

Insights into CODE-DE – Germany’s Copernicus data and exploitation platform

Tobias Storch, Christoph Reck, Stefanie Holzwarth, Bert Wieggers, Nico Mandery, Ulrich Raape, Christian Strobl, Rouven Volkmann, Martin Böttcher, Andreas Hirner, Johanna Senft, Nicolae Plesia, Thomas Kukuk, Stephan Meissl, Jens-Rainer Felske, Thomas Heege, Vanessa Keuck, Michael Schmidt & Helmut Staudenrausch

To cite this article: Tobias Storch, Christoph Reck, Stefanie Holzwarth, Bert Wieggers, Nico Mandery, Ulrich Raape, Christian Strobl, Rouven Volkmann, Martin Böttcher, Andreas Hirner, Johanna Senft, Nicolae Plesia, Thomas Kukuk, Stephan Meissl, Jens-Rainer Felske, Thomas Heege, Vanessa Keuck, Michael Schmidt & Helmut Staudenrausch (2019): Insights into CODE-DE – Germany’s Copernicus data and exploitation platform, Big Earth Data, DOI: [10.1080/20964471.2019.1692297](https://doi.org/10.1080/20964471.2019.1692297)

To link to this article: <https://doi.org/10.1080/20964471.2019.1692297>



© 2019 The Author(s). Published by Taylor & Francis Group and Science Press on behalf of the International Society for Digital Earth, supported by the CASEarth Strategic Priority Research Programme.



Published online: 03 Dec 2019.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



Insights into CODE-DE – Germany’s Copernicus data and exploitation platform

Tobias Storch^a, Christoph Reck^a, Stefanie Holzwarth^a, Bert Wieggers^b, Nico Mandery^a, Ulrich Raape^a, Christian Strobl^b, Rouven Volkmann^a, Martin Böttcher^c, Andreas Hirner^a, Johanna Senft^a, Nicolae Plesia^a, Thomas Kukuk^d, Stephan Meissl^e, Jens-Rainer Felske^f, Thomas Heege^g, Vanessa Keuck^h, Michael Schmidtⁱ and Helmut Staudenrauschⁱ

^aEarth Observation Center (EOC), German Aerospace Center (DLR), Weßling, Germany; ^bT-Systems Sfr GmbH, Bonn, Germany; ^cBrockmann Consult GmbH, Geesthacht, Germany; ^dGAF AG, Munich, Germany; ^eEOX IT Services GmbH, Vienna, Austria; ^fWerum Software & Systems AG, Lüneburg, Germany; ^gEOMAP GmbH & Co. KG, Seefeld, Germany; ^hProgramme Strategy Space R&D, German Aerospace Center (DLR), Cologne, Germany; ⁱSpace Administration, German Aerospace Center (DLR), Bonn, Germany

ABSTRACT

This article presents and analyses the modular architecture and capabilities of CODE-DE (Copernicus Data and Exploitation Platform – Deutschland, www.code-de.org), the integrated German operational environment for accessing and processing Copernicus data and products, as well as the methodology to establish and operate the system. Since March 2017, CODE-DE has been online with access to Sentinel-1 and Sentinel-2 data, to Sentinel-3 data shortly after this time, and since March 2019 with access to Sentinel-5P data. These products are available and accessed by 1,682 registered users as of March 2019. During this period 654,895 products were downloaded and a global catalogue was continuously updated, featuring a data volume of 814 TByte based on a rolling archive concept supported by a reload mechanism from a long-term archive. Since November 2017, the element for big data processing has been operational, where registered users can process and analyse data themselves specifically assisted by methods for value-added product generation. Utilizing 195,467 core and 696,406 memory hours, 982,948 products of different applications were fully automatically generated in the cloud environment and made available as of March 2019. Special features include an improved visualization of available Sentinel-2 products, which are presented within the catalogue client at full 10 m resolution.

ARTICLE HISTORY

Received 9 August 2019
Accepted 24 October 2019

KEYWORDS

Big earth data; big data processing; online data access; Sentinel; Copernicus; CODE-DE; collaborative ground segment

1. Introduction

The fleet of Copernicus Sentinel satellites provides and will provide unique opportunities for global environmental monitoring (Aschbacher, 2017). However, the capability to effectively and efficiently search, access, process and analyse big data streams from the Sentinels and other big data missions such as the Landsat program, still poses major conceptual and technical challenges. The German Aerospace Center (DLR) works towards

CONTACT Tobias Storch ✉ tobias.storch@dlr.de 📧 Earth Observation Center (EOC), German Aerospace Center (DLR), Weßling 82234, Germany

© 2019 The Author(s). Published by Taylor & Francis Group and Science Press on behalf of the International Society for Digital Earth, supported by the CASEarth Strategic Priority Research Programme.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

bridging the gap between the immense data volumes collected by modern Earth Observation missions and their application-driven, on-demand exploration through geo-information clients and services (Kiemle, Molch, Schropp, Weiland, & Mikusch, 2014).

The German Federal Government would like to achieve the broadest and most varied possible exploitation of Copernicus satellite data and services in Germany. Therefore, CODE-DE (Copernicus Data and Exploitation Platform – Deutschland, www.code-de.org) is the German entry point to the EU (European Union) Copernicus Programme and the Sentinel Satellite Systems under the framework of the Copernicus Collaborative Ground Segments from ESA (European Space Agency). CODE-DE provides its services and products full, free and open to the international user community with a focus on fulfilling the needs of German public authorities. Its aim is to operate a system for online data access and big data processing that can be re-used by projects and services. Therefore, CODE-DE is dedicated to granting fast, secure and easy-to-use access to data of all operational Sentinel satellites as well as all derived information generated by the Copernicus services (Keuck et al., 2018).

The requirements for the national Copernicus platform cover four main functions:

- Accessing the Copernicus Sentinel satellite data and products provided by ESA and the European Commission (EC) for registered users, especially users in Germany. Thereby, it is possible to list these products in the user's own infrastructure.
- User-controlled and automated processing of these data into derived products. Thereby, it is possible to include the users' own processing chains in this infrastructure.
- Providing an extended portfolio of products for registered users.
- Monitoring and operating the platform itself.

In addition, the portal to the platform shall be realized with a user-friendly GUI design in order to facilitate the handling of the system for all clients.

The creation of the operational platform service is realized by taking into account the following considerations:

- Data products of the Sentinel Earth observation satellites are obtained via dedicated ESA data hubs.
- The European Commission INSPIRE regulation prescribing standardized interfaces for the dataset discovery, view and access.
- The platform must support hosted services and hosted processing following a paradigm of bringing "the users to the data, not the data to the users". The implementation of operational interfaces and data access mechanisms are mandatory. A flexible processing approach making use of cloud computing is part of CODE-DE.
- A reliable security concept and conformance to the EU General Data Protection Regulation (GDPR).

The innovative aspect of CODE-DE – the implementation all of the above requirements – provides end-users and application developers a complete one-stop integrated, secure and performant platform for discovery and value-added processing of Copernicus data (Reck, Campuzano, Dengler, Heinen, & Winkler, 2016).

In order to view all aspects of CODE-DE, this article is structured as follows:

- Presentation of CODE-DE in the context of Copernicus and the landscape of existing platforms on data sharing and online processing in [Section 2](#).
- Methodology for the establishment of the system in [Section 3](#).
- Architecture of CODE-DE in [Section 4](#).
- Use of CODE-DE for different applications in [Section 5](#).

2. Related works

2.1. European Copernicus program

Copernicus, which is the operational European Earth Observation (EO) programme for monitoring the environment and security, is an ambitious undertaking of the EC, ESA, EUMETSAT and all their member states, to support European citizens, decision makers, scientists and industry with a constant and reliable stream of up-to-date information derived from EO satellite measurements. Copernicus is an integrated system consisting of:

- The space component: Dedicated series of EO satellites – currently six series of so-called Sentinels and other contributing missions.
- The Copernicus services, in the thematic domains: Land, Ocean, Atmosphere, Emergency Management, Climate Change, Civil Security.
- The in-situ component, supporting the Copernicus services combining information from in-situ monitoring networks.

The main Sentinel satellite data distribution is made by the Copernicus core ground segment, with its flight operations segment, the core ground stations, the processing and archiving centres and the mission performance centres. The overall challenge for the data distribution, data holding and data management is the large amount of data generated.

The free, full and open data policy stimulated a demand in the data. In order to deliver at least all data to the citizens and entities of European member states, the concept of distributing Collaborative Hubs (CoIHubs) and national Collaborative Ground Segments (CollGS) was developed. The CODE-DE system is the German national platform, which is not just mirroring incoming Sentinel data for national use, but also offering computing and further service to its users.

