

## RESEARCH PAPER

# The Norwegian National Ground Segment; Preservation, Distribution and Exploitation of Sentinel Data

Trygve Halsne, Lara Ferrighi, Bard Saadatnejad, Nico Budewitz, Frode Dinessen, Lars-Anders Breivik and Øystein Godøy

Norwegian Meteorological Institute (MET Norway), Oslo, NO

Corresponding author: Trygve Halsne ([trygve.halsne@met.no](mailto:trygve.halsne@met.no))

---

In order to take advantage of the Sentinel program, the Norwegian Space Agency decided to establish a national collaborative ground segment for satellite data with the purpose of simplifying data access, ensure support for operational national services and long term preservation of data. This is the NBS where MET Norway has the technical responsibility in terms of providing the infrastructure and storage capacity for data management. Serving the data through two separate platforms, the end users have access to the data in its original format in addition to Sentinel-1 and Sentinel-2 products in NetCDF-4/CF. Using the latter format, services like regridding, subsetting, visualization and aggregation are integrated utilizing OPeNDAP in combination with OGC WMS and OGC WPS. In addition, data uploading and retrieving operations are simplified for an end user since streaming of data by means of OPeNDAP is supported in multiple programming languages. Due to the strong coupling between space based earth observations, in-situ observation, model data etc, disseminating data in a generic data management system utilizing NetCDF-4/CF and OPeNDAP is convenient for seamless integration across branches. However, the current CF version is not mature for handling all parts of the Sentinel data but future development looks very promising.

---

**Keywords:** NetCDF; OPeNDAP; OGC WMS; OGC WPS; Sentinel; FAIR; Discovery Metadata

---

## 1. Introduction

The European Space Agency (ESA) has developed a new series of satellites which are dedicated to specific missions. These are the Sentinels which are focusing on the operational needs of the European Union's Copernicus program for Earth Observation (EO). A number of satellites have been planned and several have been launched, producing terabytes of data every day. The Sentinel products can be used for a widespread of applications within e.g. operational services and scientific research.

Data access is a crucial issue for exploiting the potential of such a large satellite constellation. Ever since the first satellite programs were launched, the access to remote sensing products has been limited in terms of costs and restrictions (Turner et al. 2015). Hence, only agencies and institutions with necessary resources and special interests (e.g. security instances and research institutions) have been able to retrieve and work with these types of geophysical data. The paradigm shift introduced by NASA in 2008 with Landsat and later through the European Sentinel program, where high resolution remote sensing products are made easy available for the public under a free data policy, enables the exploitation of data for a broader set of institutions, for small and medium-sized enterprises and also for private persons, thus also enhancing data value (Council 2012; Wulder and Coops 2014).

Norway is responsible for monitoring large sea and land areas at high latitudes and in the Arctic with few, if any, inhabitants, like the Svalbard archipelago. Thus, space-based EO data is an important, if not essential, source to gain insight on several types of phenomena and used in many environmental applications. Processing of these data can help following the development of human interactions with nature, e.g.

measuring land cover changes due to urbanization and deforestation, which have impact on the ecosystem and on the biodiversity at a global scale (Hansen and Loveland 2012; Wulder et al. 2018). Moreover, space-based EO data is a critical component for climate monitoring, e.g. recording snow cover during winter and retrieval of sea ice parameters as extent, thickness and type (Yang et al. 2013). In addition, lower level products, i.e. products requiring no further processing, can also be used directly like Synthetic Aperture Radar (SAR) imagery for ship navigation in sea ice covered areas (Dierking 2013).

Although space-based observations play a key role in describing multiple ongoing processes on the Earth's surface and in the atmosphere, they do not provide the full picture (Council 2007). They are limited in terms of e.g. the satellite instrument observation geometry not covering all parts of the Earth (like near polar orbits not covering the pole), spatio-temporal & radiometric resolution not providing full coverage and atmospheric distortion. Moreover, instrument specific limitations related to the measurement technique does also manifest itself like clouds in optical imagery and layover in SAR. In addition, in-situ measurements and airborne observations are critical for verification of space-based measurements and in the development of predictive models for e.g. air temperatures and permafrost (White-Newsome et al. 2013; Cristóbal, Ninyerola and Pons 2008; Westermann et al. 2015). This calls for generic software tools for data manipulation as well as generic ways of dealing with data management.

On behalf of the Norwegian Space Agency (NOSA), the Norwegian Meteorological Institute (MET Norway) is developing and implementing the National Ground Segment ("Nasjonalt Bakkesegment", NBS) for satellite data, following the FAIR (findable, accessible, interoperable and reusable) principles of data management (Wilkinson et al. 2016). The project focuses on i) simplifying satellite data access for end users, ii) ensuring support for national services and iii) long time preservation of satellite data products covering Norwegian areas of interests. Currently, data are served through two separate platforms: colhub.met.no and satellittdata.no. The first makes use of the Data Hub Software (DHuS)<sup>1</sup> suite for retrieval and dissemination of data while the latter, targeting both expert and non-expert satellite users, is built on the principles of an open data space where non-EO data can be easily integrated. The NBS setup in satellittdata.no is designed using lessons learned in data management efforts during the International Polar Year, WMO Global Cryosphere Watch, European projects (e.g. DAMOCLES, ACCESS, APPLICATE) as well as a number of national geoscientific e-infrastructure projects supported by the Research Council of Norway (including e.g. Norwegian Scientific Data Network – NorDataNet,<sup>2</sup> and Norwegian Marine Data Centre – NMDC<sup>3</sup>). This implies that the system is driven by discovery metadata describing content and interfaces to well documented data following the same approach as the World Meteorological Organization Information System<sup>4</sup> and INSPIRE<sup>5</sup>. As a result, the system is generic in terms of supporting heterogeneous data, i.e. data originating from various scientific branches like physics and biology, and/or generated differently like observations or numerical models. Moreover, emphasis is put on the need for semantic translations and dynamic transformation of datasets upon user request.

At the European level, a great effort is laid down into the realization of scalable EO infrastructure projects like the Data Information and Access Service's (DIAS) due to the amount of data available from the Copernicus program and earlier satellite missions. Here, the intention is to build platforms to make access to these data easier, and in addition, to offer services e.g. for cloud computing (ESA 2017). Other existing commercial projects addressing these issues are e.g. the Google Earth Engine (Gorelick et al. 2017). In NBS, it is a constant task to review national needs and user requirements up against potential costs in terms of implementing a national system versus using already existing infrastructure.

In this article the preservation, distribution and exploitation of Sentinel data will be discussed by means of the two portals with emphasis on the NBS setup. In addition, some issues concerning the need of developing a national infrastructure as a supplement to the other ongoing EO infrastructure projects on a more global scale, are addressed. The structure of this article is as follows: Section 2 describes the various technical components of the system infrastructure, Section 3 highlights the resulting products, services and the use of the system, Section 4 discusses the use of the system with both examples and status quo, and finally, the conclusion and future plans are given in Section 5.

---

<sup>1</sup> <https://sentineldatahub.github.io/DataHubSystem/>.

<sup>2</sup> <https://www.nordatanet.no/>.

<sup>3</sup> <https://nmdc.no/>.

<sup>4</sup> <http://www.wmo.int/pages/prog/www/WIS/>.

<sup>5</sup> <https://inspire.ec.europa.eu>.

