INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2022, VOL. 68, NO. 1, PP. 167–172
Manuscript received January 17, 2022; revised January, 2022. DOI: 10.24425/ijet.2022.139864

Integration of Remote Sensing Data in a Cloud Computing Environment

Yassine Sabri, Fadoua Bahja, and Henk Pet

***Abstract*—With the rapid development of remote sensing technology, our ability to obtain remote sensing data has been improved to an unprecedented level. We have entered an era of big data. Remote sensing data clear showing the characteristics of Big Data such as hyper spectral, high spatial resolution, and high time resolution, thus, resulting in a significant increase in the volume, variety, velocity and veracity of data.This paper proposes a feature supporting, salable, and efficient data cube for time-series analysis application, and used the spatial feature data and remote sensing data for comparative study of the water cover and vegetation change. In this system, the feature data cube building and distributed executor engine are critical in supporting large spatiotemporal RS data analysis with spatial features. The feature translation ensures that the geographic object can be combined with satellite data to build a feature data cube for analysis. Constructing a distributed executed engine based on dask ensures the efficient analysis of large-scale RS data. This work could provide a convenient and efficient multidimensional data services for many remote sens-ing applications.**

***Keywords*—Remote Sensing; Data integration; Cloud Computing; Big Data;**

I. Introduction

SINCE Landsat-1 first started to deliver volumes of pixels in 1972, the amount of archived remote sensing data stored by data centers has increased continuously [1, 2]. According to incomplete statistics, the total amount of data archived by the Earth Observing System Data and Information System (EOSDIS) reached 12.1 petabytes (PBs) around the year 2015 [3]. Up until August 2017, the archived data volume of the China National Satellite Meteorological Center (NSMC) reached 4.126 PBs [4], and the China Center for Resources Satellite Data and Application (CCRSDA) archived more than 16 million scenes of remote sensing images [5, 6]. Such large amounts of remote sensing data have brought great difficulties for data integration of each data center.

Due to various satellite orbit parameters and the specifications of different sensors, the storage formats, projections, spatial resolutions, and revisit periods of the archived data are vastly different, and these differences have resulted in great difficulties for data integration. In addition, the remote sensing data received by each data center arrives continuously at an ever-faster code rate. It is preferable to ingest and archive the newly received data in order to provide users with the latest data retrieval and distribution service [7]. Therefore, a unified metadata format and a well designed data integration framework are urgently needed.

Yassine Sabri and Fadoua Bahja are with Laboratoryl of Innovation in Management and Engineering for Enterprise (LIMIE), ISGA Rabat, 27l Avenuel Oqba,Agdal,Rabat,Morocco (e-mail: yassine.sabri@isga.ma).

Henk Pet is with Terra Motion Limited, 11 Ingenuity Centre, Innovation Park, Jubilee Campus, University of Nottingham, Nottingham NG7 2TU, UK

Hence, for data integration across a distributed data center spatial infrastructure, we proposed an International Standardization Organization (ISO) 19115-based metadata transform method, and then adopted the internationally popular data system framework object-oriented data technology (OODT) [8] to complete the distributed remote sensing data integration.

The rest is organized as follows: Section 2 provides an overview of the background knowledge and related work; Section 3 describes the distributed multi-source remote sensing metadata transformation and integration; Section 4 introduces the experiments and provides an analysis of the proposed program; and Section 5 provides a summary and conclusions.

II. Background on Architectures for Remote Sensing Data Integration

This section briefly reviews the distributed integration of remote sensing data, as well as the internationally popular data system framework OODT.

A. Distributed integration of remote sensing data

The most widely used data integration models include [9]: (1) The data warehouse (DW)-based integration model, which copies all data sources of each heterogeneous database system into a new and public database system, so as to provide users with a unified data access interface. However, due to the heterogeneity of each independent database system, vast data redundancy is generated, and a larger storage space is also required. (2) The federated database system (FDBS)-based integration model, which maintains the autonomy of each database system and establishes an association between each independent database system to form a database federation, then providing data retrieval services to users. However, this pattern cannot solve the problems of database heterogeneity or system scalability [10]. (3) The middleware-based integration model, which establishes middleware between the data layer and the application layer, providing a unified data access interface for the upper layer users and realizing the centralized management for the lower layer database system. The middleware not only shields the heterogeneity of each database system, providing a unified data access mechanism, but also effectively improves the query concurrency, reducing the response time. Therefore, in this paper, we will adopt the

© The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0, https://creativecommons.org/licenses/by/4.0/), which permits use, distribution, and reproduction in any medium, provided that the Article is properly cited.

middleware-based integration mode to realize the distributed remote sensing data integration.

