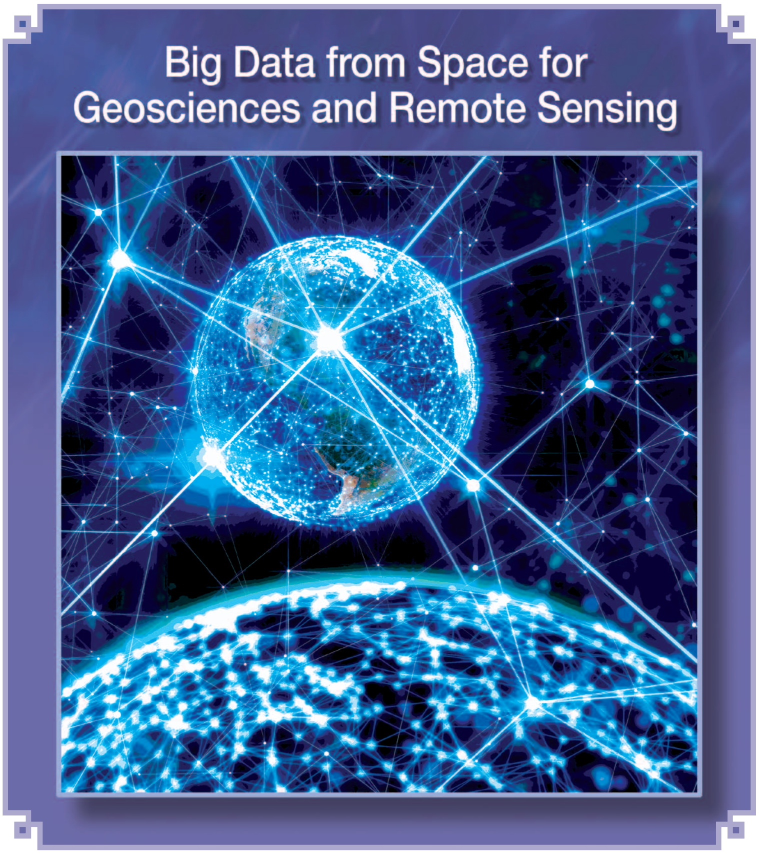
Big Data Management in Earth Observation

The German satellite data archive at the German Aerospace Center



©ISTOCKPHOTO.COM/NADLA

STEPHAN KIEMLE, KATRIN MOLCH, STEPHAN SCHROPP, NICOLAS WEILAND, AND EBERHARD MIKUSCH

The German Satellite Data Archive (D-SDA) at the German Aerospace Center (DLR) has been managing large-volume Earth-observation (EO) data in the context of EO-mission payload ground segments (PGSs) for more than two decades. Hardware, data management, processing, user access, long-term preservation, and data exploitation expertise are under one roof and interact closely. Upcoming EO-mission PGSs benefit as much from the comprehensive expertise, close interaction, and integrated infrastructure as do in-house scientific application projects requiring access, processing, and archiving of large-volume EO data. Using a number of examples, we will demonstrate how EO data life cycles benefit from the proximity of data management and application scientists and from the extensive operational experience gathered over time.

Digital Object Identifier 10.1109/MGRS.2016.2541306
Date of publication: 20 September 2016