2.2. Existing platforms on data sharing and online processing

With the launch of the Sentinels, Earth Observation entered the Big Data era. Infrastructures for processing of Sentinel data have been developed in the private sector, for example, Google Earth Engine, a web-based platform with a dedicated python-API, and Amazon Web Services, a cloud platform with direct access to each band of global Landsat products (Giuliani et al., 2017), and in the public sector, like ESA Thematic and mission Exploitation Platforms (TEPs) and by national space programmes such as PEPS in France, which is a platform similar to CODE-DE for their dedicated national needs (Garcia et al., 2018). Many systems provide only discovery services for datasets, for instance, GDI-DE (Geodateninfrastruktur Deutschland),

CEOS (Committee on EO Satellites), International Directory Network (IDN), and some have catalogues of data products, partially with on-demand online ordering and delivery like DIMS (Data and Information Management System), EOSDIS (EO System Data and Information System) Worldview. Some processing environments such as ESA G-POD (Grid Processing on Demand) reached maturity and other experiments, such as Helix Nebula, were conducted. The availability of free and large volume open data stimulated the evolution from simple download FTP servers like MODIS to the development of OGC (Open Geospatial Consortium) standardized discovery services for linking product catalogue metadata to direct download services. ESA stipulated some very application-specific TEPs to demonstrate the integration of complete workflows, from discovery, to access, visualization and processing within single fully integrated services. SentinelHub demonstrated the combination of large-scale discovery and view services. In 2016, DLR rendered its vision of unifying OGC standardized interfaces with the CODE-DE platform, going beyond the INSPIRE regulation requirements to provide a full range and large-scale service for:

- discovery service for datasets and data products
- view service with on-map data visualization
- direct download service
- processing environment with online data access
- single-sign-on user management services

The Copernicus Data and Information Access Services (C-DIAS) tender from the European Commission started a shift from the national initiatives to a European level, combining efforts of serving all users and commercial entities. The cloud-based C-DIAS concept gathers all Copernicus data and information in one virtual environment together with tools, third party information and the necessary processing environment, allowing users to exploit the information and build their own applications within their own virtual environments. In 2018, five C-DIAS platforms started with first operational components, all featuring different business models (Keuck et al., 2018).

3. Methodology to establish and operate the system

The overall vision of CODE-DE as a platform for product access and value-added processing that can be re-used by projects and services was technically described by 156 User Requirements (REQ). These are further broken down to 248 System Requirements (SR), where each one is linked to exactly one subsystem. These requirement and architecture aspects are complemented by tests for realizing the methodology to establish and operate the system (Storch, Habermeyer, Eberle, Mühle, & Müller, 2013).

To fully exploit the possibilities of the continuous data stream of Copernicus Sentinel products the CODE-DE system comprises:

- Project management, Product assurance, and Systems engineering,
- Subsystems: Infrastructure, Portal, User Management, Ingestion and Archive, Search, View and Access, Processing Environment, Value-added Products, Monitoring and Reporting, and
- Help Desk and Operations.

Each subsystem is assigned to a responsible supplier. The architecture is illustrated in [Figure 1](#) (Reck, Storch, Holzwarth, & Schmidt, 2019).

The CODE-DE architecture in [Figure 1](#) depicts the major functional blocks. Users access the system through a portal, either using a web browser or via the standards conform APIs (yellow boxes). The catalogue client runs as a browser application, seamlessly using the provided APIs, giving the user a fully comfortable browsing experience. Internally the services are modular providing the INSPIRE conform Discovery, Visualization and Download services, each implemented as components indicated in the depicted icons or names (GDAS (Geo-Data Access Services), ECSW, Nginx). The ingestion service runs a workflow extracting and feeding the service layer. The central archive is internally accessible from all services (orange boxes). The distribution service handles pre-configured subscriptions and pushes filtered data to the designated locations (internal or external to CODE-DE). The Processing is realized as a local private cloud, running Calvalus and Apache Hadoop, accessible over a web frontend or WPS (Web Processing Service) API. The ovals on the right are external cross-functional services applicable to the whole platform. More details follow in the next sections.

These subsystems are further broken down to 42 components, where each component provided a specific functionality on its own and on this level, the 66 interfaces are established – and 73 items of hardware, software, service and operations. The system is strongly based on open source elements. Configuration control is performed for items of

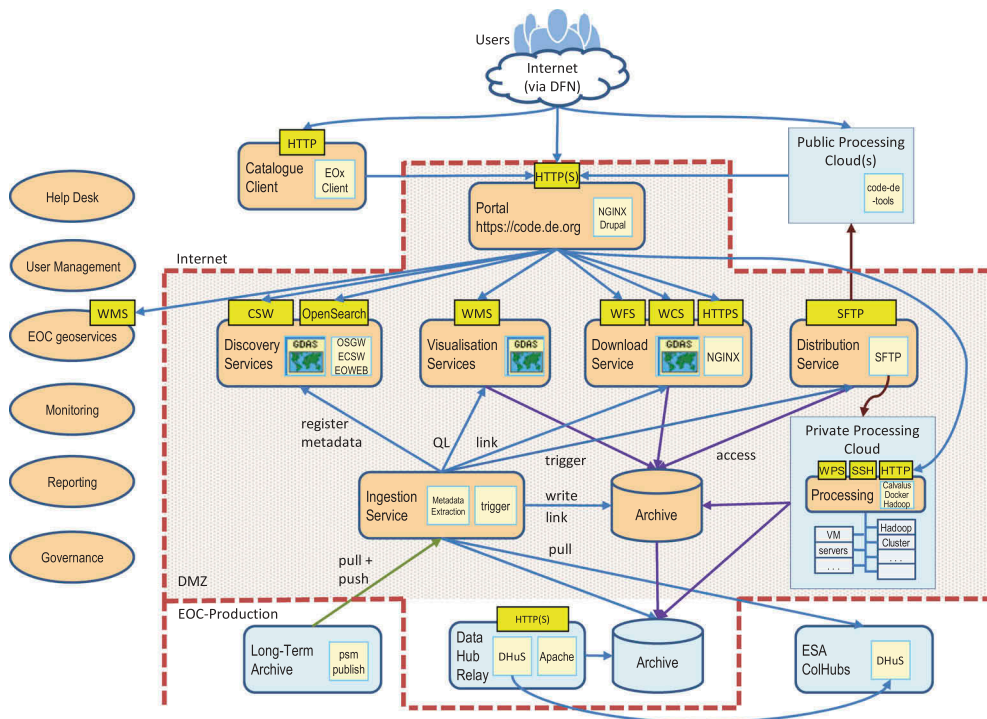


Figure 1. CODE-DE architecture (orange boxes represent CODE-DE services, blue boxes represent services strongly related to CODE-DE, yellow boxes represent standardized interfaces for users).

type hard- and software. However, items of type software are contained in Virtual Machines (VMs) which are assigned to a server, namely an item of type hardware.

The system CODE-DE was realized in two major iterations which are operational

- [Release 1.0] since 9 March 2017 focussing on online data access, where 165 SR were due, and
- [Release 2.0] since 30 November 2017 focussing on big data processing.

In order to guarantee the full functionality and high-quality of the operational platform intense test activities especially on the reference platform were conducted. The reference platform is a mirror of the operational platform with exception of redundancy, performance and network issues as well as updates under consideration. Test case specifications were linked to at least one SR and to a test cycle, especially to Release 1.0 or Release 2.0. For example, 126 test cases were executed successfully for Release 1.0 and 79 test cases had to be executed again for Release 2.0 together with 40 test cases proving the completeness and correctness of Release 2.0.

These test activities are complemented by daily manual operation tests, automatic monitoring and reporting activities, and regression tests, when the system is modified. For example, between 3 September 2018 and 7 September 2018, there was the need for a major modification as the complete infrastructure was moved from Munich, Germany, to Frankfurt, Germany, in order to guarantee the long-term availability of CODE-DE. The transition was necessary as the complete infrastructure of DLR with its CODE-DE-shared services was organized more location-centralized and provider-independent. This modification resulted in 42 selected regression tests, namely covering typical use cases and network tests, which were performed successfully and only one minor bug and one minor non-conformance were identified by monitoring and reporting as well as operation tests. These two observations concerned sending e-mails from portal to help desk; and the VM for the portal was in a critical state. However, they were closed within 4 days.

Until March 2019, 448 observations (157 bugs, 146 non-conformances, and 145 system-changes) were raised (and only 2 observations in March 2019), where only 13 observations (4 bugs, 0 non-conformances, and 9 system-changes) are not closed. These facts demonstrate the high stability of the system.