B. OODT: a data integration framework

An FS or DBMS alone are not suited for the storage and management of remote sensing data. In a DBMS-FS mixed management mode", remote sensing images are stored in the file system and their metadata are stored and managed by the DBMS. Typical examples are the European Space Agency (ESA) [11], Tiandi Maps of China, the CCRSDA, the NSMC, the China National Ocean Satellite Application Center (NSOAS), and so on. The mixed management mode both effectively solves the quick retrieval and metadata management problems and maintains the high read/write efficiency of the file system. This has been a longtime issue addressed by NASA, whose the Office for Space Science decided to fund the OODT project in 1998. Apache OODT [12] is an open-source data system framework that is managed by the Apache Software Foundation. OODT focuses on two canonical use cases: big data processing [13] and information integration [14]. It provides three core services: (1) a file manager is responsible for tracking file locations and transferring files from a staging area to controlled access storage, and for transferring their metadata to Lucene or Solr; (2) a workflow manager captures the control flow and data flow for complex processes, and allows for reproducibility and the construction of scientific pipelines; and (3) a resource manager handles allocation of workflow tasks and other jobs to underlying resources, based on the resource monitoring information from Ganglia or other monitoring software. In addition to the three core services, OODT provides three client-oriented frameworks that build on these services: (1) a file crawler automatically extracts metadata and uses Apache Tika or other self-defined toolkits to identify file types and ingest the associated information into the file manager; (2) a pushpull framework acquires remote files and makes them available to the system; (3) a scientific algorithm wrapper (called the Catalog and Archive Service Production Generation Executive, CAS-PGE) encapsulates scientific codes and allows for their execution, regardless of the environment, while capturing provenance, making the algorithms easily integrated into a production system (Figure 1).

III. Distributed Integration of Multi-Source Remote Sensing Data

With distributed multi-source remote sensing data integration, i.e., based on a unified standard, the remote sensing metadata in the distributed center will be gathered into the main center continuously or at regular intervals, either actively or passively. In this study, the unified satellite metadata standard refers to the ISO 19115-2:2009-based geographic information metadata standard. All of the remote sensing metadata in the distributed sub-centers should be transformed into the ISO 19115-based metadata format before integration to enable uniform data retrieval and management. The distributed sub-centers are mainly responsible for the storage of remote sensing images, and provide an open access interface for the main center based on the HTTP/FTP protocols. The main center is primarily responsible for the ingestion and archiving of the metadata and thumbnails of remote sensing images, and enables uniform query and access for the integrated remote sensing data.

A. The ISO 19115-based metadata transformation

Remote sensing metadata represent descriptive information about remote sensing images, as well as data identification, imaging time, imaging location, product level, quality, the spatial reference system, and other characteristic information. At present, the metadata forms of different remote sensing data vary greatly. For example, Landsat 8 collects images of the Earth with a 16-day repeat cycle, referenced to the Worldwide Reference System-2 [15]. The spatial resolution of the Operational Land Imager (OLI) sensor onboard the Landsat 8 satellite is about 30 m; its collected images are stored in GeoTIFF format, with Hierarchical Data Format Earth Observation System (HDFEOS) metadata [16, 17]. The Moderate-Resolution Imaging Spectroradiometer (MODIS) instruments capture data in 36 spectral bands ranging in wavelength from 0.4 μm to 14.4 μm and at varying spatial resolutions (2 bands at 250 m, 5 bands at 500 m, and 29 bands at 1 km). Most of the MODIS data are available in the HDF-EOS format, and it is updated every 1 to 2 days [18]. The chargecoupled device (CCD) sensor, which is carried by the Huan Jing (HJ)-1 mini satellite constellation, has an image swath of about 360 km, with blue, green, red, and near infrared (NIR) bands, 30-m ground pixel resolution, and a 4-day revisit period. Its collected images are stored in GeoTIFF format, and their customized metadata are in eXtensible Markup Language (XML) format [19]. These different metadata formats have resulted in great difficulties for data integration and management, which could be solved by transforming them into a uniform metadata format for uniform retrieval and management [20, 21]. ISO 19115-2:2009 is the geographic information metadata standard which was published by the International Standardization Organization (ISO). It mainly defines the metadata schema of geographic information and services, including the identification, quality, space range, time horizon, content, spatial reference system, distribution, and other characteristic information [22]. Currently, ISO 19115-2:2009 has been integrated into the Common Metadata Repository (CMR) as one of the most popular standards for data exchange [23], data integration, and data retrieval across international geographic information organizations and geographic data centers. On the basis of the ISO 19115-2:2009 geographic information standard, we proposed a uniform remote sensing metadata format. All of the remote sensing metadata in the distributed sub-centers should be transformed into this uniform format before data integration. In this paper, the transformational rules we established are mainly aimed at NASA EOS HDF-EOS format metadata (Aster and Landsat series satellites included) and the customized XML-based metadata of the CCRSDA (HJ-1A/B, GF and ZY series satellites included) (see Table I).