## 2. Technical Infrastructure

### 2.1 Data Retrieval, Dissemination and Preservation

The Sentinel missions produces terabytes of data every day which are disseminated to the public under a free and open data policy. Some of the member states within the Copernicus program tailor complementary data products with e.g. regional coverage and specifications like monitoring of volcano activities and landslide risks. Some of these products are considered as a supplement to the Sentinel core products and are relevant for the Copernicus program. In addition, production of local quasi-real time products from local ground stations receiving Sentinel data during the satellite overpass is also supported as long as it does not interfere with the systematic operations of the Copernicus ground segment. Hence, ESA established the Sentinel Collaborative Ground Segment (CGS)<sup>6</sup> for product dissemination and access. The CGS allows for shared distribution of Sentinel and complementary products through a distributed system, consequently supporting shared distribution between the participating nodes and international corporation. In practice, the CGS reduce the payload for the main retrieval and dissemination nodes provided by ESA with multiple pick-up points and mirror sites for Sentinel data. Norway is a participating nation in the CGS where MET Norway is responsible for running the project. Thus, the NBS is retrieving Sentinel products through the Norwegian node in CGS, which utilizes both the ESA dissemination nodes and the other CGSs for product retrieval and dissemination.

All Sentinel products covering the Norwegian areas of interest (i.e Norwegian mainland and coast, North and Barents Sea, the Svalbard archipelago with surrounding waters and Queen Maud Land in Antarctica), are preserved in a long-time archive at MET Norway's facilities. The archive is implemented using a high performance data storage with integrated integrity checking. The underlying hardware is a Lustre Storage Back-end<sup>7</sup> running on an OpenStack<sup>8</sup> infrastructure. The software used for both retrieval and dissemination of products is the DHuS software suite, allowing for data discovery and data access in a web GUI and through machine readable APIs using the Open Data Protocol (ODATA) and OpenSearch.

### 2.2 NBS Dissemination Setup

The NBS dissemination setup follows a modular approach where components serving specific functionality, as e.g. data discovery and data access, are stand-alone components which can be exchanged without changing the whole system. As shown in **Figure 1**, all the parts are wrapped together in the MET Norway Scientific Information System (METSIS), a system where the GUI is integrated in the Drupal<sup>9</sup> Content Management System (version 7) using an open source platform. METSIS consists of multiple modules where each module provides an interface communicating to the underlying web service. A number of these are REpresentational State Transfer (REST) web service compliant thus offering interoperable APIs that are utilized within the system setup as well as outside (Fielding and Taylor 2000). The key parts and services of the system are described in the following sub sections.

In order to obtain a precise description of metadata, the two following sub-categories have been used: discovery metadata and use metadata. The first one, for documentation of products/datasets. Discovery metadata describes e.g. the who, what, where and when about the products as well as the interfaces and access points to the data. Examples of discovery metadata standards are the GCMD DIF<sup>10</sup> and ISO 19115.<sup>11</sup> The latter one, i.e. use metadata, provides a definitive description of what each variable in the product/dataset represents. Use metadata serves the purpose of describing the actual content of the data themselves allowing users to understand and correctly use the datasets. Examples of use metadata are units, missing values and spatio-temporal properties of the data.

#### 2.2.1 Data Format and Distribution

The Sentinel products are originally packed in the Standard Archive Format for Europe (SAFE),<sup>12</sup> a standard compliant with the Open Archival Information System (OAIS) in terms of long time preservation, but not readily FAIR compliant. In SAFE, information can generally be divided into two main categories: the instru-

<sup>6</sup> <https://sentinel.esa.int/web/sentinel/missions/collaborative>.

<sup>7</sup> <http://lustre.org/>.

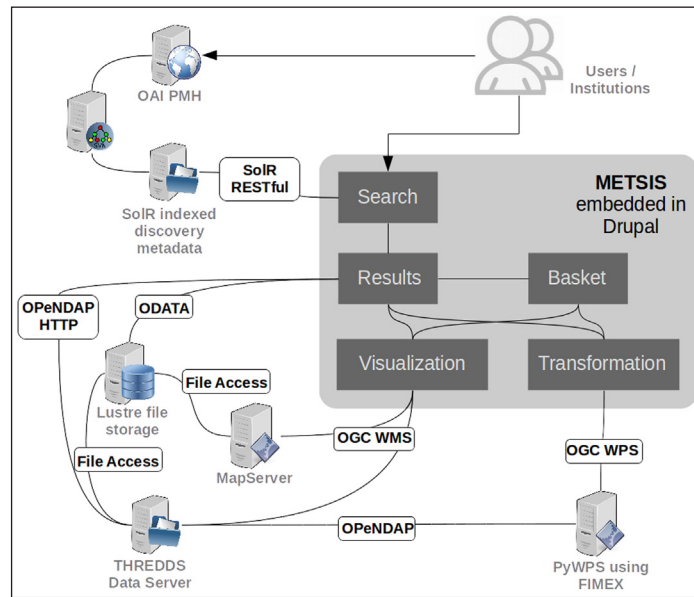
<sup>8</sup> <https://www.openstack.org/>.

<sup>9</sup> <https://www.drupal.org/drupal-7.0>.

<sup>10</sup> <https://gcmd.nasa.gov/DocumentBuilder/defaultDif10/guide/index.html>.

<sup>11</sup> <http://www.dcc.ac.uk/resources/metadata-standards/iso-19115>.

<sup>12</sup> <http://earth.esa.int/SAFE/>.



**Figure 1:** Technology infrastructure diagram in the NBS dissemination setup. All functionality within the system is available for the users through METSIS which contains multiple modules embedded in Drupal. The most common work flow is to use the search module to search for products. The module queries the indexed discovery metadata in SolR and lists the results on a new page. Data access points and services are given in dedicated elements in the discovery metadata. It is also possible to harvest the discovery metadata directly using OAI PMH. The METSIS modules access data and services by means of the dedicated elements in the discovery metadata.

ment measurements and the metadata. The instrument measurements from e.g. SAR and the Multi-Spectral Instrument (MSI) onboard the Sentinel-1 and Sentinel-2 satellites respectively, are stored in various file formats within the SAFE structure such as GeoTiff and JPEG 2000. Metadata and product auxiliary information are stored in eXtensible Markup Language (XML) files. SAFE for Sentinel data is based on the XML Formatted Data Units (XFDU) within the restrictions made for the EO domain.

In the first version of the NBS setup, focus have been put on performing a lossless transformation of data from SAFE to NetCDF-4<sup>13</sup> (Network Common Data Form v. 4). NetCDF is an open-source software, developed and supported by UCAR's Unidata Program,<sup>14</sup> that allows for creation, access and sharing of scientific data through machine-independent formats. Traditionally, NetCDF has been heavily used in communities working with atmospheric and ocean modelling, but the widespread of use of the format is growing in terms of branches and communities (Hankin et al. 2018). NetCDF-4 results from a collaboration with the group developing HDF5 (Hierarchical Data Format version 5) and makes use for the same data management layer (Folk et al. 2011). HDF5 is the prescribed file format for standard data products in EO missions operated by NASA (MuQun, Robert & Mike 2005).

NetCDF and HDF5 are powerful tools in terms of data storage and sharing. However, a certain level of both discovery and use metadata is needed to understand and interpret the content of the files. Hence, the NBS has focused on following the Climate and Forecasts convention (CF)<sup>15</sup> and the Attribute Convention for Data Discovery (ACDD).<sup>16</sup> If combined with the CF convention, the NetCDF-4 data files are self-describing and machine readable. The CF convention defines how earth science data should be encoded in terms of data structure and use metadata (Rew et al. 2019; Brian et al.). When following the ACDD convention, the NetCDF products also contains global metadata, i.e. metadata applicable for all the data in the dataset, in order to be compliant with standardized discovery metadata formats for geophysical data as e.g. ISO19115 and GCMD DIF. HDF5 files produced in the context of EO are often referred to as HDF5-EOS (EO System) where EOS defines a convention similar to CF for NetCDF.

<sup>13</sup> <https://www.unidata.ucar.edu/software/netcdf/>.