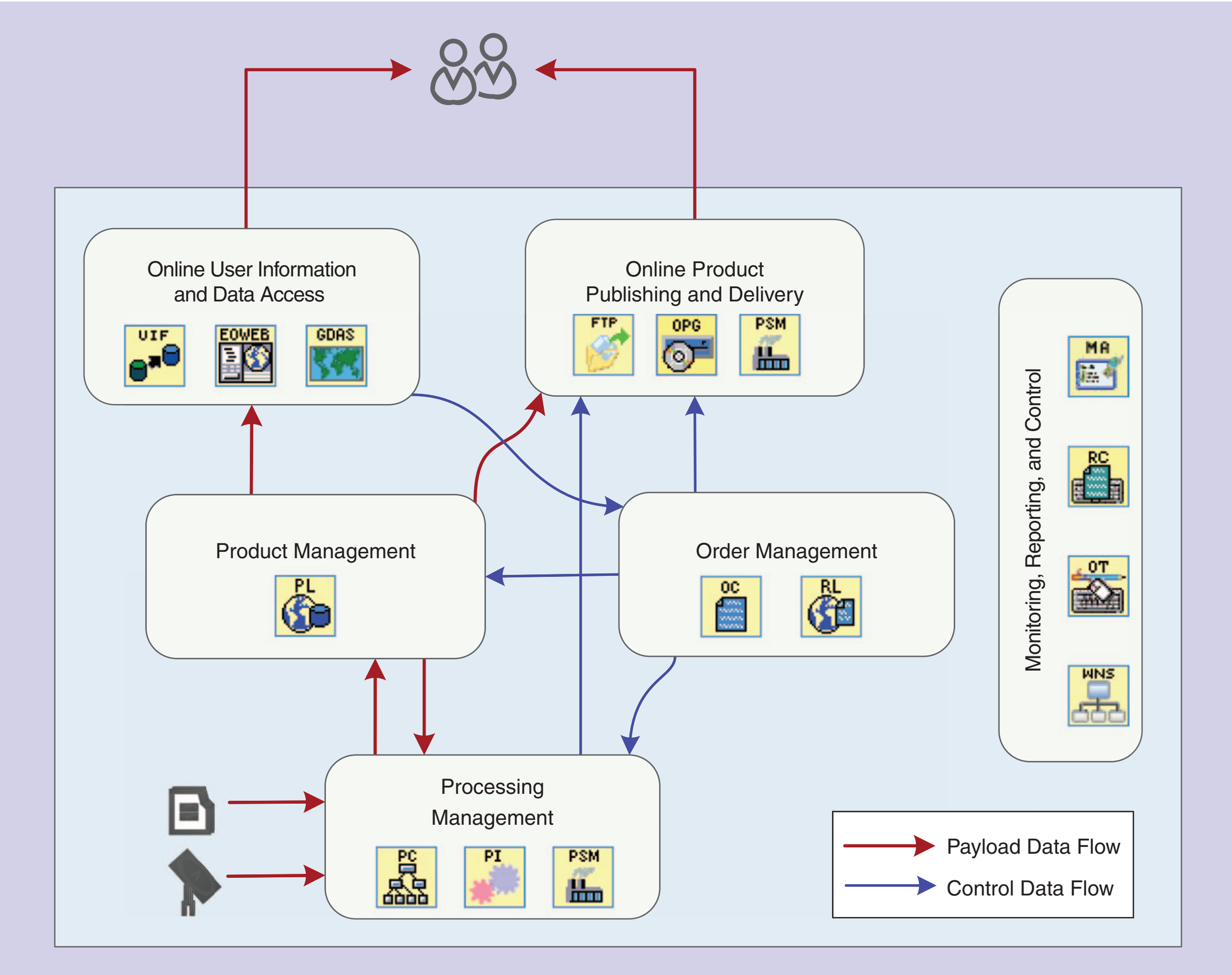


FIGURE 1. The functional components and modules of the DIMS.

CHARACTERISTICS OF THE GERMAN SATELLITE DATA ARCHIVE

The D-SDA, which is operated at the German Remote Sensing Data Center of DLR, is a central German infrastructure providing large-volume EO data archiving and access functionality. For many years, it has been a key component in the PGS of numerous national and European EO missions, such as the Shuttle Radar Topography Mission (SRTM), TerraSAR-X, TanDEM-X, and the environmental satellite ENVISAT. Additionally, it serves as a data center for scientific and civil-service EO applications. The D-SDA currently holds approximately 6 PB of EO data and spatial information products in its robotic tape library, the key component of the hierarchical storage system having a total capacity of 50 PB. In response to current mission requirements, 170 TB of online cache accelerate the data access by the processing systems with maximum throughput rates in the range of 6 TB per day. This, as well as the transfer of data to external users, is assisted by a high-capacity local area (10 Gb/s) and wide area (1 Gb/s) network infrastructure.

DATA AND INFORMATION MANAGEMENT

Integrated with the storage systems and network infrastructure, the Data and Information Management System (DIMS) (developed by DLR in collaboration with an industrial partner [12]) handles all aspects of EO data management in a multimission environment. The DIMS is a highly modular system with individual components that handle tasks in processing management, order management, and online product publishing and delivery (Figure 1). The components encapsulate the range of functionality required for efficiently managing EO data and ensure flawless data ingestion into the archive as well as delivery to the processing systems and users.

Two essential components of the DIMS are the product library (PL) and the processing system management (PSM). The PL manages the individual archived EO products in a structured database model that combines primary data and metadata into one information package. Metadata extracted from the archive database are used for populating the product catalog that is then published via a standardized catalog service. The PSM component controls complex processing workflows and is used to set up and orchestrate distributed tasks and services. As a comprehensive, integrated system, the DIMS ensures dedicated and swift data flows between archives, processing systems, and users during active EO missions. Data management at DLR, however, extends beyond the mission lifetime, encompassing the entire data life cycle. The archived EO data continue to be curated and valorized after the mission has ended, ensuring long-term accessibility and usability.

