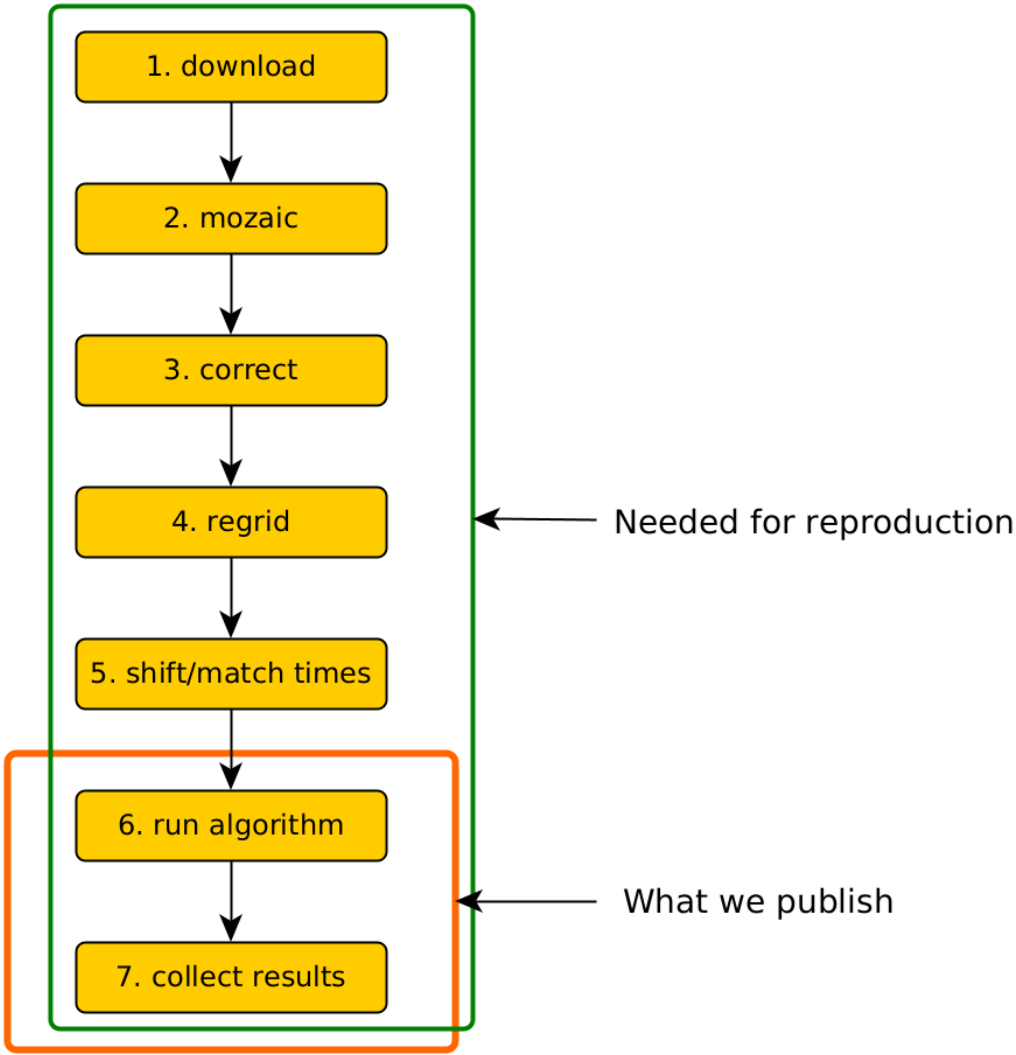
# Where I come from

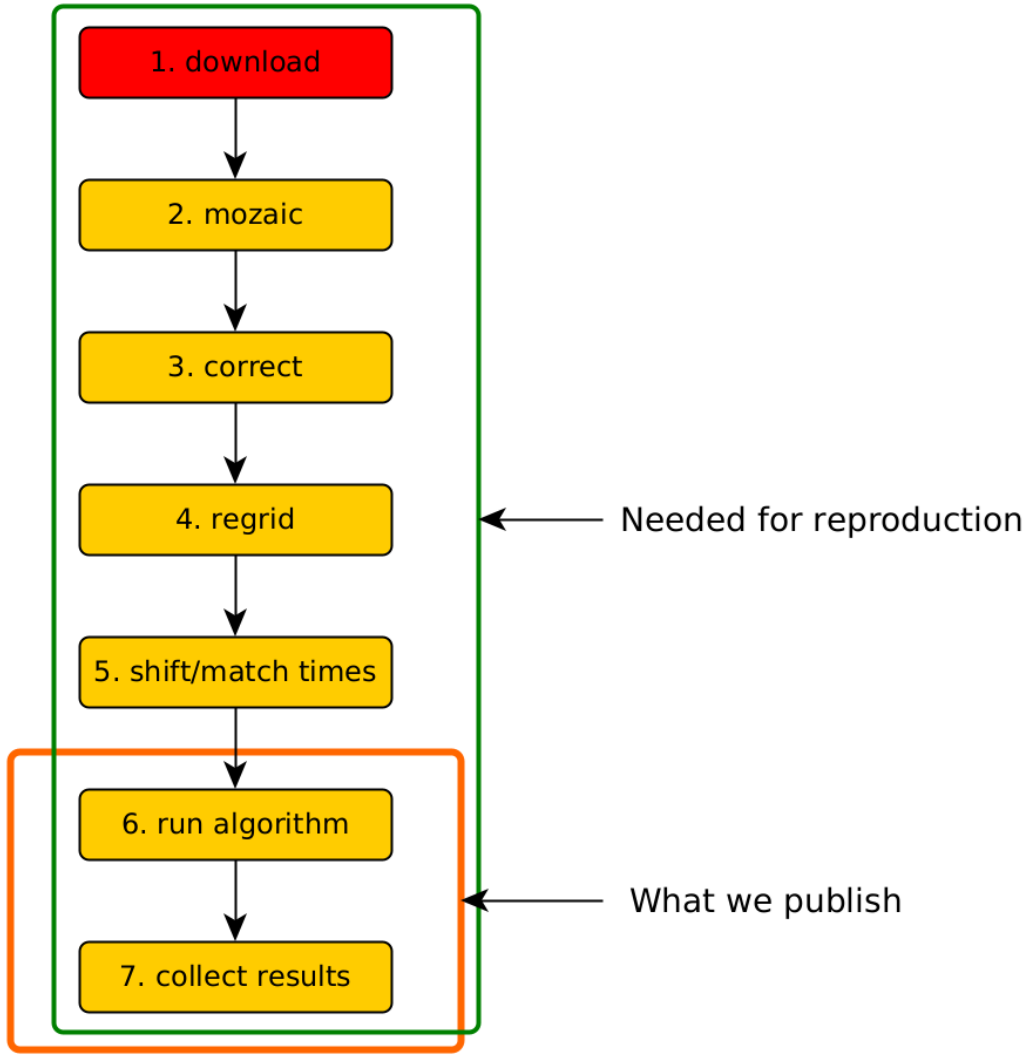
- 1997- Open source geospatial developer & contributor
- 2007- Institute for Geoinformatics, University of Muenster
- Editor of *Computers & Geosciences* (2014-17), *Spatial Statistics* (EB, 2011-19), *Journal of Statistical Software* (2015-21)
- 2016 Initiator of openEO (H2020, ESA, Horizon Europe)
- 2016-21 *Opening Reproducible Research* (DFG)
- 2018- R Foundation ordinary member
- 2019- openEO PSC member