<sup>14</sup> <https://www.unidata.ucar.edu/>.

<sup>15</sup> <http://cfconventions.org/>.

<sup>16</sup> [http://wiki.esipfed.org/index.php/Category:Attribute\\_Conventions\\_Dataset\\_Discovery](http://wiki.esipfed.org/index.php/Category:Attribute_Conventions_Dataset_Discovery).

Serving NetCDF data through a server supporting the Data Access Protocol (DAP), the utilization of OPeNDAP (Open source Project for a Network DAP)<sup>17</sup> gives access to data stored at remote locations through data streaming, removing the need for downloading data prior to their usage and implementing efficient data reduction by the data provider. Also, creation of virtual datasets by means of aggregation, i.e. combining multiple datasets, and/or subsetting of a potentially large dataset, becomes readily possible (Hankin et al. 2018). As a result of collaborative efforts between Unidata and OPeNDAP, a tight fusion of NetCDF with the DAP is achieved as both use the generic data model: UNIDATA's Common Data Model<sup>18</sup>. In the current NBS setup, data access is provided using Unidata's Thematic Real-time Environmental Distributed Data Services (THREDDS) Data Server (TDS). In addition to OPeNDAP, the TDS client-service architecture also supports Open Geospatial Consortium Web Map Service (OGC WMS) using ncWMS, HTTP File Server and NetCDF Subset Service.

### 2.2.2 Data Visualization and Transformation Services

For implementation of services within the system, the open standards provided by OGC have been utilized. The OGC WMS standard for geo-referenced scalable data is used for visualization. In addition to the ncWMS client-service on the TDS, a MapServer (version 7.0.0.) have been configured to fit the project purposes like visualization of high spatial resolution RGB imagery from Sentinel-2. Visualization of the data is carried out in an OpenLayers version 3 client. For all predefined quick-looks (e.g. the RGBs from Sentinel-2), JPEG is chosen as data format. These quick-looks are stored locally in product raster resolution in contrast to the quick-looks inherent in SAFE which are thumbnails (i.e. low spatial resolution imagery) of the product.

Exploiting OPeNDAP, an OGC Web Processing Service (WPS) has been implemented using the framework of pyWPS<sup>19</sup>. The transformation process, running on a separate virtual machine in a private cloud environment, allows for product reprojection, resampling, reformatting and subsetting of temporal/spatial extent and variables. Concerning reprojection and reformatting, all target projections supported by Proj<sup>20</sup> can be implemented and currently supported output formats are NetCDF and GeoTIFF. Only registered users have access to this service due to back-end resource management and asynchronous delivery of final products to the users. The main underlying software carrying out the transformation is FIMEX,<sup>21</sup> a C/C++ software developed by MET Norway performing file manipulation, interpolation and extraction on gridded geospatial data. The software is build around the Unidata CDM and supports conversion to/from WMO GRIB as well. Hence, the NetCDF products on the TDS can be accessed using OPeNDAP and transformed according to the user request.

### 2.2.3 Discovery Metadata

The system relies on discovery metadata records storing information about the scientific datasets. This is the cornerstone of the whole system where all the necessary information for carrying out the steps described above are present by means of semantic notation. The metadata records are following the MET Norway MetaData format (MMD), a standard compliant with GCMD DIF, ISO19115/ISO19139 – standards imposed by WMO and Norge Digitalt/INSPIRE, but also extends these by adding dedicated elements like configuration metadata for OGC WMS. The purpose of the format is to document datasets and not web services, i.e. it is not made for documentation of e.g. REST services. Information on the web services available for a dataset are provided through a dedicated element. As described in **Figure 1**, all available metadata records are ingested in an Apache SolR<sup>22</sup> database and queried/filtered by means of a search module developed in METSIS. SolR is used for a large number of web applications, offering faceted searching, spatio-temporal querying and text search, among others. It is also flexible in terms of providing access to metadata across different cores (i.e. project specific repositories of metadata). In addition, the MMD records are pushed to an Apache subversion server for version control, where the discovery metadata is served through an OAI-PMH<sup>23</sup> client for other institutions to harvest.

<sup>17</sup> <https://www.opendap.org/>.

<sup>18</sup> <https://www.unidata.ucar.edu/software/thredds/current/netcdf-java/CDM/>.

<sup>19</sup> <http://pywps.org/>.

<sup>20</sup> <http://proj4.org/>.

<sup>21</sup> <https://wiki.met.no/fimex/start>.

<sup>22</sup> <http://lucene.apache.org/solr/>.

<sup>23</sup> <https://www.openarchives.org/pmh/>.

### 3. Results

Initially, the primary focus of NBS has been on delivering all Sentinel products through the DHuS while simultaneously making Sentinel-2 Level-1C and Sentinel-1 Ground Range Detected (GRD) products available in the NBS setup (i.e. as NetCDF). Concerning the utilization of DHuS in the CGS, MET Norway currently retrieves the GRD, Single Look Complex, level 2 Ocean & raw Sentinel-1 products data from the Sentinel retrieval and dissemination nodes provided by ESA, along with the Sentinel-3 land products. The marine products from Sentinel-3 (i.e. sea-surface topography, sea-surface temperature etc.) are produced and disseminated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), and hence retrieved utilizing their distribution network. Sentinel-2 data are retrieved from the Deutsches Zentrum für Luft- und Raumfahrt (DLR), a partner in CGS. In addition, NOSA has made an agreement with Kongsberg Satellite Services (KSAT)<sup>24</sup> and ESA through the CGS to provide the NBS with Sentinel-1 GRD Quasi Real-Time products, i.e. products processed and delivered no later than one hour after image acquisition. KSAT is an under contractor of ESA for reading and processing Sentinel data at their ground stations in Svalbard, Norway and Inuvik in Canada, and hence utilizing the CGS by reading and processing Sentinel data over Norwegian areas of interests during the satellite overpass, while not disturbing the systematic operations of the Copernicus ground segment. MET Norway is also the primary distribution node for Sentinel-1 data in CGS, thus disseminating global Sentinel-1 products through a dedicated mirror site for CGS partners. All Sentinel products covering the Norwegian areas of interest are ingested to the high-performance storage system and made available for the public through colhub.met.no. A structured overview of the data available in NBS is listed in **Table 1**.

In the NBS setup, which relies on discovery metadata records and satellite products in NetCDF, the initial focus has been put on making Sentinel-1 GRD and Sentinel-2 L1C available. Thus, in-house file format conversion tools for each of the Sentinel missions using Python (Python Software Foundation. Python Language Reference, version 2.7. Available at <http://www.python.org>) with standard libraries as e.g. gdal, numpy and NetCDF4 were developed to generate Sentinel products in NetCDF (Halsne and Dinessen 2018). The tools traverse the Sentinel product structure by means of the XFDU profile and manipulate the variables (i.e. interpolates raster bands and tables) in order to create a NetCDF-4/CF product. For Sentinel-2 L1C products, the frequency bands (13 in total) appear in three categories of spatial resolution, i.e.  $10 \times 10$  m,  $20 \times 20$  m and  $60 \times 60$  m. In the NetCDF version, all frequency raster bands are resampled to the highest resolution. In order to retain the original data in SAFE from NetCDF without losing information, a nearest neighbor interpolation was chosen to support lossless downsampling to the native resolution. All the view and solar angles were also resampled to fit with the dimensions of the frequency bands. As for the current setup, cirrus and opaque cloud information also follows this strategy. Cloud information is originally stored as Geographic Markup Language (GML) files in SAFE, but rasterized as variables in NetCDF to be compliant with CF (version 1.6). The layering of all these variables in NetCDF is carried out to add support for subsetting and aggregation directly. Hence, the products are processed to a higher level following the principle of Analysis Ready Data (Dwyer et al. 2018).