CODE-DE is established and operated by a team under the lead of the DLR Earth Observation Center (EOC), supported by T-Systems Sfr GmbH, Brockmann Consult GmbH, GAF AG, EOX IT Services GmbH, Werum Software & Systems AG, and EOMAP GmbH & Co. KG contracted by the Federal Ministry of Transport and Digital Infrastructure (BMVI).

4. Architecture

4.1. Infrastructure

The infrastructure consists of a dedicated archive running GPFS (general parallel file system) with 1,444 TByte HD and 19 VMs deployed on six service hosts with 244 cores and 1,536 GByte RAM located in Frankfurt, Germany. In particular, User Management as well as Monitoring and Reporting rely on external environments and services which are shared with other projects to increase the security and responsiveness for a reliable

system. The hosted processing environment consists of four nodes with 112 cores and 512 GByte RAM. The system is linked to the internet with a 5 GBit/s connection. [Table 1](#) lists the operational infrastructure virtual machines (VMs), hosts and storage nodes (Storch, Reck, Holzwarth, & Keuck, 2018).

The colours in [Table 1](#) reflect groups of service hosts placed in separate network zones to protect the services and data from attacks and modification. The distribution of virtual machines on different service hosts also serves for resource and load balancing. The communication between the service hosts and network zones is strongly regulated over local and central firewalls. The communication from the internet to CODE-DE is completely routed and filtered through the HTTPS proxy service. Entitled external project administrators are enabled to access the private cloud and processing environment via a hop server.

The complete operational environment is replicated with an additional set of VMs serving as the test environment. These test environment hosts are not listed in [Table 1](#); these are running on the service host number 6, also serving as a hot spare.

Average performances of the system are above 10 MByte/s for external access to single files and over 100 MByte/s for multiple files (tested with 10 parallel transfers), whereas for internal access, it is 350 MByte/s for single files and rates above 500 MByte/s are achieved accessing multiple files (tests with dd, cp and curl demonstrated peak sustained rates of up to 2 GByte/s for one host and 3 GByte/s with a distributed load over three hosts). These values emphasise the necessity to typically process products directly at their location.

Table 1. CODE-DE infrastructure (hosts named according to the internal provides service, along with the assigned resources).

Host	vCPU	RAM	Storage	Service Host
cd-repository	2	4	1,000 GB	1
cd-ingestion	8	20	60 GB	2
cd-distribution	4	12	50 GB	2
cd-oper	4	20	50 GB	3
cd-catalog	20	32	1,000 GB	3
cd-hub	16	48	1,000 GB	3
cd-access1	20	48	1,000 GB	1
cd-access2	20	48	1,000 GB	2
cd-access3	20	48	1,000 GB	3
cd-tilecache	32	48	100 GB	4
cd-proxy	16	26	50 GB	4
cd-portal	8	16	500 GB	4
cd-discovery	8	8	1,000 GB	5
cd-hop	4	4	0 GB	5
cd-project	16	40	5,000 GB	5
cd-cvportal	4	12	100 GB	4
cd-cvfeeder	4	16	100 GB	4
cd-cvmaster	4	28	100 GB	5
cd-cvprod	6	20	100 GB	5
cd-node01	28	128	960 GB	cluster
cd-node02	28	128	960 GB	cluster
cd-node03	28	128	960 GB	cluster
cd-node04	28	128	960 GB	cluster
bigdatacube	24	64	100 GB	6
gpfs1	-	-	1,002 TB	SAN
gpfs2	-	-	408 TB	SAN

The system availability was 99.8% between March 2017 and March 2019 (99.8% in March 2019) which is defined as average availability of all VMs, where the unavailability of a VM considers unplanned downtimes only and not planned ones.

4.2. Portal and user management

The portal www.code-de.org (CODE-DE, 2019) is visible to all users. It provides information on all available Sentinel and value-added products including pivotal access to search, view and online data access services. A portray service which is realized based on Nginx Web Server allows for configuring and proxying these different services, but it also logs the usage. Furthermore, links to facts on Copernicus, available tools and projects in this context together with public cloud big data processing are displayed. Each such service item and product collection is described by a corresponding ISO (International Organization for Standardization) metadata set which allows harvesting them by the portal using a CSW client to create corresponding queries for the discovery of the system. The query then returns INSPIRE- and GDI-DE-conform metadata entries (Craglia & Annoni, 2007) which are validated using the GDI-DE test suite provided by the Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie; BKG), to be consistent with other standardized services. BKG harvests the CODE-DE collection metadata catalogue, integrating the entries into the Geodatenkatalog.DE for discovery via the Geoportal.DE, the German central national access point to the GDI-DE. For the portal of CODE-DE, the entries are displayed in groups according to their category on separate pages or based on a dedicated search by the user. The details of an entry are shown on user request as a popup overlay or on a separate page. Moreover, the portal includes access to the user help desk and frequently asked questions (FAQs) as well as news and contact information. The portal is illustrated in [Figure 2](#).

Although the portal is accessible anonymously, it interfaces a user management system especially for authentication and authorization of user credentials to prevent access to protected products or services by unauthorized users. The user management system also serves as single sign-on service to the portal and other services and it ensures management of user data compliant to GDPR standards. For this purpose, it offers user self-registration and self-management functionalities with password and user profile updates, where they apply online to optional access rights using group memberships. The user management system is based on an LDAP server which validates and stores the user, group and rights data and performs password policy checks. Some services need to directly access the LDAP server to retrieve user data and perform authentication. However, a CAS server performs authentication and provides interfaces for authentication of users and services. Furthermore, it issues and validates tickets for the single sign-on process. These elements are complemented by a dedicated web service and access service for LDAP administration.

4.3. Online data access

4.3.1. Ingestion and archive

The Sentinel products are retrieved from the various ESA Sentinel data hubs, especially from dedicated data hubs operated for the national collaborative ground segments.



DEUTSCH ENGLISH REGISTER LOGIN

MARKETPLACE ABOUT NEWS HELP Q

CODE-DE

COPERNICUS DATA AND EXPLOITATION PLATFORM - DEUTSCHLAND

Featured Content

The Copernicus Data and Exploitation Platform – Deutschland (CODE-DE) is the German entry point to the Sentinel Satellite Systems, their data products and the products of the Copernicus Services.

read more >

Sentinel-1 Sentinel-2 Tools

MAP MARKETPLACE TOOLS PROJECTS PROCESSORS

DATASETS SERVICES

Latest News

March 18, 2019 - 09:00
Geohazard Supersites TerraSAR-X Download Service

March 13, 2019 - 09:00
CODE-DE: Presentation of Interfaces for Search, View and Download of Copernicus Data at FOSSGIS 2019 Conference on March 14th, 2019, HTW Dresden

January 16, 2019 - 16:00
CODE-DE Demo on "big data from space" 2019 Conference, 19th - 21st Feb. 2019, Munich

GET IN CONTACT WITH OUR USER-HELPDESK ...

... or check out our USER MANUAL

Figure 2. CODE-DE portal www.code-de.org (as seen on 31 March 2019).

These data hubs are realized based on the DHuS software (Tona, Bua, Arcari, Borgia, & Montieri, 2018) to synchronize and distribute the Sentinel products. In CODE-DE it also serves to provide subscription and compatible interfaces of the ESA open access hub. The four major performance indicators with corresponding values for CODE-DE from 7th May 2018 are:

- Delay (average): Sentinel-1 by DHR-NO: 00 h 06 m, Sentinel-2 by CollaborativeHub2: 13 h 21 m (for A)/00 h 09 m (for B), Sentinel-3 by DHR-AT: 00 h 30 m
- Completeness (for 3 May 2018): Sentinel-1 by DHR-NO: 2,694/2,709, Sentinel-2 by CollaborativeHub2: 10,958/11,082, Sentinel-3 by DHR-AT: 378/386
- Availability (average): 100%
- Network (average): 158.74 MByte/s

As the online data storage disk space available for CODE-DE does not allow to keep all this data online, a specific rolling archive scheme is established. Namely, the data of more-interest for CODE-DE users are always kept online, whereas data of less-interest – either for geographic location or age – is evicted from the online disk space. To obtain a high flexibility, this scheme is implemented differently for each product type and it is adapted based on the availability and amount of products as well as user requests. Therefore, the required information is frequently obtained from Monitoring and Reporting and then analysed. For example, the Sentinel-1 SLC products outside Europe were kept online for 2 months in November 2017 and reduced to 0.25 months in March 2019 due to the huge size of products and limited requests by users. All Sentinel-2 Level 1C products (Drusch et al., 2012) over the wider area of Germany are always kept online and specific products obtained from the long-term archive and held online for 1 week. The retention times for eviction per mission product type are illustrated in Table 2.