In response to the initiation of the Copernicus Sentinel missions and the maturity of a number of EO data exploitation chains, e.g., in atmospheric remote sensing, the user requirements for EO data access and processing are changing. Users request the direct download of large data volumes and access to hosted processing capabilities. Functionality, which was previously closely integrated with the data-management infrastructure, is moving closer to the end user who expects flexible interaction. In response, the D-SDA has recently upgraded its online data access capacity to 0.5 PB. In the near future, the D-SDA will be further integrated with hosted processing infrastructures, extending the current data-processing capabilities of subsetting, reprojecting, and reformatting during data download.

Collaborating closely with and driven by the needs of the user community, a mission that PGS is inherently concerned with managing the data content and applications as it accompanies the data life cycle. The following three examples will introduce and discuss specific competencies accumulated at the D-SDA over time. These are considered essential for providing comprehensive and efficient EO data-management services in the future.

PAST EXPERIENCE FACILITATING THE SETUP OF A BIG DATA PAYLOAD GROUND SEGMENT

On behalf of the European Space Agency (ESA), DLR develops and operates the Sentinel-5 Precursor PGS [1]. In the mission PGS, the D-SDA provides archiving and data access services for lower-level and higher-level processing systems. Long-standing experience with a variety of aspects and scenarios of EO data management, especially the dynamic data modeling and workflow management, facilitate the PGS setup and operations for this new mission.

The Sentinel-5 Precursor is an atmospheric mission that will generate products for monitoring air quality, ozone concentrations, ultraviolet radiation, and climate variables [2]. In setting up the PGS, DLR system engineers were able to draw on extensive experience with atmospheric data workflows and with user-centered data modeling in support of these specific workflows, both of which were gained from close interactions with scientists developing complex retrieval algorithms to be transferred into the operational processing environment. This experience was built up over time with other long-lasting projects such as the Satellite Application Facility for Atmospheric Composition and UV Radiation (O3M SAF) of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) [3], for which near-real-time trace gas retrieval, offline processing, product quality assurance, and reprocessing workflows are being developed and operated. Surpassing the O3M SAF requirements, the processing workflows within the Sentinel-5 Precursor PGS are based on numerous product interdependencies and make intensive use of auxiliary data. Figure 2 illustrates the complex network of dependencies and the auxiliary data used for generating atmospheric composition products.

THE DIMS IS A HIGHLY MODULAR SYSTEM WITH INDIVIDUAL COMPONENTS HANDLING TASKS IN PROCESSING MANAGEMENT, ORDER MANAGEMENT, AND ONLINE PRODUCT PUBLISHING AND DELIVERY.

A generic scheduling engine on a large, distributed computing facility is not sufficient for controlling this type of workflow. In addition, specific EO metadata are required for selecting the input data and for verifying the consistency of the output data, as well as knowledge on the behavior of the individual processor modules to optimize the data throughput and product quality. This allows the scientists to specify the numerical complexity of algorithms and the capability for parallelization. Based on this input, the system engineers plan the assignment of system resources for computation and data transfer. In addition, they design intelligent scheduling schemes, taking into account the various input dependencies and resource constraints to meet the timeliness requirements for the output products as defined by the mission or project management.

The engineers of the Sentinel-5 Precursor PGS specifically strive for a swift and efficient transfer of increasingly large volumes of input data and processed products between the archive and the processing systems. The German national radar EO-mission TanDEM-X is equally demanding in terms of large-volume data management and systematic processing involving large-volume data transfers. Data transfer rates reach a maximum of 10.2 TB per day during peak processing phases [4], which has been made possible not only by setting up a dedicated 10-Gb/s fiber-channel-based data transfer local area network, but also by scaling up the storage area network (SAN) storage systems. More spinning disks are now used in parallel via additional partitioning, which allows for a higher throughput. With similar throughput requirements, as detailed in Table 1, the design and setup of the Sentinel-5 Precursor PGS were based on these experiences gained during the TanDEM-X mission.