In general, Sentinel-1 GRD products in NetCDF are read and manipulated in the same way as for Sentinel-2. However, the differences in modality and processing level give rise to differences in the products. E.g. Sentinel-1 SAR is an active sensor illuminating the earth with a transmit pulse of 5.405 GHz. The portion of the energy backscattered from the earth surface is the basis of the resulting radar image. Since the instrument is side looking, the backscattered values are usually normalized to the geometry in which unit area the user is interested in. In SAFE, these values are systematically added as calibration vectors in XML annotation

**Table 1:** Schematic overview of all products currently available in SAFE format in NBS as for 2019-30-01.

Mission	Product Types	Retrieval Node	Approximate Number of Products	Approximate Size (PB)
Sentinel-1	RAW, SLC, GRD, OCN	ESA Node 1, Node 2, Node 3	600k	0.8
Sentinel-2	L1C, L2A	DLR	740k	0.7
Sentinel-3	SLSTR, OLCI, Synergy, Altimetry	EUMETSAT, ESA	210k	0.1
Total			1550k	1.6PB

<sup>24</sup> <https://www.ksat.no/>.

files. However, traversing the annotation files to retrieve all this information and make it ready in raster resolution requires quite some effort and understanding of the product structure. Hence, all these vectors are interpolated and made ready as variables in raster resolution in the NetCDF version. In addition, thermal noise matrices and sub swath numbers are rasterized as variables in NetCDF, the first generated according to the thermal de-noising procedure developed by the Sentinel-1 Instrument Processing Facility (Riccardo, Nuno & Hajduch 2017) and the latter according to the specifications within the SAFE product. Hence, subsetting and aggregation is supported for all these variables.

For data visualization, using ncWMS directly from TDS is not sufficient to meet the project needs. The default projection in the OL3 client is EPSG:32661, WGS84/Up North. The Sentinel-1 GRD products are projected to a WGS84 earth ellipsoid model. Hence, ncWMS has to calculate the target projection coordinates from the spherical latitude and longitude coordinates. Thus, the time latency for ncWMS to calculating the grid points is too large (i.e. in the order of one minute) in order to offer a good and responsive service for the user. Consequently, quick-looks for each polarization are stored in EPSG:32661 in advance and used for data visualization through the MapServer client, as shown in **Figure 1**. Instead of using the original unsigned 16 bit data type an GeoTIFF file format, an 8 bit color depth and JPEG was used to decrease the time latency for product visualization. For Sentinel-2 products, the same approach in terms of preprocessing quick-looks in JPEG was used, generating three RGB composites per product available for product visualization. Map projection was not an issue here and visualization of the RGB composites was just an add-on to each of the raster layers which are available through ncWMS. However, visualization of the RGB quick-looks in JPEG are faster compared with the single raster bands due to the selection of data type and the compression in JPEG.

All the NetCDF satellite products are available through the METSIS search module embedded in Drupal. The module exposes a user-friendly interface where offering tools and boxes to constrain the search by means of filtering for e.g. cloud coverage, temporal and/or spatial extent, satellite platform and free text. The final query is carried out on the discovery metadata records in SolR and the resulting products are listed on a results page in the web portal. Functionality for sharing searches between users is already developed and will be deployed in the next upgrade. For each product listed in the search result, the end user can choose to explore the discovery metadata, visualize the product in the OL3 client, transform the product or download the entire product directly in NetCDF using HTTP File Server on TDS or the SAFE version utilizing ODATA. The METSIS basket module allows for collecting multiple products in a virtual shopping cart where all the above mentioned operations can be carried out on all selected products simultaneously, as shown in **Figure 1**.

#### 4. Discussion

The NetCDF versions of both the Sentinel-1 GRD and Sentinel-2 L1C products are processed to a higher level in terms of preparing the data for further analysis. ESA has developed free and open source toolboxes for the scientific exploitation of EO missions. The toolboxes for Sentinel-1, Sentinel-2 and Sentinel-3 are accessible in a shared architecture called The Sentinel Application Platform (SNAP)<sup>25</sup>. SNAP has a wide variety of applications and processing algorithms implemented for post-processing of Sentinel products, as well as support for selected legacy missions as the ENVISAT ASAR. E.g. oil spill detection and wind field estimation from SAR data is possible in addition to lower level data operations as e.g. calibration and thermal noise correction. Despite all the functionality and tools available, the existence of e.g. scientific communities having their own algorithms for processing higher order products should address the issue of decoupling data processing steps from specific tools. The Geospatial Data Abstraction Library (GDAL) has developed a reader of Sentinel SAFE products in e.g. Python, but large portions of the product information and variables are missing. Hence, making e.g. solar angle, thermal noise and calibration information available in raster resolution makes it possible to carry out product subsetting and aggregation directly, but also makes the Sentinel products more generic in terms of processing platform independence. Batch processing of Sentinel products is already possible by either utilizing the Graph Processing Framework or Snappy – the Python implementation of SNAP – embedded in SNAP. However, for many applications it is convenient to avoid this step in-between reading and analysis of the data (Holmes 2018). In addition, as we stand before one of the most emerging challenges within scientific research today, i.e. dealing with data-intensive research, addressing the need of generic software tools as well as generic data management systems with support for heterogeneous data is great (Hey, Tansley and Tolle 2009). In the literature, this is referred to as “The Fourth

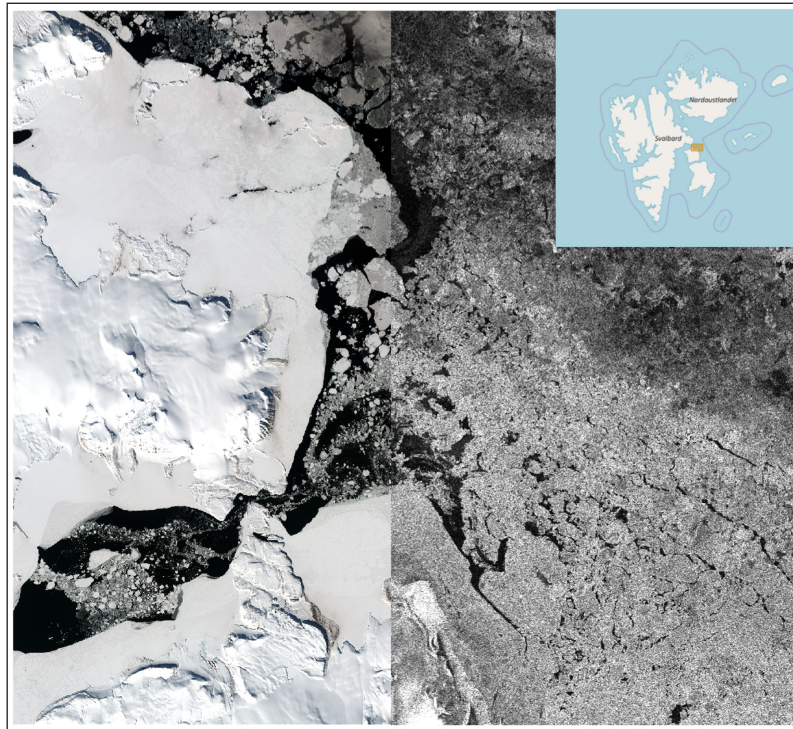
<sup>25</sup> <http://step.esa.int/main/toolboxes/snap/>.