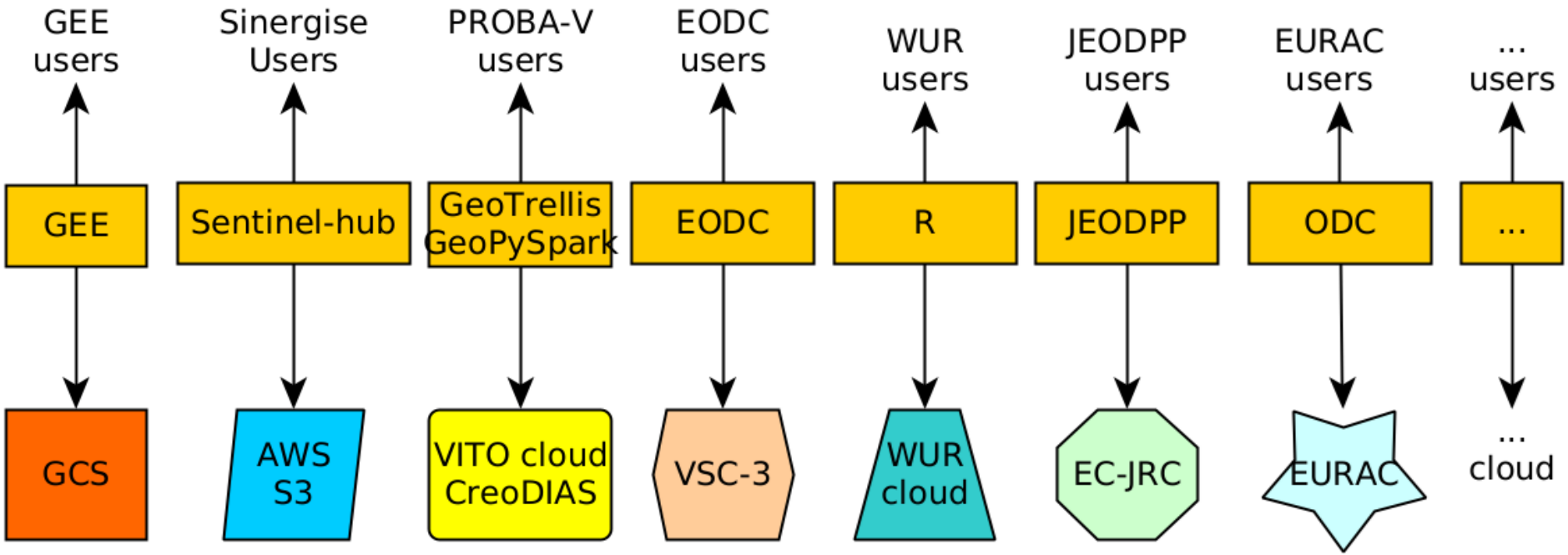

“Breaking down barriers in the scientific use of EO Data”, EODC forum, May 30, 2016


1. download
2. mozaic
3. correct
4. regrid
5. shift/match times
6. run algorithm
7. collect results
What we publish