However, metadata of all evicted products are available via the catalogue and all Sentinel-2 Level 1C products evicted from the online disk space are long-term archived at the German Satellite Data Archive (D-SDA) of the Remote Sensing Data Center (DFD). They are marked as offline and not as unavailable in the catalogue for a high transparency. In March 2019 a volume of 3,246.2 TByte or 6,543,370 Sentinel-2 Level 1C products was stored in the D-SDA. In case an offline product is selected for download or processing, CODE-DE invokes an archive retrieval mechanism, which is able to fetch this

Table 2. CODE-DE storage retention times (per mission product type for March 2019 based on location and age with information on products online in March 2019 and total evicted between March 2017 and March 2019).

Mission Product Type	Germany (months)	Europe (months)	Global (months)	Online (TByte/Products)	Evicted (TByte/Products)
Sentinel-1	36	12	1	121.9/ 156,955	641.4/ 853,969
L1 GRD					
Sentinel-1	36	2	0.25	97.2/ 23,029	2,773.0/ 651,309
L1 SLC					
Sentinel-1	n/a	n/a	12	2.1/ 153,544	1.8/ 179,728
L2 OCN					
Sentinel-2 L1C	36	12	1	362.9/ 1,302,529	3,382.1/ 6,156,875 **
Sentinel-2 L2A	36	1	0.5	92.1/ 152,054	607.4/ 920,729
Sentinel-3 OLCI	36	36	3	46.8/ 76,628	328.0/ 438,422
Sentinel-3 SLSTR	36	36	1	50.9/ 134,028	111.0/ 574,183
Sentinel-3 SRAL	36	36	1	39.3/ 67,802	995.8/ 224,043
Sentinel-5P L2	24	24	24	n/a*	n/a*

*no information because of late availability, **products evicted to D-SDA

demanded data from the archive and transfer it to the online disk space for 5 days. Until March 2019, a total of 1,328 products with a total volume of 0.8 TByte were reloaded in 2164 s per products on average (270 products with a total volume of 0.1 TByte were reloaded in 887 s per products on average in March 2019). The process of the reload mechanism from the long-term archive performs the following activities:

- Request is initiated by the authorized user accessing a product that is not online.
- The HTTPS Download Service calls the Reload Manager for the desired product. The Reload Manager then checks the status of the product in the catalogue and replies with unknown, unavailable, reloading, retry later. Afterwards, it generates a batch request and records the request status.
- The Delivery Service picks up the request and uploads status messages until the file has been uploaded; taking means to avoid partial file transfers.
- The Ingestion Service detects the product in the pick-up point, moves it to the Download Service access point and updates the status in the Reload Manager and in the catalogue.
- The authorized user can perform the direct download.

Value-added products are directly received by the processing systems of the platform.

The administration and operation are also in charge to balance the download between the different sources which was adapted due to availability of, for example, product types, satellite series and performances of the sources. Another challenge was to consider changes of product formats, for instance in December 2016, the Sentinel-2 Level 1C format changed from multiple- to single-tile packages.

4.3.2. Search, view and access

Through the portal, not only are all the product collections searchable with CSW, but all products are searchable as well. The product search is additionally simplified with the support of an OpenSearch protocol binding. All products can be viewed at quicklook or full-resolution level using WMS (Web Mapping Service) and accessed using HTTPS and WCS (Web Coverage Service) with the web-based catalogue client of CODE-DE and via supported APIs based on OGC-compliant catalogues, too.

For a user-friendly search, time, spatial and additional filters like polarization mode or cloud cover, are applied in combination with the various layers such as Sentinel-2 Level 1C and overlays. Based on the search results products can be selected and downloaded either as single products, via a metalink, or URL listing to allow for a higher automatization. The features of the provided powerful yet simple access service range from simple directory browsing, direct download in original formats via HTTPS to quota handling, which limits the parallel downloads per user and throttles the bandwidth depending on the groups the user belongs to. The large series of online data products of the archive service is organized in directory structure in the form of mission/year/month/day/<files>. The majority of the users preferred the HTTPS services instead of using something else, like WCS. The number of accessed products per mission is illustrated in [Table 3](#). The numbers of downloaded products by users are given for March 2017 to March 2019 and for March 2019 to highlight the

Table 3. CODE-DE access, number of downloaded products (by users; per mission product type with information for March 2017 to March 2019 and March 2019).

Mission Type	Product	Download (TByte/Products) (March 2017 to March 2019)	Download (TByte/Products) (March 2019)
Sentinel-1	L1 GRD	11.0/64,500	0.4/2,037
Sentinel-1	L1 SLC	0.2/7,896	0.0/12
Sentinel-1	L2 OCN	0.0/307	0.0/7
Sentinel-2	L1C	5.9/449,099	0.5/163,945
Sentinel-2	L2A	0.2/639	0.0/38
Sentinel-3	OLCI	0.2/743	0.0/10
Sentinel-3	SLSTR	0.1/1,929	0.0/78
Sentinel-3	SRAL	0.0/6	0.0/0
Sentinel-5P	L2	n/a*	n/a*

* no information because of late availability

- increasing acceptance of the system, namely the average monthly downloads between March 2017 and March 2019 are a factor of 0.8/0.1 (TByte/Products) of the downloads in March 2019 and the
- stability of the demand of the different product types, for example, it is relatively fixed for Sentinel-1 L1 GRD and strongly changing for Sentinel-2 Level 1C.

The stability of the demand of the different locations is similarly highlighted in [Figure 3](#) using heat maps for the geographical centre coordinates of the downloaded products by users. There surely is a fixed and high demand for products covering Germany as well as a fixed and medium demand for products covering Europe – and Artic/Antarctica because of a continuous service hosted to map sea-ice for ship navigation. Further, especially sporadic demands arise from projects using the system for download and which focus on specific regions-of-interest such as on Peru for analysis of volcanic activities and Vietnam for analysis of the Mekong Delta.

The search, view and access service is based on the GDAS software to administer process, discover, visualize, and distribute geospatial data for CODE-DE.

On collection level, namely for data set series, ISO metadata are interactively generated and provided via an OGC-compliant Catalogue Service for the Web (CSW). On individual product level, namely for data sets, OGC EOP-compliant metadata are automatically generated and provided via eBRIM-compliant CSW (ECSW) interface, which is also comfortably accessible via an OpenSearch query interface.

Using the WMS geospatial data are visualized in various GIS clients, geospatial data portals, or virtual globes, as well as the CODE-DE Catalog Client. EO-WMS adds the ability to browse a collection of products via the time parameter. As a result, a single WMS layer can be provided which allows the selected products to be shown via the time parameter. Only products with a validity time, typically the acquisition time, intersecting with the requested time are rendered.

Using Web Coverage Service (WCS) and the Web Feature Service (WFS) direct data access is also available to download via HTTPS. By generating a specified request the data for a defined area and time interval are extracted in XML-format for further usage. EO-WCS