Paradigm” within scientific research or “Big Data” in the most recent years. This is related to what is called the three V’s, i.e. data “Volume, Velocity and Variety”, denoting the growing volume of data, the speed at which the data becomes available and the great variety of branches and access mechanisms covered by the data (Laney 2001). Being able to combine different types of datasets, like EO, in-situ observations and model data, is of great importance within geophysical science to achieve the full potential in terms of synergy, data validation and data veracity (Teillet et al. 2007). The barrier for a user wanting to combine e.g. in-situ observations and EO data can include diversity in data formats, data delivery chains, data access points and lack of description. Hence, preparing data and aligning discovery, access and exploitation of different datasets in one system might reduce this barrier. The underlying data management system in NBS supports an open data space approach facilitating access to the data and services through the dedicated elements in the discovery metadata. To give an example: Land surface temperature measurements are single point observations spanning in a one-dimensional time domain. The data is made FAIR by means of utilizing the support of discrete sampling geometries in CF using timeSeries as feature type and disseminating the data through a server supporting DAP. Hence, the NBS system is generic in terms of supporting heterogeneous data and the data can be accessed directly through OPeNDAP. Moreover, integration of other geoscientific products such as model data or higher order EO products is easily achievable. For instance, numerical weather prediction models processed at MET Norway is already under a free an open policy, disseminated in NetCDF/CF and shared on a TDS. In general, the data flow in the system is based on a distributed architecture where data can be geographically distributed around the world, but accessed through OPeNDAP directly. Thus, data duplication is not necessary due to the use of common protocols, services and APIs. This, however, comes with a cost if low bandwidth and/or downtime at the data hosts is a problem (Cinquini et al. 2014).

The Copernicus Sentinels find their respective places in the chronology of EO satellite missions. In parallel, while being an operational program for future decades, some of the missions are also a continuation of EO measurements from other, heritage, satellite missions allowing for long time-series analysis. For instance, the Sentinel-2 constellation complements and extends the Landsat mission with a new combination of frequencies and spatial resolutions. The same applies for the Sentinel-3 and MODIS missions. The Sentinel-1 continues and complements the ESA European Remote Sensing (ERS) one & two missions, the ENVISAT ASAR mission, and the Canadian RADARSAT one & two missions. C-band SAR has proven very good for monitoring sea ice and hence sea ice drift at high latitudes. It is thus possible to build up a relatively long time-series of continuous sea ice drift measurements from these missions. For this particular application, inter-comparison and verification of the satellite derived drift products with in-situ observations is very important. Drifters and floaters are often used for this purpose. These devices are only producing single point observations and thus covers only a very small portion of the satellite swaths. Using OPeNDAP on the Sentinel-1 GRD products in NetCDF/CF, the verification procedure could extract a geographic subset of the relevant variables like the HV polarisation, thermal noise, relevant calibration coefficient and sub swath number for the Sentinel-1 and collocate with the in-situ observation by means of e.g. interpolation in time and space. This is not readily possible using the Sentinel-1 SAFE products since the thermal noise, calibration coefficient and sub swath number variables are not provided as raster data but in various look-up tables. Additionally, since the most common methodologies for computing drift in image analysis is by means of cross-correlations, this comes with a computational cost which is reduced when applying subsetting. The standardizing of in-situ measurements in terms making observations and the data management according to the FAIR principles is a great task in itself with a lot problems to overcome. However, standards like the WMO Integrated Global Observing System (WIGOS) is addressing this and hence contributing to FAIRness by means of standardizing both the discovery metadata for the observations conditions (like measurement location and instrument specific details) and the use metadata for the observations (WMO 2017). Combining this with data dissemination in NetCDF/CF (using relevant feature types) and thus OPeNDAP would benefit both the user community and the data publishers in terms of providing data according to the FAIR guiding principles. In particular, this addresses the I and R in FAIR, i.e. interoperability and reusability, in terms of integrating heterogeneous data and keeping track of data provenance (Wilkinson et al. 2016).

Building services using OGC technologies has proven to be of high value (Cinquini et al. 2014). After searching for data, visualization and transformation are the two services that can be carried out on either single products, or multiple products utilizing the basket interface. In particular, the latter would allow for co-visualization of Sentinel-1 SAR and Sentinel-2 MSI products, as shown in **Figure 2**. In this case, the images are acquired over areas with land fast and drifting ice in the east part of Spitzbergen with approximately 2.5 hours time delay. The land fast sea ice does not move much but drifting sea ice can be revealed in the center of the image. Co-visualization can be of great value for inter-comparison of certain features (e.g.

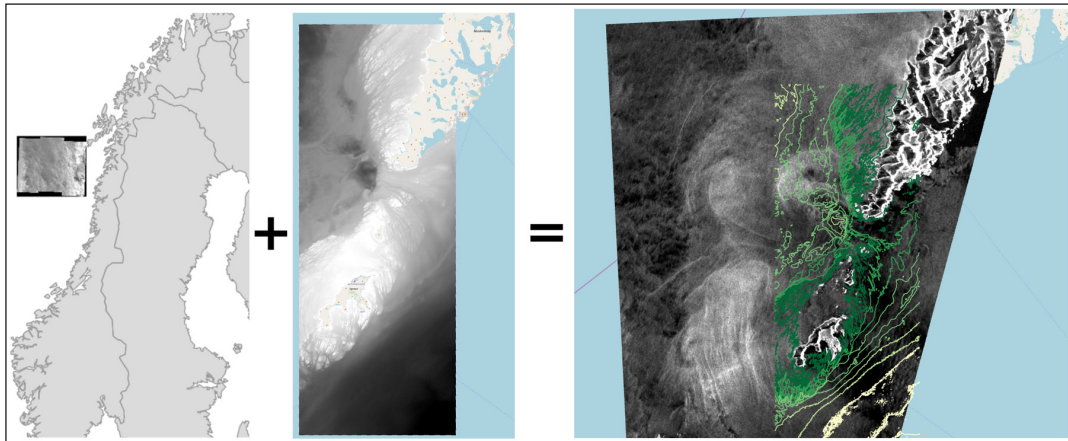




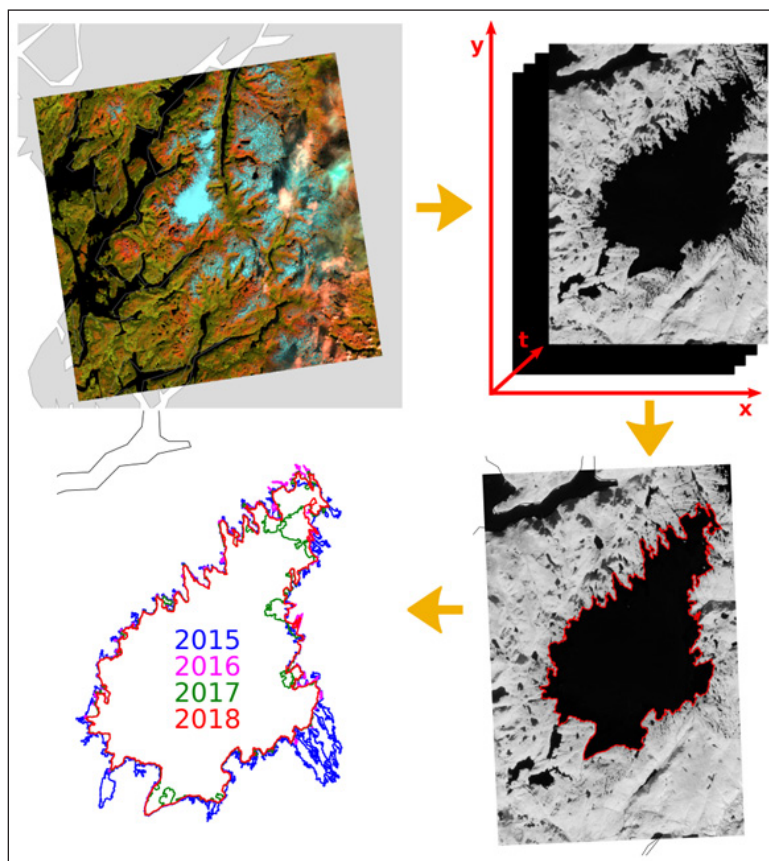
**Figure 2:** Co-visualization of Sentinel-2 MSI true color vegetation composite (left) and Sentinel-1 Interferometric Wide Swath SAR VV polarization data (right) covering Einhyrningbukta between the East coast of Spitzbergen and The Barents Island in Svalbard. Both images are acquired 29<sup>th</sup> April 2018 with approximately 2.5 hours time delay and visualized through the basket functionality. The land fast ice does not move but the drift of sea ice can be seen in the center of the image.