The DLR PGS system engineers' knowledge about ESA-specific ground segment architectures and interfaces for discovery and access facilitated the setup of the infrastructure for the Sentinel-5 Precursor. Similar components had previously been developed for the Copernicus Space Component Data Access and the long-term archives for Sentinel-1 and Sentinel-3. As a result, comparable solutions were pursued

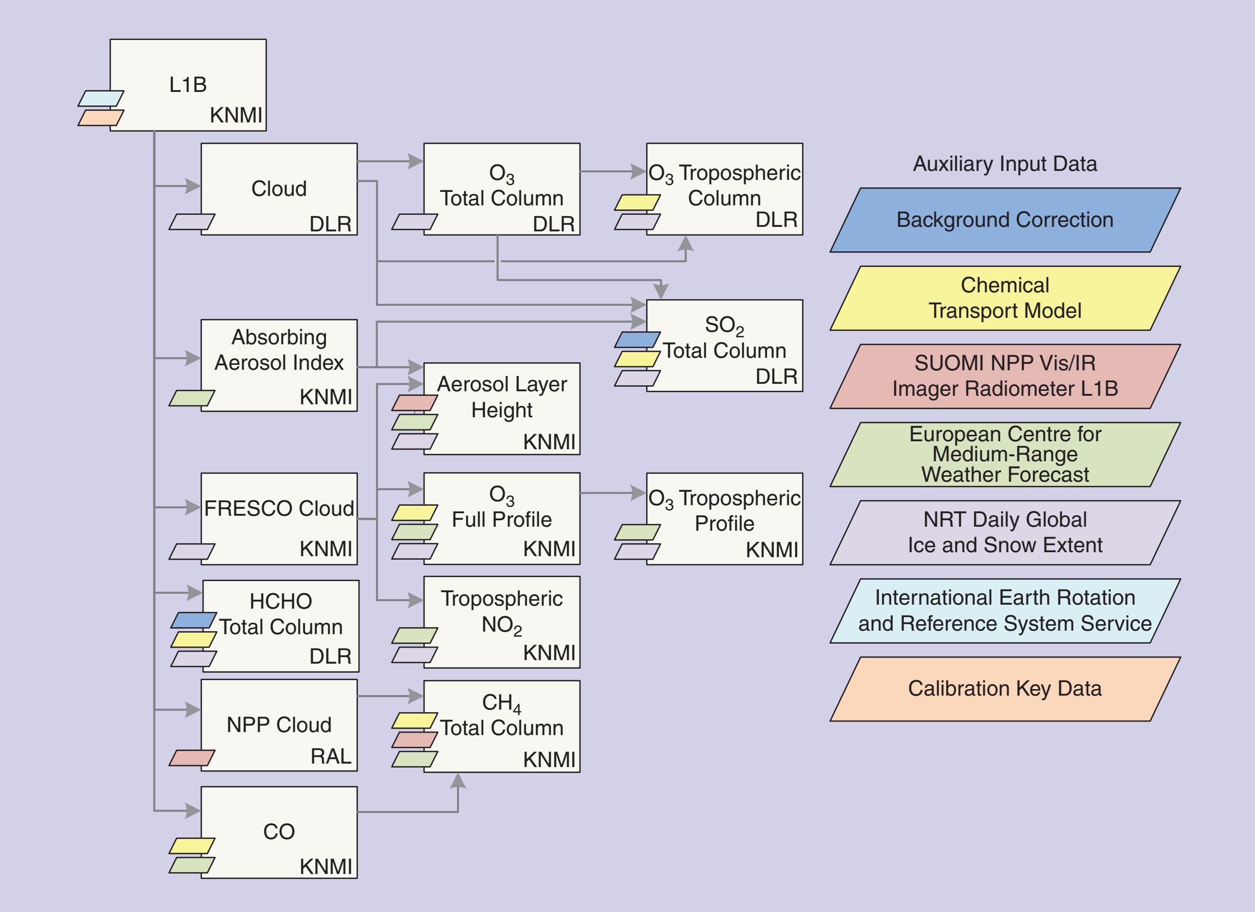


FIGURE 2. The expected product interdependencies, auxiliary data, and organizational responsibilities for generating level-two atmospheric composition products in the Sentinel-5 Precursor mission PGS. KNMI: Koninklijk Nederlands Meteorologisch Instituut (Royal Netherlands Meterological Institute); FRESCO Cloud: fast retrieval scheme for cloud from oxygen A band; NRT: near real time; HCHO: formaldehyde; CO: carbon oxide; O_3: ozone; NO_2: nitrogen dioxide; L1B: level 1B.

for integrating DLR- and Copernicus-wide area networks for all three missions, i.e., the Sentinel-1, Sentinel-3, and Sentinel-5 Precursors. The PGSs of the Sentinel Precursor missions are being monitored closely. A systematic reporting of system activities, such as the data processing completeness, auxiliary data usage, product quality, and service performance, is requested by the ESA as the contracting entity. The D-SDA has established a reporting control service that systematically generates reports based on the information collected from all PGS components.

TABLE 1. THE SENTINEL-5 PRECURSOR DAILY DATA THROUGHPUT.