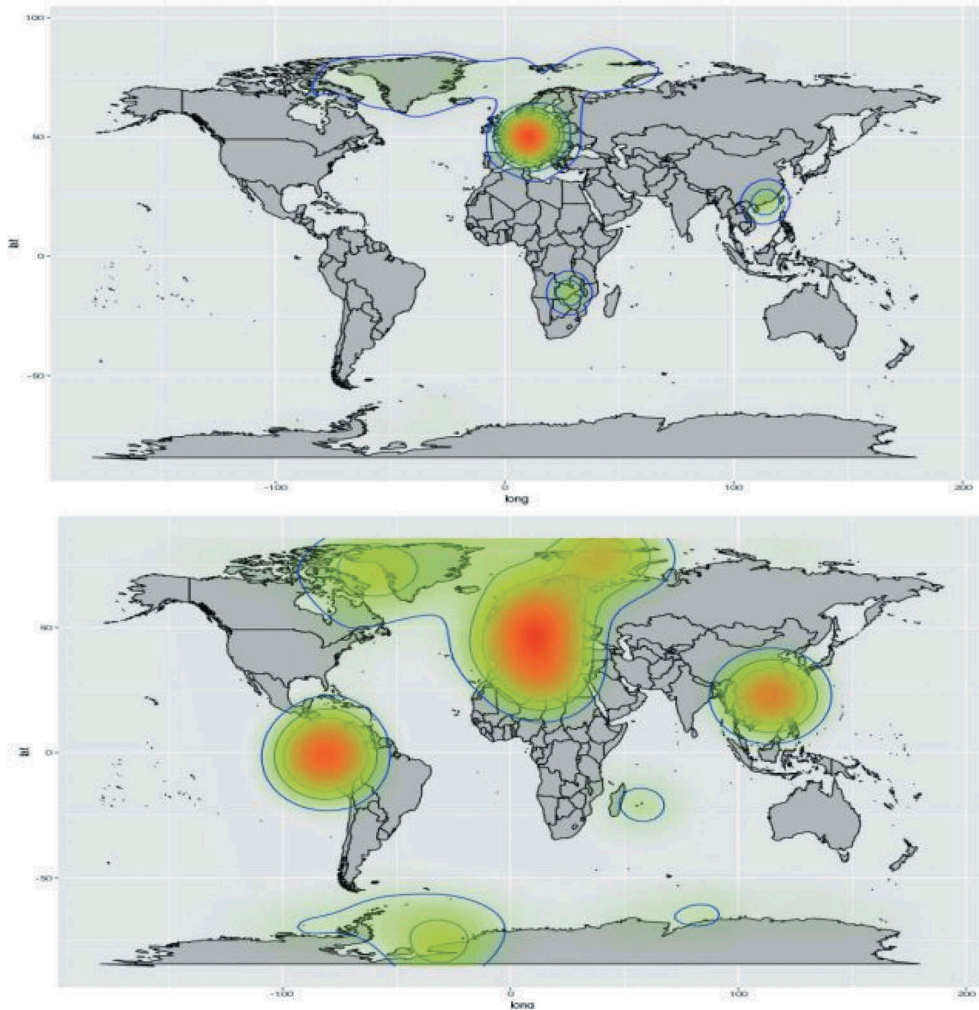


Figure 3. CODE-DE access, location of downloaded products (by users; land in dark grey, water in light grey, number of downloaded products for a location increases from transparent via green to red; with information for March 2017 to March 2019 (top) and March 2019 (bottom)).

adds the capability to group coverages and to perform spatial and temporal searches on those groups.

The CODE-DE Catalog Client provides a user-centred solution to search, view, and access available EO data (Meissl & Triebnig, 2010). Immediately after loading the web-client in the browser it provides a first view on the available products on the map (centre), as list (right), the applied filters (left), and the time slider (bottom) as illustrated in Figure 4. Searches are performed automatically whenever a user interaction like zooming or panning happens, either on the map or on the time-slider, or when the applied filters are modified, and in background without interrupting the user interactions. This behaviour is achieved by a combined exploitation of OpenSearch and WMS/EO-WMS, allowing the immediate display of available mosaics or single scenes. Because both simple searches based on default



Figure 4. CODE-DE Catalog Client (as seen on 31 March 2019; for Sentinel-1 L1 GRD over central Europe for 6 days based on quicklooks (top), Sentinel-2 L1C over Munich, Germany, for 6 days based on processing for high-resolution browsing (middle), Sentinel-3 SLSTR over central and northern Europe for 1 day based on processing for low-resolution browsing (bottom)).

parameters as well as complex searches based on advanced filters are supported, non-expert as well as expert users are catered for, here.

A subscription service to products based on the user needs (for example, specified areas) is available to automatically transfer files of interest to a remote location like an external cloud. By connecting to the Open Telekom Cloud (OTC) any user may obtain access to additional and dedicated processing resources. Both the use of pre-installed Earth observation toolboxes provided by CODE-DE such as SNAP (Sentinel toolboxes) as well as the deployments of customer developed external processors are supported by the OTC (Niggemann, Appel, Bach, de la Mar, & Schirpke, 2014).

4.4. Big data processing

The processing services provide an environment to generate higher level data products from the data products hosted on the CODE-DE platform. Instead of downloading the data users and projects can process the existing Sentinel data products and other hosted data on-site and run their analysis with existing (for example, SNAP) as well as provided data processors. Processing includes on-demand use of the service using the web-client in the browser with data processors and generic processing workflows called public cloud processing and the systematic generation of products and services by projects or CODE-DE called private cloud processing. These processing services are based on a combination of a processing cluster and a virtualised environment with VMs. The VMs run continuous services or control processing jobs on the cluster. While projects running inside the CODE-DE system are restricted in number to the resources available, external projects can add own resources and combine them with remote processing calls within CODE-DE. The processing cluster runs Calvalus (Fomferra, Böttcher, Zühlke, Brockmann, & Kwiatkowska, 2012), and Apache Hadoop (Shvachko, Kuang, Radia, & Chansler, 2010) supports parallel processing and storage of large series of data. Calvalus allows for the submission of processing jobs to the cluster, the control of processing workflows with dedicated tools, namely orchestration with scripting, or specific workflow engines, and the use of the CODE-DE services for systematic data-driven processing. Apache Hadoop is the software used for job queuing, processing task scheduling, and data-local processing with automated failover and re-attempts as well as fair allocation of processing resources. Until March 2019, 195,467 core and 696,406 memory hours were utilized to generate 982,948 output products, where 6,636,100 input products were used. (In March 2019, 27,152 core and 99,418 memory hours were utilized to generate 132,574 output products, where 1,062,272 input products were used.)

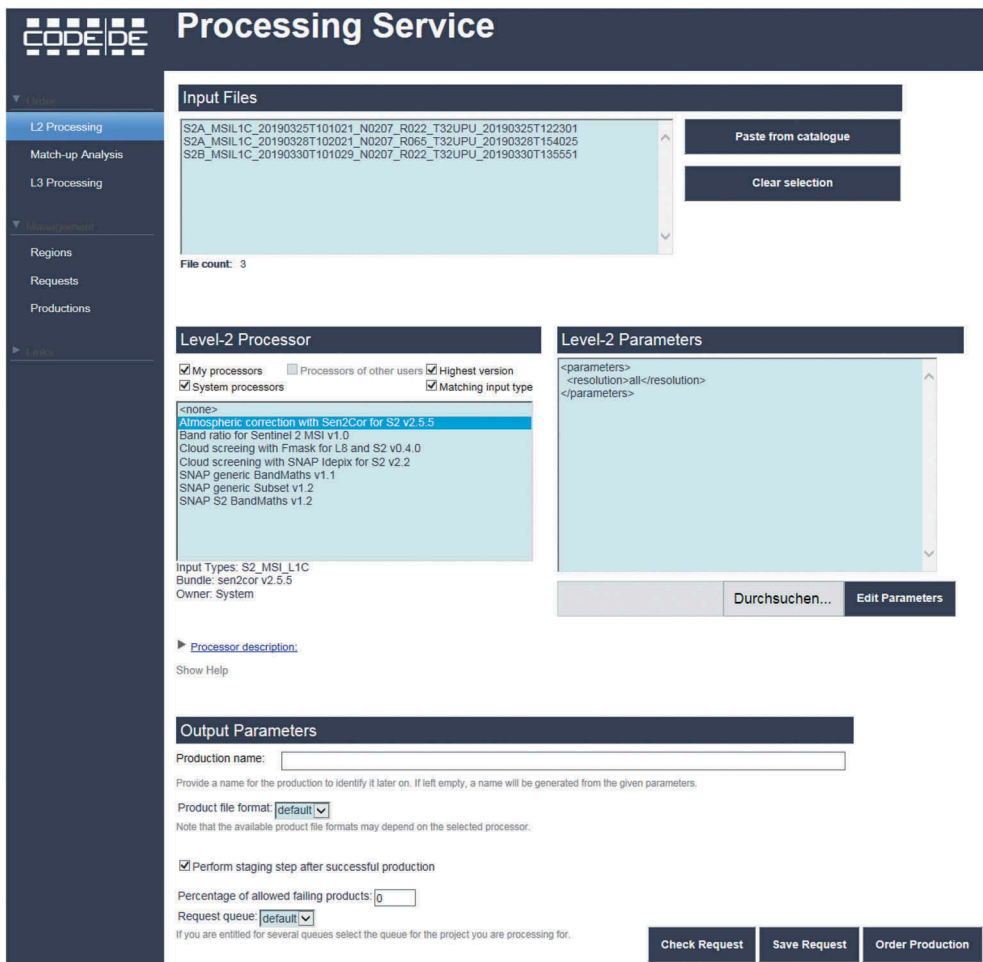
Processing in CODE-DE has relations to the portal and user management for access control, processors integrated or to be integrated, search and access of input data as well as output data publishing.

4.4.1. Web client

The processing web client provides a guided access to the processing environment with its processors and chains of processors. It comprises several web-forms for processing request composition, submission, result access, and monitoring. Namely, registered users freely process the data themselves with the help of various methodologies and then download the generated higher level products. This environment allows the individual selection of algorithms, user-specified parameters as well as a spatial and temporal search for the remote sensing data to which the methodology is to be applied. The methods available to all registered users are mainly excerpts from SNAP such as re-projection

operators which create single input data products in a target coordinate reference system; binning operators which create aggregates of multiple input data products; and generic band math (Zuhlke et al., 2015) to calculate, for example, the NDVI (Normalized Difference Vegetation Index). Sensor-specific processors are also available, such as Sen2Cor for the atmospheric compensation of Sentinel-2 Level 1C products (Main-Knorn et al., 2017). The individual results can be downloaded or are optionally mosaicked to cover larger areas or a larger period with different processing options, for instance, maximum composite. The current status of the processing can be viewed in a processing tab after tasking the job. The graphical user interface of the public cloud processing environment is illustrated in Figure 5 with options for Level 2 and Level 3 processing as well as user-customized processing options with code in XML-format parsed to SNAP via Calvalus.