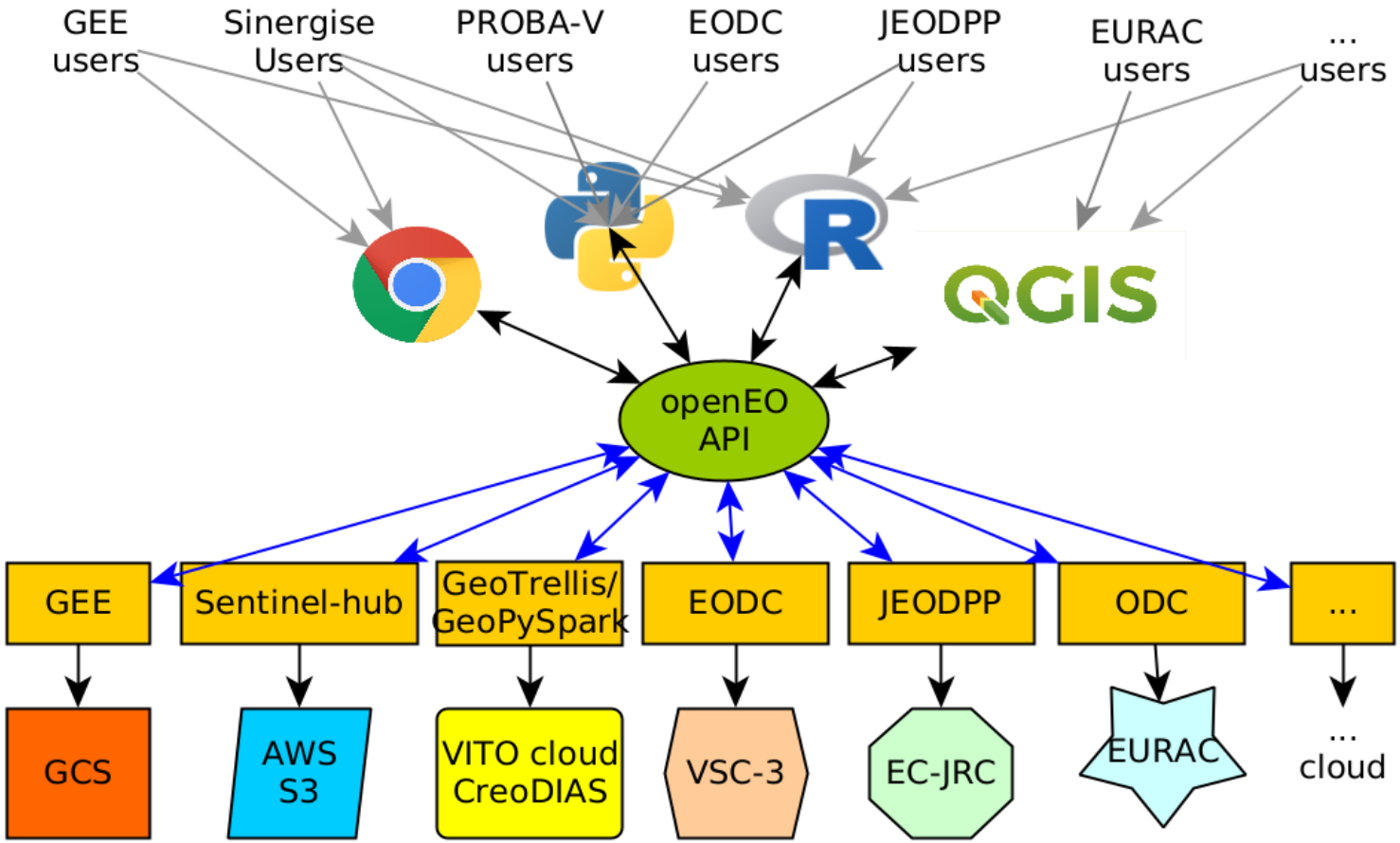

1. download
2. mozaic
3. correct
4. regrid
5. shift/match times
6. run algorithm
7. collect results
Needed for reproduction
What we publish


1. download
2. mozaic
3. correct
4. regrid
5. shift/match times
6. run algorithm
7. collect results
Needed for reproduction
What we publish