DATA TYPE	DAILY DATA THROUGHPUT
Level 0	250 GB
Level 1b near real time	510 GB
Level 1b offline	905 GB
Auxiliary	100 GB

TIMELINE: THE SERVICE-DRIVEN DATA MANAGEMENT IN A PGS-STYLE ENVIRONMENT AND LESSONS LEARNED

In support of monitoring global change, the DLR TIMELINE project will generate a consistent 30-year timeline of 18 environmental variables over Europe, including sea-surface temperature, snow cover, and cloud properties, from the 1-km data of the Advanced Very High Resolution Radiometer flown onboard the series of Polar Orbiting Environmental Satellites of the American National Oceanic and Atmospheric Administration. To support the TIMELINE project, the D-SDA gradually extends its PGS data-management functionality toward servicing large-volume archiving and access service requirements of scientific projects.

Thus, the PGS-focused D-SDA is now more closely collaborating with scientific users, which requires a mutual adaptation of approaches, architectures, processes, and procedures. The project highlighted the advantages of scientists, system engineers, and operations engineers working under one roof. Collaborating closely and exchanging views without delay allowed for a coherent design and operations of the complete service chain, from data acquisition to archiving via basic and thematic processing to product access. Additionally, existing PGS data-management components could be reused in a different but technically related context. Another advantage is the common use of the DIMS Operating Tool, which is a unified graphical user interface for system operators supporting the configuration, monitoring, and control of PGS components. With the help of this tool, operational procedures can be defined and shared across multiple missions.

During the initial project phases and the close interaction between data managers and application scientists, processor developers, and system operators, existing constraints and the difference in approaches and cultures on either side were recognized. A PGS is usually set up and operated based on clearly defined roles, requirements, and procedures that are laid down in a series of technical documents. Any modifications to the system require a formal change of the defined procedures. The scientific context of this project, however, asked for a more flexible approach, involving, e.g., on the technical side, the ability to frequently change or upgrade the systems. In response, a distinct test and integration environment for the preparation and verification of system upgrades was set up in parallel to the stable operational environment. Thus, data models and processing schemas could be enhanced and validated during the project in an iterative approach without impeding operational processing.

On the organizational side, defining and assigning clear roles and responsibilities to each project participant improved the efficient collaboration between scientists and data-management engineers. Thus, a mutually agreed view on roles and responsibilities was established that resulted in a constructive division of duties. The overall project coordination and communication within the project benefitted from the clear responsibilities, which also furthered the progress of the project. The roles and related responsibilities as established within the TIMELINE project are provided in Table 2.

On the technical side, the D-SDA data management has evolved toward a more flexible, service-driven archiving and access infrastructure in response to the project's requirements. The improvements include creating a project-specific customization of the generic D-SDA data portal EO on the web Geoportal, offering project-specific discovery and download services, as well as developing new interfaces for bulk data handling. These two sample improvements are briefly described in the following paragraphs.

With an intention to further increase interoperability, the project provided an incentive for the D-SDA to move toward standardizing product metadata based on the EO Metadata Profile of Observations and Measurements as defined by the Open Geospatial Consortium (OGC) [5]. Standard OGC catalog and web mapping services were also implemented and are accessible by any compatible client. Standardized catalogs and spatial data services allow the project to better promote its results within the scientific community. Scientists can more easily discover data sets of interest for their application and integrate them smoothly into their system for further analysis. The assignment of digital object identifiers [8] to TIMELINE products further facilitates discovery and citability.

THE SENTINEL-5 PRECURSOR IS AN ATMOSPHERIC MISSION THAT WILL GENERATE PRODUCTS FOR MONITORING AIR QUALITY, OZONE CONCENTRATIONS, ULTRAVIOLET RADIATION, AND CLIMATE VARIABLES.

The large-volume data flows between the services in support of the TIMELINE project are powered by the DIMS. The DIMS supports specific data-handling functions such as data subscription based on specific metadata and bulk data retrieval from the archive. Randomly accessing data that is archived on tape for processing is inefficient because tapes have to be loaded and positioned. Bulk retrieval, as implemented in the DIMS, makes use of the metadata and the processing order defined by scientists for creating and organizing large data requests. With these requests, the data are efficiently prestaged, i.e., transferred from the tape to the disk cache in bulk for fast access by the processing system. The DIMS supports the configuration of individual product data models adapting to any kind of mission, sensor, and processing level by allowing metadata structures that are specific to the individual product type. Engineers can select standard metadata structures or reuse the structures defined in previous missions to reduce redundancy and implement best practices.