Besides using the graphical user interface, any registered user may invoke processors to work directly on the CODE-DE data offerings by submitting their processing requests using a WPS interface, supported by the provided tools.



CODE-DE Processing Service

Input Files

S2A_MSIL1C_20190325T101021_N0207_R022_T32UPU_20190325T122301
 S2A_MSIL1C_20190328T102021_N0207_R065_T32UPU_20190328T154025
 S2B_MSIL1C_20190330T101029_N0207_R022_T32UPU_20190330T135551

Paste from catalogue
 Clear selection

File count: 3

Level-2 Processor

My processors Processors of other users Highest version
 System processors Matching input type

<none>
 Atmospheric correction with Sen2Cor for S2 v2.5.5
 Band ratio for Sentinel 2 MSI v1.0
 Cloud screening with Fmask for L8 and S2 v0.4.0
 Cloud screening with SNAP Idefix for S2 v2.2
 SNAP generic BandMaths v1.1
 SNAP generic Subset v1.2
 SNAP S2 BandMaths v1.2

Input Types: S2_MSI_L1C
 Bundle: sen2cor v2.5.5
 Owner: System

Level-2 Parameters

<parameters>
 <resolution>all</resolution>
 </parameters>

Durchsuchen... Edit Parameters

Output Parameters

Production name:

Provide a name for the production to identify it later on. If left empty, a name will be generated from the given parameters.

Product file format:

Note that the available product file formats may depend on the selected processor.

Perform staging step after successful production

Percentage of allowed failing products:

Request queue:

If you are entitled for several queues select the queue for the project you are processing for.

Check Request Save Request Order Production

Figure 5. CODE-DE public cloud processing (as seen on 31 March 2019; for atmospheric correction with Sen2Cor of three Sentinel-2 L1C products over Munich, Germany, for 6 days using standard parameters).

4.4.2. Cloud

Access to a dedicated service host within the hosted processing environment is offered to qualified users and projects. They can deploy their own processors, submit processing requests directly to the processing cluster, access the file system and processing results, automate processing and set-up own service. Access is provided through dedicated VMs or Docker containers, run in the processing environment. While the service host is not necessarily a powerful machine itself, it can use the processing cluster as powerful computing facility. The VMs which have interfaces to the processing environment may use a command-line client to flexibly submit processing requests and may be used as host for services to be linked from the CODE-DE web portal to provide data or processing services to external.

The roles and interfaces of a service host are illustrated in [Figure 6](#).

The VMs or Docker containers may implement workflows to automate the generation and submission of jobs for complex processing tasks either for bulk- or data-driven processing. They have access to the storage with data and processing results, where each processor execution is done in a separate working directory. The working directory is provided upon instantiation by Apache Hadoop and contains input product, (links to) unpacked processor installation package, wrapper script to start processing which depends on the type of framework used, and storage for intermediate data and output products. In addition to the original Sentinel products, CODE-DE processes and provides a set of standardized value-added products as input product for further information extraction.

A processor installation package is a set of compressed and packaged files (.tar.gz or .zip files) containing the runtime software of one or several processors and an optional descriptor file identifies data processors, their input product type, parameters, output product type, and bands of its output product (to be archived). Several conventions are available for processor packages to allow for a high flexibility, among them Docker images, Linux executables (for CentOS), SNAP operators (as jar files), and processors

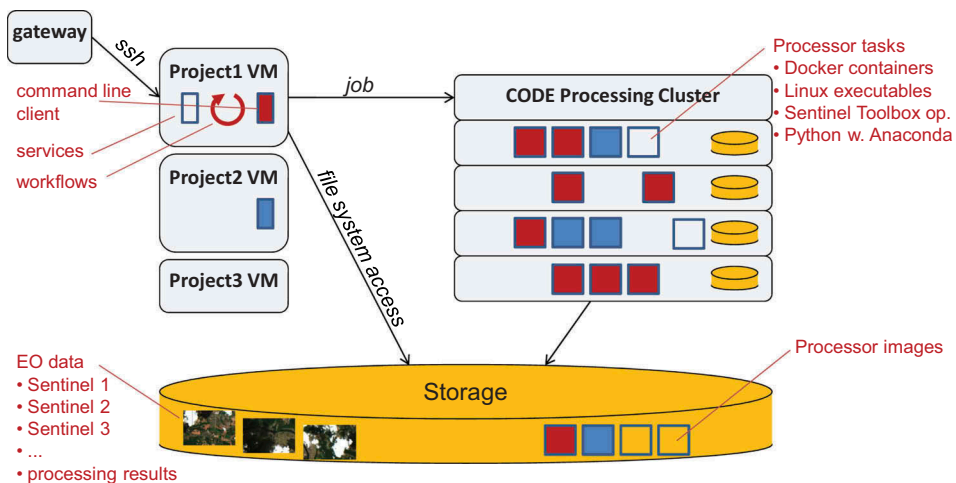


Figure 6. CODE-DE private cloud processing (access over ssh to a project VM, submitting jobs to the cluster, and direct access to the data in the storage).

implemented in Python using Anaconda as runtime environment. The wrapper script, which is to be provided with the processor, is an adapter between the calling convention of the processing infrastructure and the processor, including respective dependencies for automated deployment and parallelisation.

The processing cluster is a shared facility. One queue per project on this cluster in combination with fair scheduling ensures that each project at least gets its share of the cluster computing and memory resources. A project gets more whenever not all projects are processing at a certain time, up to the full cluster capacity, with dynamic adaptation as new requests are submitted.

4.5. Operations

The stable system for CODE-DE operations is divided into three levels, where for the third level, the system and subsystem engineers are directly responsible.

The first level support is covered by the help desk activities, where until March 2019, 1,606 tickets are raised (and 53 tickets in March 2019). But it is also responsible for the news and frequently asked questions (FAQ) as well as trainings and tutorials. Questions are mainly concerning the use of the system, automatic access to the data repositories including questions to interfaces and scripting, user registration and rights, processing and tools. Geographic distribution of requests shows – as expected – a maximum coming from Germany mainly from public entities and universities, then from European users, but also a number of international users.

The second level support performs a continuous daily monitoring of the CODE-DE operational hard- and software. Nagios and Ganglia are employed for monitoring (Wu, Zhang, & Li, 2013), while the operational activities are performed using several JIRA tasks (Filion, Daviot, Le Bel, & Gagnon, 2017), that are created each day. Nagios reports on 26 hosts and 230 services, where the tactical overview section is checked for alerts by means of a dedicated task. Another JIRA task is used for monitoring the information provided by Ganglia on system load and memory allocation for 19 hosts. A dedicated JIRA task is used to check the Sentinel data in the CODE-DE Catalog Client against gaps. Finally, the functionality of addressing the first level support is verified by another JIRA task, where emails are sent from the portal to the help desk to make sure that the communication loop closes successfully.

5. Applications

Several applications are hosted in CODE-DE producing baseline and value-added products. The output data is either offered for direct online download, or rendered in a web visualization service. The CODE-DE own and the EOC GeoService integrate the aspects of visual online data access with an extraction and download service. Beside products of the six Copernicus Services which are linked via the portal directly as well as of Supersites with TerraSAR-X products (Eineder, Roth, Minet, & Amelung, 2011) and RapidEye products for scientists (Borg, Daedelow, Missling, & Apel, 2013), CODE-DE provides an additional variety of products and services.