sea ice in this case) in different image modalities. In addition, visualization of satellite products in high resolution is a valuable service for e.g. ship navigation in ice-covered seas and route planning for snow mobile in remote locations (Dierking 2013). There are currently no limitations on how many products than can be visualized simultaneously, hence visualizing entire satellite passes and creating virtual mosaics is possible. For more quantitative studies, one might utilize the transformation service in order to align products. In addition to spatio-temporal and variable subsetting, FIMEX also supports re-projection of data in satellite swath geometry, i.e. interpolating latitude and longitude values in spherical coordinates to a Cartesian projection. An example of an oceanographic application is shown in **Figure 3** where a subset of Sentinel-1 GRD data is extracted and aligned with ocean bathymetry data provided by The Norwegian Mapping Authorities in EPSG:25833 projection by means of FIMEX. SAR is well known for detection of ocean features which in coastal areas often is related to the ocean bathymetry. In this example, you can see a plume feature generated between the island of Mosken and the cape of Lofoten. The strong Lofoten Maelstrom is mainly caused by the pressure gradient put up by the diurnal tidal cycle due to the funnel shape of the Lofoten peninsula. The ocean bottom topography states the boundary conditions for the ocean currents, and large topographic gradients has high impact on the velocity fields (Gjevik, Moe and Ommundsen 1997). In terms of data management for further processing, the size ratio of the output NetCDF subset and original Sentinel-1 SAFE product was 6 %. Hence both time and disk space is saved by the end user.

Within geophysical science, researchers will rather prefer to concentrate their efforts on using data, both for model development and data processing, than on data uploading and retrieving operations, which are nowadays easily achievable by exploiting OPeNDAP (Holmes 2018). To a large extent, commonly used programming languages, like Python, R, Matlab, C and Fortran, have modules that support reading of data from OPeNDAP. Hence, distributing data through DAP reduces both time and disk consumption for the user by means of data streaming. In **Figure 4**, an example of efficient utilization of OPeNDAP from a research perspective is shown. The example objective is to decide the spatial extent of the Folgefonna glacier in Norway and store the result as a shapefile using open source tools in Python (v. 2.7). Due to the predefined tiles of Sentinel-2 L1C data, an end user can aggregate a virtual product spanning three dimensions i.e. x, y and time building a virtual data cube with Analysis Ready Data – a very strong concept in terms of time-series analysis (Dwyer et al. 2018; Lewis et al. 2017). The data cube concept is also further strengthened by means of the



**Figure 3:** The result of subsetting and aligning a Sentinel-1 GRD Interferometric Wide Swath SAR product by means of FIMEX with ocean bathymetry data provided by The Norwegian Mapping Authorities in EPSG:25833 projection. Left image show the extent of the Sentinel-1 scene in the OL3 visualization client embedded in the NBS dissemination system. In the middle, ocean bathymetry data is plotted in a GIS tool in gray-scale where black and white denotes deep and shallow waters respectively. Right image show the subset from Sentinel-1 SAR swath product covering the Lofoten peninsula in northern Norway aligned with the ocean bathymetry data overlaid as contours. For the contours, the green gradient goes from dark to bright indicating shallow and bright respectively. In both the middle and right image, the products are overlaid on an OpenStreetMap background layer.



**Figure 4:** Processing steps from extracting the extent of the Folgefonna glacier in Norway at different times from subsetting of Sentinel-2 L1C NetCDF-4/CF files using OPeNDAP. Upper left, upper right, lower right and lower left show 1) coverage of tile T32VLM in the OL3 client in a false color composite, 2) concept of temporal aggregating of products covering the same area, 3) the Folgefonna glacier extent decided by means of density slicing from the computation of the Normalised Difference Snow Index and 4) polygons showing the glacier extent from four different years in the same plot, respectively.

Analysis Ready Data processing of the Sentinel-2 products in NetCDF/CF since additional variables are available in raster resolution. In this particular example, the number of pixels used to calculate the Normalized Difference Snow Index is 0.2% of all pixels available in the product. In addition, the frequency bands used for the calculation, i.e. band B3 and B11, are originally offered in  $10 \times 10$  m and  $20 \times 20$  m pixel resolution. Currently, OAI-PMH is the only API implemented for Data Discovery. However, the project aims on integrating OpenSearch on top of SolR. Hence, retrieval of all OPeNDAP data access URL's covering the same area will be easy achievable utilizing OpenSearch for building the virtual time stack, as shown in the upper right plot of **Figure 4**, which is carried out manually by means of adding URLs in a numpy array in this example. The source code for computing the glacier extent is open and available as a github gist.<sup>26</sup> An extension of this example, where the users also would gain from subsetting and transformation for aligning data, could be to align the Sentinel-2 glacier data either with altimetry data or a digital elevation model for calculation of the glacier volume.

In summary the NBS takes care of discovery, access, visualization, transformation and long time preservation of Sentinel data. There are, however, challenges. Although the CF convention gets more and more mature, an easy way of adding plain text, e.g. XML annotation and auxiliary files for Sentinel products, is not in place in NetCDF. There are however workarounds like e.g. splitting text into string variables, although this might not be an optimal solution. However, recent development of the CF convention (current development version is 1.8) makes it more flexible and suitable for EO data by e.g. adding support for geometries. In addition, standardized encoding of Earth Science swath data in instrument viewing geometry in CF is under development and a proposal was provided in late 2017.<sup>27</sup> On the other side, the current CF version does only support a flat structure in NetCDF. However, NetCDF has support for a hierarchical structure and the utilization of groups in NetCDF would simplify handling of satellite data and make them more lucid in terms of product structure. The intention of CF version 2.0 is to add support for groups and other features of the underlying storage model in NetCDF. Currently, the conversion of Sentinel-1 and Sentinel-2 data is not lossless in terms of preserving all information from SAFE in NetCDF. However, this is the ultimate goal, and further development of the CF convention, supporting more of the flexibility in NetCDF, underpins this. Thus, SAFE is currently used for the archiving of Sentinel data. But when a lossless conversion process is in place, SAFE products could be generated from the NetCDF files and the latter format could, for some of the products, be used as archive format.

Regarding national services for Norway, the Quasi Real-Time Sentinel-1 products are of great value when it comes to near real-time monitoring of the large areas covered by open ocean outside the coastline of Norway along with the Svalbard Archipelago and its surrounding areas. E.g. both monitoring ship traffic and early detection of oil spills from SAR are valuable supplemental services for the Norwegian Coastal Administration and their work (Grønnestad 2016). Another national service being implemented address the problems related to the elevation models used for the Sentinel-1 products, which has proven to yield poor results in the steep mountainous areas at high latitudes areas. This is the case for large portion of the surface topography on the Norwegian mainland. Hence, a national project has been initiated for building a processor for ortho-rectifying Sentinel-1 GRD products. These products will be disseminated through the NBS system. How they will be disseminated has however yet to be decided. Since this is a processor, products can be processed on demand by setting up a dedicated WPS available for the users in the NBS setup. Thus, the products could be ordered and disseminated by means of the basket directly to the user. Alternatively, the products could also be ingested and disseminated through the TDS and hence be available through OPeNDAP. Ingestion and dissemination through DHuS is an option. The latter option could provide a new product available for all DHuS systems in CGS, but at a cost of products being less FAIR compliant compared with the NetCDF/CF versions. Moreover, the two latter options using TDS or DHuS should lead to building up a data base where all available products are ortho-rectified since users may be confused if only some products are available. Hence, storage space will be an issue founding the premise for further discussion. Both the above mentioned services are of national importance and is unlikely to be served as a service on a European level. Hence, the need for covering national services and interests are important arguments for building both national infrastructure and competence to deal with these types of problems.