# How it started:



# How it’s going:



# openEO concept: *virtual data cube*

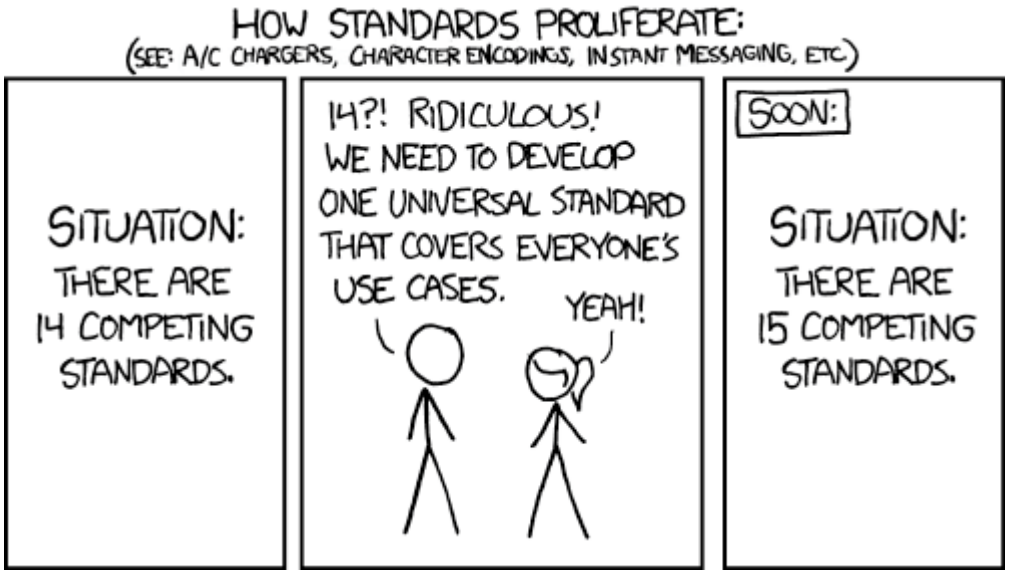
- Data are provided as image collections (STAC collections, these are less structured than Pangeo data cubes or OGC coverages)
- Constrain the problem by defining a target (virtual) data cube:
  - Select region, time period and/or bands (filter)
  - Choose spatial and temporal target CRS, resolution & aggregation method
- Process this data cube further
  - Compute indexes
  - Classify / segment time series etc.
  - Download or view results

# openEO: API & processes

*API*: **how** is the interaction done with a back-end? Managing jobs, accounting,...

*Processes*: **what** is being done with the data?

- All math & logical operations found in common programming languages
- Data cube operations such as filter, apply, reduce_dimension
- A process graph (DAG) of arbitrary complexity reflects expressions
- After connecting a back-end, clients know which image collections and processes are available and how they should be used
- Lazy evaluation: computations are postponed until results are asked for (e.g. for downloading or plotting/viewing)




https://xkcd.com/927/


openEO platform provides intuitive programming libraries to process a wide variety of earth observation datasets. This large-scale data access and processing is performed on multiple infrastructures, which all support the openEO API. This allows use cases from explorative research to large-scale production of EO-derived maps and information.

# Why has this been a success (so far)?

- We built an open, running system (https://openeo.cloud, with support from ESA)
- We built upon / connected to existing open data science communities, while resolving the language/system barrier (Python, R, JS, QGIS, …)
- We embraced existing technological heterogeneity in back-ends (ODC, GEE, Pangeo, GeoTrellis, P.C., …) rather than fighting or augmenting it
- We adopted OpenAPI
- We helped building new technology (STAC collections, now an emerging standard) rather than adopting legacy OGC standards (such as OpenSearch extension for EO or CSW)
- Very early on we communicated the goals and later on the funded project proposal, while sharing progress over social media, conferences, workshops, and hackatons
- We exclusively used open channels (GitHub) for discussions and code development
- We invited everyone to engage, and actively shared this message on conferences and social networks, gave workshops, trainings, monthly **open** developer meetings
- openEO architecture resembles an ecosystem: not everything needs to be supported; clients and backends may appear, or disappear, without disqualifying the system

# Challenges ahead

- Convincing users that open science platforms/APIs/software are a more sustainable proposition, and get them to engage (who provides free plans?)
- Convincing developers and companies that open science approaches are a good investment of resources
- Expanding features, while avoiding feature creep
- Bring in / develop more light-weight and easy to deploy back-ends
- Sustained Funding
- Increasing diversity, equity

Slides:  bit.ly/edzer_esa

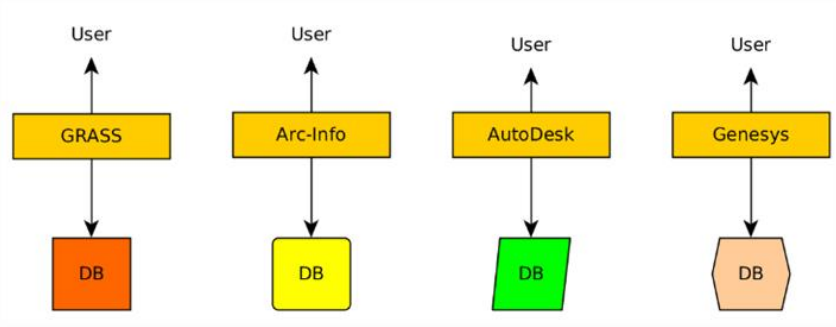
# OpenEO: a GDAL for Earth Observation Analytics