An additional challenge faced during the TIMELINE project involved the conversion of scientific processing algorithms into large-volume PGS-style processing chains while providing sufficient flexibility for ad hoc modifications as required by the scientists. Specific issues arise from the different cultures of scientific and operational PGS-style data processing. The project scientists are predominantly concerned with developing and fine tuning algorithms for accurate thematic product generation, which involves frequent test runs on sample data sets for algorithm verification and validation. For systematic, PGS-style processing, these algorithms then have to be converted into processing tools and ultimately into high-capacity, high-efficiency operational processing systems for deployment in DLR's high-capacity computation infrastructure GeoFarm, which consists of virtualized machines on clustered servers with SAN systems. Experience shows that this transfer is best done by an information technology (IT) specialist within the scientific application environment, whose proximity and understanding of the data and processing algorithms and close interaction with the scientists will result in short evolution cycles and add to the flexibility of the resulting processing system.

USER COMMUNITY-DRIVEN, LONG-TERM DATA PRESERVATION AND ACCESS

Beyond providing data-management services to current missions and projects, the D-SDA provides sustainable, long-term data accessibility. EO data are unique snapshots of the condition of Earth or the atmosphere at a specific point in time. DLR and the D-SDA, therefore, put particular emphasis on long-term data preservation with the objective to keep the valuable data and products accessible and usable for future generations. Preservation is an inherent part of the data life cycle, with its scope extending beyond keeping the instrument data safe from loss.

THE PROJECT SCIENTISTS ARE PREDOMINANTLY CONCERNED WITH DEVELOPING AND FINE TUNING ALGORITHMS FOR ACCURATE THEMATIC PRODUCT GENERATION.

One aspect for ensuring long-term data accessibility in the D-SDA is its stable and sustainable archive infrastructure. The hardware and software are maintained and upgraded on a regular basis following technology evolution cycles. Tape media, which for economic and safety reasons today are still widely adopted for archiving with hierarchical storage management at the petabyte scale, are being replaced by a new generation of tapes at intervals of about five to seven years. Tape drives and archive servers follow a similar technology evolution cycle, while the tape library is being replaced approximately every 15–20 years. The modules of the data-management software are maintained in-house on a continuous basis and upgraded as new technology or protocols emerge.

On the data side, active data curation contributes to long-term usability. As new requirements arise—from technology evolution, application scientists, or end users—the data are migrated; converted to new formats; reprocessed using improved algorithms; transformed into user-friendly, higher-level products; and made accessible via the required interfaces. The data curation is based on an appraisal of the data sets and a preservation plan detailing activities and assessing associated cost and risks.

TABLE 2. THE ROLES AND RESPONSIBILITIES DEFINED WITHIN THE TIMELINE PROJECT.

ROLE	DETAILS AND RESPONSIBILITIES
Scientist	• Is a member of the research department • Develops processing/exploitation algorithms • Performs algorithm tests on a limited sample data set • Provides reference data(base) if needed • Provides level-four support
Tool developer	• Is usually a software engineer within the research department • Transfers the algorithm into an operational tool • Performs testing of the tool on a larger number of sample data sets • Provides level-three support
Software engineer	• May be with the research department or the IT department • Has knowledge about existing modules and overall system architecture • Encapsulates and transfers tools into operational modules and processing chains • Tests modules/chains with a significant number of sample data sets • Delivers modules to system engineer for approval and integration • Provides level-two support
System engineer	• Is a member of the IT department • Analyzes requirements in collaboration with scientist and software engineer • Is responsible for the main data repository and for the system framework • Decides on the development of modules/processing chain • Approves new modules from software engineers • Performs integration tests in a semi-operational test environment • Is responsible for module/processing chain versioning and deployment
Data librarian	• Is responsible for the reference databases • Verifies and makes available reference databases received from scientists, tool developers, or software engineers
Operational system engineer	• Is part of the ground segment department or with the customer • Installs the system received from the system engineer on the production site • Performs operational system tests with a large number of sample data (~1,000 scenes) • Approves readiness for mass production
Operations engineer	• Is part of the ground segment department or with the customer • Runs the system in the operational environment • Is responsible for quality assurance • Provides level-one support

For data sets to remain usable for future generations, preserving the instrument data and metadata is not sufficient. The curation of the associated knowledge, such as mission and sensor documentation, data structure and format specifications, or calibration and processing information, is as essential as maintaining the capability to visualize and reprocess the data [6], [7]. For new missions, appropriate data-management principles, including the generation of the associated information, should already be established during the planning and preparation phases. The D-SDA stores product metadata within the individual archival information package as components of the data model. Collection metadata are stored in a separate database. The concepts for preserving the associated documentation foresee a central document management with access provided via D-SDA servers.