5.1. Sentinel-2 High- and Sentinel-3 Low-Resolution Browsing

For Sentinel-2 High- and Sentinel-3 Low-Resolution Browsing more than 10,000 Sentinel-2 Level 1C products are produced daily, each with a size of approximately 500 MByte. This results in more than 50 TByte per month or 600 TByte each year. Serving this amount of globally and continuously acquired high resolution satellite data in full resolution is a challenging task. The original Sentinel-2 Level 1C data are available in zipped SAFE file format. Each zip file contains the imagery from one of the $100 \times 100 \text{ km}^2$ grids. It contains the image data, metadata, quality indicators (such as defective pixels masks) and auxiliary data. The image data are provided with JPEG2000 lossless compression resulting in a file size of approximately 50% compared to conventional lossless compression methods. While this is beneficial for storing and providing the raw data, it has negative implications on rendering the imagery. Reading and rendering JPEG2000 compressed data are time-consuming and requires much processing power. Although this is acceptable for non-recurring tasks where small file sizes and good quality outweigh performance needs, it becomes impracticable for continuous on-the-fly operations. Therefore, the following approach was realized. The Geospatial Data Abstraction Library (GDAL) is used for the geographic data processing needs and for the conversion of the original Sentinel-2 raster data stored in JPEG2000 format to the widely supported GeoTIFF container format. For rendering and serving the data, GeoServer is used. GeoServer is a web-server based on Apache Tomcat for rendering and serving geographic data using the OGC WMS together with WCS standards. Therefore, the data are merged to a mosaic in a global target coordinate reference system and the bit depth scaled to 8 bits. Because the visual representation is of interest, here, values above 4,096, that are mainly clouds and glaciers, are cut and stretched down allowing for JPEG compression to receive much smaller file sizes. JPEG2000 leads to best compression (350 MByte) the worst viewing performance compared to other lossless algorithms such as LZW (535 MByte). In contrast, JPEG offers much smaller file size (50 MByte with quality set to 75%) and best viewing performance (Volkman, Strobl, Twele, Heinen, & Reck, 2017). Besides being a valuable service on its own it may also support other existing services (Kempeneers & Soille, 2017).

5.2. Value-added products of CODE-DE and of projects such as AGRO-DE

The value-added products of CODE-DE to support the user community comprise:

- Temporal feature extraction products which are Germany-wide datasets including temporal statistics of Sentinel-2 Level 1C products of an entire year and derived spectral indices. It may, for example, be used to detect changes in land use and land cover in a quantitative manner (Esch et al., 2018).
- Cloud-filtered monthly composites of Sentinel-2 Level 1C products for a more convenient monitoring of, for example, vegetation (Main-Knorn et al., 2017).
- Atmospheric-corrected water albedo of Sentinel-2 Level 1C products including a water-land-cloud mask based on analysis of the atmospherically corrected data and spectral properties as illustrated in [Figure 7](#) (Dörnhöfer, Klinger, Heege, & Oppelt, 2017).

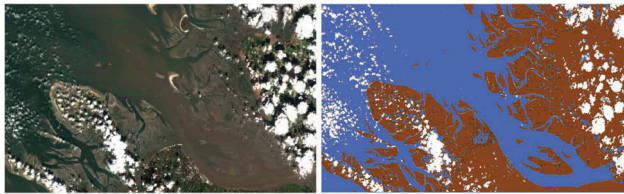


Figure 7. Products of maritime processor (based on Sentinel-2 Level 1C over an area north-west of Hamburg, Germany on 17 September 2017; water albedo (left) for blue, green and red band, mask (right) for water (blue), land (brown), cloud (white)).

The value-added products of the project AGRO-DE are generated to enable farms, consultants, contractors, and service providers to use pre-processed up-to-date open access remote sensing information promptly and integrate it into their business processes. Here, the products comprise:

- Sentinel-2 Level 2A products generated by means of PACO (Reyes, Richter, Langheinrich, Pflug, & Schwind, 2018) in 10 m spatial resolution. Furthermore, cloud, cloud-shadow, snow and water masks are provided which have been calculated by means of the Fmask algorithm (Zhu, Wang, & Woodcock, 2015).
- Sentinel-2 Level 2B vegetation indices products.
- Soil Composite Mapping Processor (SCMaP) makes use of per-pixel compositing to overcome the issue of limited soil exposure due to vegetation and generates product levels that allow for long-term assessment and distribution of soils including the distribution of exposed soils, statistical information related to soil use and intensity as well as the generation of exposed soil reflectance image composites (Rogge et al., 2018) as illustrated in Figure 8.

5.3. Services of projects such as BigDataCube and of EOC GeoService

With the service of the project BigDataCube, the innovative paradigm of the data cube – i.e. analysis-ready massive spatio-temporal raster data – is established (Strobl et al., 2017). Raster data manager (rasdaman) is used to populate and access the data cube (Baumann, Misev, Merticariu, Pham Huu, & Bell, 2018). The query and access language is WCPS (Web Coverage Processing Service), enabling flexible and scalable data retrieval from Sentinel-1 and Sentinel-2 data at any time and without programming. By consistently using the OGC standards, the user remains in the comfort zone of their own clients.

With EOC GeoService products (Dengler, Heinen, Huber, Molch, & Mikusch, 2013) such as SRTM X-SAR Digital Elevation Mosaic – an aggregation of DLR's SRTM X-SAR DTED files – is displayed (Roth, Eineder, Rabus, Mikusch, & Schättler, 2001). The files have been generated from data acquired by the German-Italian X-band interferometric SAR system during the Shuttle Radar Topography Mission (SRTM) in 2000.

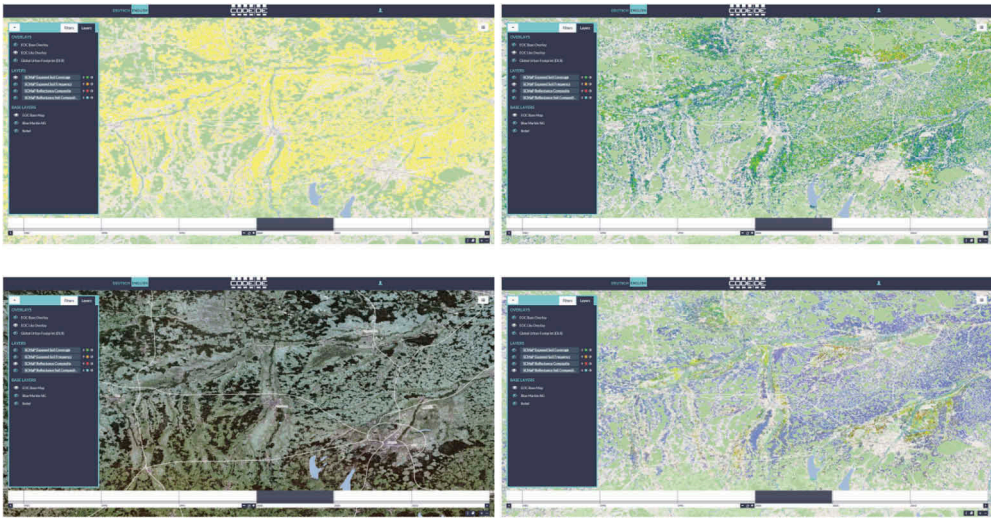


Figure 8. Products of Soil Composite Mapping Processor (SCMaP) (distribution of exposed soils depicted in yellow (top left), statistical information related to soil use and intensity highlighting areas that are prone to soil erosion in red (top right), exposed soil reflectance image composites where different colours can be linked to different soil types (bottom right), reflectance composites (bottom left)).

6. Conclusions and outlook

The methodology, architecture, and various functionalities of CODE-DE (Copernicus Data and Exploitation Platform – Deutschland, www.code-de.org) are presented and analysed to obtain a high-quality system. In order to realize and operate a harmonized infrastructure providing geo-services, the focus is on the major challenges of user-friendly online data access and high-performance big data processing. It is investigated how CODE-DE meets the demands to serve the national platform promoting the utilization of Copernicus products for the public authorities, where major challenges are to fulfil security and interface standards. Nevertheless, various complex applications are already successfully onboarded providing innovative added value.

Since March 2017, CODE-DE has been operational with ingestion capabilities of 4,971 TByte in total until March 2019, including 394 TByte for 571,684 Sentinel products ingested in March 2019. Within the first two years 1,682 registered users, where 1,361 are from Germany, have downloaded 654,895 Sentinel products from the rolling archive which has a total capacity of 814 TByte for Sentinel products. Via the portal, search, view and access services, access is provided to products of the Copernicus Services, Supersites with TerraSAR-X products, RapidEye products for scientists, from CODE-DE internal (such as temporal feature extraction, cloud-filtered mosaics and maritime products) or from CODE-DE external (for example, from the AGRO-DE and BigDataCube projects) and from the EOC GeoService (like the SRTM X-Band DEM).

Since November 2017, a big data on-demand processing environment for high-performance cloud computation is additionally operational. The algorithms available are mainly part of SNAP (Sentinel toolboxes) and the results can optionally be mosaicked. All registered users are able to use this processing environment in a generic way, namely

sharing the available resources of the infrastructure for free. Furthermore, it is possible for selected applications based on dedicated Virtual Machines (VMs) using Docker to use the CODE-DE infrastructure and resources. This flexible processing approach supports different use cases, namely allowing developers, processing and data experts to work on the platform in support of their own processing scenarios.