Due to the large volume of data, there are also emerging interests in batch processing on large amounts of data for creation of higher order products. In addition to have high storage capacity, MET Norway accesses

<sup>26</sup> <https://gist.github.com/hevgyrt/d5d29f5b2d691d8ed3c4b3cd65e2009e>.

<sup>27</sup> <https://github.com/Unidata/EC-netCDF-CF/blob/master/swath/swath.adoc>.

a private cloud used for post-processing of data with dedicated nodes for GPU and CPU computing. Hence, offering a Virtual Research Environment as a part of the NBS system is feasible and supports the concept of moving processing close to the data for limiting costs, I/O operations and duplication of data since large amounts of Sentinel data are stored locally (Balasubramonian and Grot 2016). However, algorithms for some operational services and retrieval of parameters from EO as e.g. sea-ice extent, land deformation and crop damage, are not critical to run on national IT infrastructure and could potentially be run in a cloud processing environment somewhere else, like on one of the DIAS'es which will be up and running in the future. The DIAS'es will also offer several supplemental datasets to the Sentinel missions like data from The Copernicus Services (ESA 2017). However, products with national specifications, such as the above mentioned ortho-rectified Sentinel-1 products, are unlikely to be served in other places. Therefore, a cost-benefit analysis is required to figure out how to align national initiatives like the NBS with European, or outside European, initiatives to maximize the exploitation of the systems the best way possible.

## 5. Conclusion and Future Remarks

The NBS project takes care of discovery, access, visualization, transformation and long time preservation of Sentinel data covering Norwegian areas of interest. Disseminating various Sentinel products using NetCDF-4/CF enhances product value and the data can easily be accessed, combined and subsetted by means of OPeNDAP. Additionally, the products have larger catchment in terms of making more variables available in product raster resolution in a relatively common file format supported by a large number of tools and software's compared with SAFE. Disseminating Sentinel data through the NBS setup allows for easy integration with non-EO data, thus achieving more of the potential in terms of synergy, data validation and data veracity. Future development of the CF standard makes the format more suitable for EO products and thus enhance the quality of the NetCDF4/CF versions of Sentinel data.

To relate the NBS to other ongoing EO infrastructure projects in a global context, it is of particular interest to further investigate a potential future exploitation of global infrastructure for running algorithms in a cloud processing environment. However, further processing of specific products is required in order to overcome problematic regional conditions like the steep surface topography in certain areas in Norway. These products are unlikely to be processed and disseminated in a global scale. Hence, building national competence and infrastructure is important in order to offer the best suitable products, and services, covering the Norwegian areas of interest.

The first milestones in future development will be to: i) Add Sentinel-3 Ocean and Land Cover Imagery (OLCI) products along with Sea and Land Surface Temperature Radiation (SLSTR) products which already are in NetCDF. However, in contrast to the Sentinel-3 level 2 products processed by EUMETSAT, level 1 products do not follow the CF convention and must be re-configured in order to fit optimal in the NBS system. ii) Integrate an OpenSearch API on top of the discovery metadata in SolR. iii) Further development on the utilization of MapServer to visualize no-data values as a transparent layer. iv) Retrieve and implement the software taking care of Sentinel-1 ortho-rectification. v) Identify the need for providing a Virtual Research Environment for national users as a supplement to the services provided by commercial contractors as e.g. the DIAS'es.

## Acknowledgements

This research activity is supported by the Norwegian Space Agency from establishing the National Ground Segment ("Nasjonalt Bakkesegment", NBS) for satellite data in Norway.

## Competing Interests

The authors have no competing interests to declare.

## Authors Contributions

- Halsne, Trygve – lead development and implementation of The NBS dissemination system.
- Ferrighi, Lara – developing various models for Drupal in satellittdata.no and takes care of system configuration.
- Saadatnejad, Bard – lead METSIS developer.
- Budewitz, Nico – main responsible on running Norway's node in the Sentinel Collaborative Ground Segment.
- Dinessen, Frode – contributes to the development of the SAFE to NetCDF conversion tools.
- Breivik, Lars-Anders – project management in NBS and remote sensing expert.
- Godoy, Øystein – project leader and expert within data management and remote sensing.

## References


- Balasubramonian, R** and **Grot, B.** 2016. Near-data processing. *IEEE Micro*, 4–5. DOI: <https://doi.org/10.1109/MM.2016.1>
- Brian, E, Jonathan, G, Bob, D, Karl, T, Steve, H, Jon, B, John, C, Rich, S, Phil, B, Greg, R, Heinke, H, Alison, P, Martin, J** and **Martin, R.** *NetCDF Climate and Forecast (CF) Metadata Conventions*. Available at <http://cfconventions.org> [Last accessed 16 11 2019].
- Cinquini, L, Crichton, D, Mattmann, C, Harney, J, Shipman, G, Wang, F, Ananthakrishnan, R, Miller, N, Denvil, S, Morgan, M, Pobre, Z, Bell, GM, Doutriaux, C, Drach, R, Williams, D, Kershaw, P, Pascoe, S, Gonzalez, E, Fiore, S** and **Schweitzer, R.** 2014. The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data. *Future Generation Computer Systems*, 36: 400–417. DOI: <https://doi.org/10.1016/j.future.2013.07.002>
- Cristóbal, J, Ninyerola, M** and **Pons, X.** 2008. Modeling air temperature through a combination of remote sensing and GIS data. *Journal of Geophysical Research*, 113. DOI: <https://doi.org/10.1029/2007JD009318>
- Dierking, W.** 2013. Sea Ice Monitoring by Synthetic Aperture Radar. *Oceanography*, 26. DOI: <https://doi.org/10.5670/oceanog.2013.33>
- Dwyer, JL, Roy, DP, Sauer, B, Jenkerson, CB, Zhang, HK** and **Lymburner, L.** 2018. Analysis Ready Data: Enabling Analysis of the Landsat Archive. *Remote Sensing*, 10. DOI: <https://doi.org/10.20944/preprints201808.0029.v1>
- ESA.** 2017. *Accessing Copernicus Data Made Easier*. Available at [https://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Accessing\\_Copernicus\\_data\\_made\\_easier](https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Accessing_Copernicus_data_made_easier) [Last accessed 16 01 2019].
- Fielding, RT** and **Taylor, RN.** 2000. Architectural styles and the design of network-based software architectures. Irvine, USA: University of California.
- Folk, M, Heber, G, Koziol, Q, Pourmal, E** and **Robinson, D.** 2011. An Overview of the HDF5 Technology Suite and Its Applications. *Proceedings of the EDBT/ICDT 2011 Workshop on Array Databases*, 36–47. DOI: <https://doi.org/10.1145/1966895.1966900>
- Gjevik, B, Moe, H** and **Ommundsen, A.** 1997. Sources of the Maelstrom. *Nature*, 388: 837–838. DOI: <https://doi.org/10.1038/42159>
- Gorelick, N, Hancher, M, Dixon, M, Ilyushchenko, S, Thau, D** and **Moore, R.** 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202: 18–27. DOI: <https://doi.org/10.1016/j.rse.2017.06.031>
- Grønnestad, KS.** 2016. Nyttige satellittar. *Barentswatch*. Available at <https://www.barentswatch.no/artikler/nyttige-satellittar/> [Last accessed 16 01 2019].
- Halsne, T** and **Dinessen, F.** 2018. `safe_to_netcdf`. *GitHub repository*. Available at [https://github.com/hevgyrt/safe\\_to\\_netcdf.git](https://github.com/hevgyrt/safe_to_netcdf.git). (commit: 027f3ffd0106289f84a523c48e36abc5b64c4177) [Last accessed 18 11 2019].
- Hankin, SC, Blower, J, Carval, T, Casey, K, Donlon, C, Lauret, O, Loubrieu, T, Srinivasan, A, Trinanes, J, Godøy, Ø, Mendelssohn, R, Signell, R, De La Beaujardiere, J, Cornillon, P, Blanc, F, Rew, R** and **Harlan, J.** 2018. NETCDF-CF-OPENDAP: standards for ocean data interoperability and object lessons for community data standards processes.
- Hansen, MC** and **Loveland, TR.** 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122: 66–74. DOI: <https://doi.org/10.1016/j.rse.2011.08.024>
- Hey, T, Tansley, S** and **Tolle, K.** 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery. *Microsoft Research*.
- Holmes, C.** 2018. *Analysis Ready Data Defined*. Available at <https://medium.com/planet-stories/analysis-ready-data-defined-5694f6f48815> [Last accessed 16 01 2019].
- Laney, D.** 2001. 3D data management: controlling data volume, velocity and variety. *META Group Research Note*.
- Lewis, A, Oliver, S, Lymburner, L, Evans, B, Wyborn, L, Mueller, N, Raevksi, G, Hooke, J, Woodcock, R, Sixsmith, J, Wu, W, Tan, P, Li, F, Killough, B, Minchin, S, Roberts, D, Ayers, D, Bala, B, Dwyer, J, Dekker, A, Dhu, T, Hicks, A, Ip, A, Purss, M, Richards, C, Sagar, S, Trenham, C, Wang, P** and **Wang, L-W.** 2017. The Australian Geoscience Data Cube – Foundations and lessons learned. *Remote Sensing of Environment*, 202: 276–292. DOI: <https://doi.org/10.1016/j.rse.2017.03.015>
- MuQun, Y, Robert, EM** and **Mike, F.** 2005. HDF5 – a high performance data format for earth science. *21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography and Hydrology*.
- National Research Council.** 2007. *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond*. Washington, DC: The National Academies Press.