In addition to implementing in its archive-sustainable data curation measures, the D-SDA contributes to developing ESA-wide harmonized and interoperable data-management principles and preservation procedures for EO data. In collaboration with partner organizations, operational know-how is being transferred into guidelines and best practices. Extending beyond Europe, these documents are being introduced to the Group on EOs, with a number of them having been approved as best practices by the Committee on EO Satellites [6]–[8].

LOOKING BEYOND EARTH OBSERVATION: FACILITATING INTEGRATED DATA EXPLORATION AND ANALYSIS

Making use of state-of-the-art technologies and implementing international standards, as described in the previous sections, the D-SDA is capable of providing data discovery and access services to the emerging collaborative exploitation platforms. In these centralized infrastructures, scientists and other users will be able to interact with data and algorithms, contribute and share information, and process and exploit data. The ability to contribute will increase the use and reuse of the EO data and products held in the D-SDA and facilitate cross-fertilization between different application communities.

Data exploitation across archives requires seamless interoperability. The use of standardized metadata formats and discovery and access protocols, such as those provided by OGC [9], the Heterogeneous Mission Accessibility Project [9], or OpenSearch [11], which are gradually being implemented in the D-SDA, become mandatory for ensuring smooth interoperability between distributed data repositories and the exploitation infrastructure. Federated data discovery across distributed Earth-science data archives and access to higher-level products based on international standards is already common. However, interoperable data formats and lower-level data-exchange formats are still in their infancy. Data preservation and curation will soon require operational solutions for scenarios such as consolidating time series across archives or for physically relocating complete historical archive holdings in case of a transfer of data-preservation responsibilities [7].

SUMMARY AND CONCLUSIONS

Big data management in EO is not just an IT issue; the responsibilities go beyond handling large data volumes, providing high-throughput processing capacity, and making available large network bandwidths for swift data access. This article has shown that EO data sets as well as the valuable scientific results are best handled by a comprehensive, competent data life cycle center in which data managers, IT engineers, and scientists closely collaborate. It is the knowledge of the data set content and the experience with data-management systems, both serving specific EO applications and workflows, that make for optimum, integrated end-to-end service chains and ensure tailored, sustainable curation of data and systems. Thus, the maximum value is extracted out of EO-mission data, during the mission lifetime and beyond, for current and future generations.

ON THE DATA SIDE, ACTIVE DATA CURATION CONTRIBUTES TO LONG-TERM USABILITY.

AUTHOR INFORMATION

Stephan Kiemle (stephan.kiemle@dlr.de) received the Dipl. degree in computer science from the Technical University of Munich, Germany, in 1995. Since 1995, he has been with the German Remote Sensing Data Center (DFD), German Aerospace Center, Oberpfaffenhofen, Wessling, Germany, where he started as a software developer for the Interactive Satellite Information System. In 1997, this work merged into the development of the DFD's multimission Earth-observation data system, the Data and Information Management System (DIMS), where he was responsible for the product library. Since 2003, he has been leading the DIMS development team and has been working as a system engineer and project manager for a couple of Earth-observation payload ground-segment and infrastructure projects.

Katrin Molch (katrin.molch@dlr.de) received her M.A. degree in geosciences from the University of Freiburg, Germany, in 1997. Since 1997, she has been working as a remote sensing scientist with Canadian and European Earth-observation research organizations and in industry. She has developed and given a number of training courses on satellite image exploitation and radar interferometry to international civilian and military clients. Since 2011, she has led the data-management services team of the satellite data archive at the German Aerospace Center. She is working toward sustainable data curation and improving accessibility to the organization's data holdings. She is a member of several national and international working groups, coordinating and harmonizing efforts in Earth-observation data curation to ensure long-term usability. Her scientific focus is on exploiting radar satellite imagery for geological, agricultural, and urban applications.

Stephan Schropp (stephan.schropp@dlr.de) received the Dipl.-Ing. degree in electrical engineering from the Technical

University of Munich, Germany, in 2008. Shortly after graduating, he joined the German Remote Sensing Data Center of the German Aerospace Center, Oberpfaffenhofen, Germany. He is responsible for the management of hardware and operating system-level software of the German Satellite Data Archive. In 2012, he also took responsibility in the field of project management of the Sentinel-5 Precursor payload data ground segment, planning and deploying the necessary communication networks and archiving facilities.