Nov 29, 2016 • Edzer Pebesma, Wolfgang Wagner, Jan Verbesselt, Erwin Goor, Christian Briese, Markus Neteler

Earth observation data, or satellite imagery, is one of the richest sources to find out how our Earth is changing. The amount of Earth observation data we collect Today has become too large to analyze on a single computer. Although most of the Earth observation data is available for free, practical difficulties we are currently facing when we try to analyze it seriously constrains the potential benefits for citizens, industry, scientists, or society. How did we get here?

## GIS: the 80's

To understand the current difficulty when analyzing big Earth observation data analysis, let us look how geographic information systems (GIS) developed over the past decades. In the early days, they would be isolated structures:




where one would get things done in isolation, without any chance of verifying or comparing it with another system: these were expensive systems, hard to set up and maintain, and (with the exception of GRASS) closed databases and closed source software.

zenodo.org/record/1065474



November 23, 2017

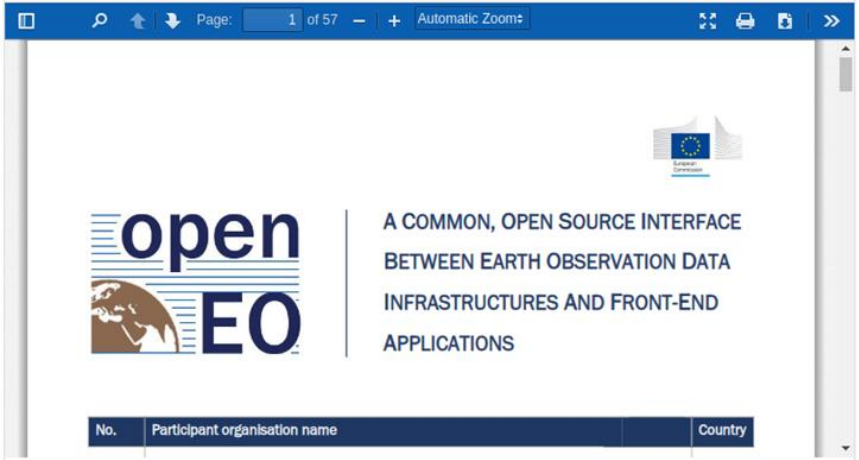
Proposal Open Access

# OpenEO - a Common, Open Source Interface Between Earth Observation Data Infrastructures and Front-End Applications

Pebesma, Edzer; Wagner, Wolfgang; Schramm, Matthias; Von Beringe, Alexandra; Paulik, Christoph; Neteler, Markus; Reiche, Johannes; Verbesselt, Jan; Dries, Jeroen; Goor, Erwin; Mistelbauer, Thomas; Briese, Christian; Notarnicola, Claudia; Monsorno, Roberto; Marin, Carlo; Jacob, Alexander; Kempeneers, Pieter; Soille, Pierre

Project proposal (without effort tables) of the openEO project, see http://openeo.org/ . *openEO - A Common, Open Source Interface between Earth Observation Data Infrastructures and Front-End Applications* is a H2020 project funded under call EO-2-2017: EO Big Data Shift, under proposal number 776242. It will run from Oct 2017 to Sept 2020.

Preview




1,643 views
1,517 downloads
See more details...

Indexed in OpenAIRE

**Publication date:**
November 23, 2017

**DOI:**
DOI 10.5281/zenodo.1065474

**Keyword(s):**
Earth observation, cloud computing, API

**Grants:**
European Commission:
- openEO - openEO - a common, open source interface between Earth Observation data infrastructures and front-end applications (776242)

**Communities:**
openEO - a Common, Open Source Interface between Earth Observation Data Infrastructures and Front-End Applications

**License (for files):**
Creative Commons Attribution Share Alike 4.0 International