- National Research Council, Mathae, KB and Uhler, PF** (Ed.). 2012. *The Case for International Sharing of Scientific Data: A Focus on Developing Countries: Proceedings of a Symposium*. Washington, DC: The National Academies Press.
- Rew, R, Davis, G, Emerson, S, Davies, H, Hartnett, E, Heimbigner, D and Fisher, W.** 2019. *NetCDF User's Guide*. Available at <https://www.unidata.ucar.edu/software/netcdf/docs/index.html> [Last accessed 16 01 2019].
- Riccardo, P, Nuno, M and Hajduch, G.** 2017. *Thermal Denoising of Products Generated by the S-1 IPF*. Available at <https://sentinel.esa.int/documents/247904/2142675/Thermal-Denoising-of-Products-Generated-by-Sentinel-1-IPF> [Last accessed 16 01 2019].
- Tillet, PM, Chichagov, A, Fedosejevs, G, Gauthier, RP, Ainsley, G, Maloley, M, Guimond, M, Nadeau, C, Wehn, H and Shankaie, A.** 2007. An integrated earth sensing sensorweb for improved crop and range-land yield predictions. *Canadian Journal of Remote Sensing*, 33: 88–98. DOI: <https://doi.org/10.5589/m07-012>
- Turner, W, Rondinini, C, Pettorelli, N, Mora, B, Leidner, AK, Szantoi, Z, Buchanan, G, Dech, S, Dwyer, J, Herold, M, Koh, LP, Leimgruber, P, Taubenboeck, H, Wegmann, M, Wikelski, M and Woodcock, C.** 2015. Free and open-access satellite data are key to biodiversity conservation. *Biological Conservation*, 182: 173–176. DOI: <https://doi.org/10.1016/j.biocon.2014.11.048>
- Westermann, S, Østby, TI, Gislås, K, Schuler, TV and Etzelmüller, B.** 2015. A ground temperature map of the North Atlantic permafrost region based on remote sensing and reanalysis data. *The Cryosphere*, 9: 1303–1319. DOI: <https://doi.org/10.5194/tc-9-1303-2015>
- White-Newsome, JL, Brines, SJ, Brown, DG, Dvonch, JT, Gronlund, CJ, Zhang, K, Oswald, EM and O'Neill, MS.** 2013. Validating satellite-derived land surface temperature with in situ measurements: a public health perspective. *Environmental health perspectives*, 121: 925–931. DOI: <https://doi.org/10.1289/ehp.1206176>
- Wilkinson, MD, Dumontier, M, Aalbersberg, IJJ, Appleton, G, Axton, M, Baak, A, Blomberg, N, Boiten, J-W, da Silva Santos, LB, Bourne, PE, Bouwman, J, Brookes, AJ, Clark, T, Crosas, M, Dillo, I, Dumon, O, Edmunds, S, Evelo, CT, Finkers, R, Gonzalez-Beltran, A, Gray, AJG, Groth, P, Goble, C, Grethe, JS, Heringa, J, 't Hoen, PAC, Hooft, R, Kuhn, T, Kok, R, Kok, J, Lusher, SJ, Martone, ME, Mons, A, Packer, AL, Persson, B, Rocca-Serra, P, Roos, M, van Schaik, R, Sansone, S-A, Schultes, E, Sengstag, T, Slater, T, Strawn, G, Swertz, MA, Thompson, M, van der Lei, J, van Mulligen, E, Velterop, J, Waagmeester, A, Wittenburg, P, Wolstencroft, K, Zhao, J and Mons, B.** 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3: 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>
- World Meteorological Organization.** 2017. *WIGOS Metadata Standard*. Available at [https://library.wmo.int/doc\\_num.php?explnum\\_id=3653](https://library.wmo.int/doc_num.php?explnum_id=3653) [Last accessed 18 11 2019].
- Wulder, MA and Coops, NC.** 2014. Satellites: Make Earth observations open access. *Nature*, 513: 30–31. DOI: <https://doi.org/10.1038/513030a>
- Wulder, MA, Coops, NC, Roy, DP, White, JC and Hermosilla, T.** 2018. Land cover 2.0. *International Journal of Remote Sensing*, 39: 4254–4284. DOI: <https://doi.org/10.1080/01431161.2018.1452075>
- Yang, J, Gong, P, Fu, R, Zhang, M, Chen, J, Liang, S, Xu, B, Shi, J and Dickinson, R.** 2013. The role of satellite remote sensing in climate change studies. *Nature Climate Change*, 3: 875–883. DOI: <https://doi.org/10.1038/nclimate1908>

**How to cite this article:** Halsne, T, Ferrighi, L, Saadatnejad, B, Budewitz, N, Dinessen, F, Breivik, L-A and Godøy, Ø. 2019. The Norwegian National Ground Segment; Preservation, Distribution and Exploitation of Sentinel Data. *Data Science Journal*, 18: 61, pp. 1–14. DOI: <https://doi.org/10.5334/dsj-2019-061>

**Submitted:** 31 January 2019    **Accepted:** 26 November 2019    **Published:** 27 December 2019

**Copyright:** © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 