Nicolas Weiland (Nicolas.weiland@dlr.de) received the Dipl. degree in computer science from the Ludwig Maximilian University of Munich, Germany, in 2008. Since 2009, he has been with the German Remote Sensing Data Center of the German Aerospace Center, Oberpfaffenhofen, Germany. He started as a software developer for the Data and Information Management System, where he was responsible for the product library. Since 2011, he has been working as a system engineer for the payload data ground segment of the European Space Agency/European Union Sentinel-5 Precursor mission within the European Copernicus Program.

Eberhard Mikusch (eberhard.mikusch@dlr.de) received the Dipl. degree in computer science from the University of Applied Sciences in Munich, Germany, in 1983. Since 1983, he has been with the German Aerospace Center (DLR). From 1983 to 1987, he was a development engineer for the Electrical Ground Support Environment and designed DLR's image processing system for the Halley multicolor camera of the European Giotto Mission. From 1987 to 1992, he was responsible for designing the Ground Support Program Equipment of the Anthrorack experiment of the Spacelab Mission D-2. From 1996 to 2003, he lead the software development team establishing the multimission Data and Information Management System used within several Earth-observation missions including TerraSAR-X. Since 2004, he has been the head of the Information Technology Department of DLR's German Remote Sensing Data Center. His research interests include optimizing Earth-observation archiving and access infrastructures for high-volume data provision in distributed multimission environments.

REFERENCES

[1] S. Kiemle, R. Knispel, M. Schwinger, and N. Weiland, "Sentinel-5 Precursor payload data ground segment," presented at the Proc. Advances in Atmospheric Science and Applications (ATMOS), Bruges, Belgium, June 18–22, 2012, ESA Special Publication SP-708.

[2] European Space Agency. (2015, June 12). Sentinel-5P. [Online]. Available: https://earth.esa.int/web/guest/missions/esa-future-missions/sentinel-5P

[3] S. Kiemle, P. Valks, M. Boettcher, D. Loyola, W. Zimmer, T. Ruppert, and T. Erbertseder, "DLR data services for GOME-2/MetOp atmospheric trace gas monitoring," presented at the Proc. Joint 2007 EUMETSAT Meteorological Satellite Conference, Amsterdam, The Netherlands, Sept 24–28, 2007.

[4] S. Kroeger, M. Schwinger, M. Wegner, and M. Wolfmueller, "Data handling and preservation for the TanDEM-X satellite mission," in *Proc. Ensuring Long-Term Preservation and Adding Value to Scientific and Technical Data* (*PV*), Villafranca, Spain, Dec. 1–3, 2009, pp. 1–7.

[5] Open Geospatial Consortium. (2012, June 12). *Earth Observation Metadata Profile of Observations and Measurements*, v. 1.0, OGC 10-157r3. [Online]. Available: https://portal.opengeospatial.org/files/?artifact_id=47040

[6] Committee on Earth Observation Satellites, Data Stewardship Interest Group. (2015, Sept.). *Earth observation preserved data set content*, v. 1.0. [Online]. Available: http://ceos.org/document_management/Working_Groups/WGISS/Interest_Groups/Data_Stewardship/Recommendations/EO%20Preserved%20Data%20Set%20Content_v1.0.pdf

[7] Committee on Earth Observation Satellites, Data Stewardship Interest Group. (2015, Sept. 15). *EO data preservation guidelines*, v.1.1.[Online].Available:http://ceos.org/document_management/Working_Groups/WGISS/Interest_Groups/Data_Stewardship/Recommendations/EO%20Data%20Preservation%20Guidelines_v1.1.pdf

[8] Committee on Earth Observation Satellites, Data Stewardship Interest Group. (2015, Mar.). *Persistent identifiers best practices*, v. 1.0. [Online]. Available: http://ceos.org/document_management/Working_Groups/WGISS/Interest_Groups/Data_Stewardship/Best_Practices/CEOS%20Persistent%20Identifier%20Best%20Practices_v1.0.pdf

[9] Open Geospatial Consortium. (2015, Dec. 15). *OGC standards and supporting documents*. [Online]. Available: http://www.opengeospatial.org/standards

[10] T. Usländer, Y. Coene, and P. G. Marchetti, *Heterogeneous Missions Accessibility* (ESA TM-21). Noordwijk, The Netherlands: ESA Communications, 2012.

[11] P. Gonçalves and U. Voges, "OGC OpenSearch extension for Earth observation," Open Geospatial Consortium, Wayland, MA, OGC Draft Implementation Standard OGC 13-026r5, 2015.

[12] Werum Software & Systems. (2015, Dec. 21). DIMS-EO: Data information and management system for Earth observation. [Online]. Available: www.werum.de/en/platforms/dims-eo.jsp

GRS