A first phase of CODE-DE is funded until 31 March 2020, and as a cloud-based platform for Sentinel data access and processing, it is foreseen to enter a second phase taking into account the further developed framework conditions on the part of the Copernicus program and evolving user requirements, and adjust the functions accordingly. An even stronger focus on the applications of users with their derivation of information based on Earth observation (EO) measurements is planned. Continuity of operations for at least five more years with some upgrades, an expanded portfolio of products, such as monthly Sentinel-1 Gamma 0 and Sentinel-2 Level 2A with EVI mosaics covering Germany, and a scalable processing environment, is envisaged. This is with a particular focus on assisting users in national public service institutions with access to analysis-ready Sentinel data and Copernicus service products, as well as with onboarding their processing algorithms. It is envisaged that this CODE-DE transition will make a best possible use of synergies with the European level Copernicus Data Information and Access Services (C-DIAS).

Acknowledgments

The authors thank the three reviewers for their valuable suggestions.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

CODE-DE is realized with funding from the German Federal Ministry of Transport and Digital Infrastructure (BMVI).

References

- Aschbacher, J. (2017). ESA's earth observation strategy and Copernicus. In M. Onoda & O. R. Young (Eds.), *Satellite earth observations and their impact on society and policy* (pp. 81–86). Singapore: Springer.
- Baumann, P., Misev, D., Merticariu, V., Pham Huu, B., & Bell, B. (2018) Rasdaman: Spatio-temporal datacubes on steroids. *Proceedings of 26th ACM SIGSPATIAL* (pp. 604–607). Seattle, WA, USA.
- Borg, E., Daedelow, H., Missling, K. D., & Apel, M. (2013) RapidEye science archive: Remote sensing data for the German scientific community. *Proceedings of 5th RESA Workshop* (pp. 5–20). Neustrelitz, Germany.

CODE-DE (2019) www.code-de.org

- Craglia, M., & Annoni, A. (2007). INSPIRE: An innovative approach to the development of spatial data infrastructures in Europe. In H. Onsrud (Ed.), *Research and theory in advancing spatial data infrastructure concepts* (pp. 93–105). Redlands, CA: ESRI Press.
- Dengler, K., Heinen, T., Huber, A., Molch, K., & Mikusch, E. (2013) The EOC GeoService: Standardized access to earth observation data sets and value added products. *Proceedings of PV* (pp. 1–5). Frascati, Italy.
- Dörnhöfer, K., Klinger, P., Heege, T., & Oppelt, N. (2017). Multi-sensor satellite and in situ monitoring of phytoplankton development in a eutrophic-mesotrophic lake. *Science of the Total Environment*, 612C, 1200–1214.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25–36.
- Eineder, M., Roth, A., Minet, C., & Amelung, F. (2011) DLR's activities to support the GEO supersite initiative with terraSAR-X data. *ESA FRINGE 2011 Workshop* (p. 1). Frascati, Italy.
- Esch, T., Üreyen, S., Zeidler, J., Metz-Marconcini, A., Hirner, A., Asamer, H., ... Marconcini, M. (2018). Exploiting big earth data from space – First experiences with the timescan processing chain. *Big Earth Data*, 2(1), 36–55.
- Filion, L., Daviot, N., Le Bel, J., & Gagnon, M. (2017) Using Atlassian tools for efficient requirements management: An industrial case study. *2017 IEEE International Systems Conference* (pp. 1–6). Montreal, QC, Canada.
- Fomferra, N., Böttcher, M., Zühlke, M., Brockmann, C., & Kwiatkowska, E. (2012) Calvalus: Full-mission EO cal/val, processing and exploitation services. *2012 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5278–5281). Munich, Germany.
- Garcia, V., Paulin, M., Poupard, E., Taillan, C., Gasperi, J., & Louge, C. (2018) PEPs: French platform for usage and distribution of Sentinel products. *Proceedings of 2018 SpaceOps Conference* (pp. 1–6). Marseille, France.
- Giuliani, G., Chatenoux, B., De Bono, A., Rodila, D., Richard, J.-P., Allenbach, K., ... Peduzzi, P. (2017). Building an earth observations data cube: Lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data*, 1(1–2), 110–117.
- Kempeneers, P., & Soille, P. (2017). Optimizing Sentinel-2 image selection in a big data context. *Big Earth Data*, 1(1–2), 145–158.
- Keuck, V., Storch, T., Reck, C., Holzwarth, S., Strobl, C., Böttcher, M., ... Schreier, G. (2018) The German copernicus data and exploitation platform "CODE-DE" - online data access and big data processing. *69th International Astronautical Congress (IAC)* (pp. 1–4). Bremen, Germany.
- Kiemle, S., Molch, K., Schropp, S., Weiland, N., & Mikusch, E. (2014) Big data management in earth observation: The German satellite data archive at DLR. *Proceedings of 2014 Conference on Big Data from Space* (pp. 46–49). Frascati, Italy.
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilms, U., & Gascon, F. (2017) Sen2Cor for Sentinel-2. *Proceedings of SPIE 20427, Image and Signal Processing for Remote Sensing* (pp. 1–3). Warsaw, Poland.
- Meissl, S., & Triebnig, G. (2010). Land monitoring network services based on international geospatial standards: SOSI and Geoland2/SDI Projects. *International Journal of Digital Earth*, 3, 70–84.
- Niggemann, F., Appel, F., Bach, H., de la Mar, J., & Schirpke, B. (2014) The use of cloud computing resources for processing big data from space for operational land surface monitoring in Germany. *Proceedings of 2014 Conference on Big Data from Space* (pp. 210–213). Frascati, Italy.
- Reck, C., Campuzano, G., Dengler, K., Heinen, T., & Winkler, M. (2016) German copernicus data access and exploitation platform. *Proceedings of 2016 Conference on Big Data from Space* (pp. 1–4). Auditorio de Tenerife, Spain.
- Reck, C., Storch, T., Holzwarth, S., & Schmidt, M. (2019) Online data access and big data processing in the german copernicus data and exploitation environment (CODE-DE). *Proceedings of 2019 Conference on Big Data from Space* (pp. 1–4). Munich, Germany.

- Reyes, R., Richter, R., Langheinrich, M., Pflug, B., & Schwind, P. (2018) Validation of a new atmospheric correction software using AERONET reference data PACO: Python-based atmospheric correction. *Workshop on Land Product Validation and Evolution* (p. 1). Frascati, Italy.
- Rogge, D., Bauer, A., Zeidler, J., Müller, A., Esch, T., & Heiden, U. (2018). Building an exposed soil composite processor (SCMaP) for mapping spatial and temporal characteristics of soils with Landsat imagery (1984–2014). *Remote Sensing of Environment*, 205, 1–17.
- Roth, A., Eineder, M., Rabus, B., Mikusch, E., & Schättler, B. (2001) SRTM/X-SAR: Products and processing facility. *2001 IEEE International Geoscience and Remote Sensing Symposium* (pp. 745–747). Sydney, Australia.
- Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010). The Hadoop distributed file system. *26th IEEE Symposium on Massive Storage Systems and Technologies* (pp. 1–10). Incline Village, NV, USA.
- Storch, T., Habermeyer, M., Eberle, S., Mühle, H., & Müller, R. (2013). Towards a critical design of an operational ground segment for an earth observation mission. *Journal of Applied Remote Sensing*, 7(1), 1–12.
- Storch, T., Reck, C., Holzwarth, S., & Keuck, V. (2018) CODE-DE – The Germany operational environment for accessing and processing copernicus sentinel products. *2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1–4). Valencia, Spain.
- Strobl, P., Baumann, P., Lewis, A., Szantoi, Z., Killough, B., Purss, M., ... Dhu, T. (2017) The six faces of the datacube. *Proceedings of 2017 Conference on Big Data from Space* (pp. 1–4). Toulouse, France.
- Tona, C., Bua, R., Arcari, C., Borgia, B., & Montieri, A. (2018) Open source datahubsystem: From Sentinels data access to an agnostic data distribution tool. *Processing of 42nd COSPAR Scientific Assembly* (p. 1). Pasadena, CA, USA.
- Volkman, R., Strobl, C., Twele, A., Heinen, T., & Reck, C. (2017) Serving continuous and global high resolution satellite data – An example based on Sentinel-2 data. *Proceedings of 2017 Conference on Big Data from Space* (pp. 1–4). Toulouse, France.
- Wu, M., Zhang, Z., & Li, Y. (2013) Application research of Hadoop resource monitoring system based on Ganglia and Nagios. *2013 IEEE International Conference on Software Engineering and Service Science* (pp. 684–688). Beijing, China.
- Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269–277.
- Zuhlke, M., Fomferra, N., Brockmann, C., Peters, M., Veci, L., Malik, J., & Regner, P. (2015). SNAP (Sentinel Application Platform) and the ESA Sentinel-3 toolbox. *Proceedings of Sentinel-3 for Science Workshop* (p. 21). Venice, Italy.