It should be noted that in Table I, the strike-through (-) shows the field does not exist, and it will be assigned a null value after metadata transformation. In the ISO metadata

www.czasopisma.pan.pl PAN www.journals.pan.pl
POLSKA AKADEMIA NAUK

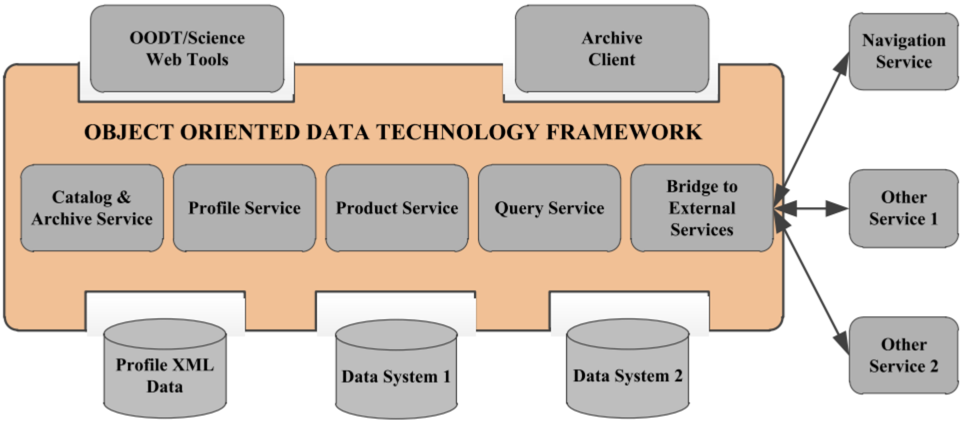

Fig. 1. An object-oriented data technology (OODT) framework

TABLE I
THE ISO 19115-2:2009-BASED UNIFORM METADATA FORMAT AND TRANSFORMATIONAL RULES. ISO: INTERNATIONAL STANDARDIZATION ORGANIZATION; CCRSDA: CHINA CENTER FOR RESOURCES SATELLITE DATA AND APPLICATION; HDF-EOS: HIERARCHICAL DATA FORMAT EARTH OBSERVATION SYSTEM.

Categories	ISO Metadata	HDF-EOS Metadata	CCRSDA Metadata
Metadata information	Creation	FILE_DATE	-
	LastRevision	-	-
Image Information	MD_Identifier	LOCALGRANULEID	-
	TimePeriod_beginposition	RangeBeginningDate+RangeBeginningTime	imagingStartTime
	TimePeriod_endPosition	RangeEndingDate+RangeEndingTime	imagingStopTime
	Platform	AssociatedPlatformShortName	satelliteId
	Instrument	AssociatedinstrumentShortName	-
	Sensor	AssociatedsensorShortName	sensorId
	Datacenter	PROCESSINGCENTER	-
	recStationId	STATION_ID	recStationId
	spatialResolution	NADIRDATARESOLUTION	pixelSpacing
	westBoundLongitude	WESTBOUNDINGCOORDINATE	productUpperLeftLong
	eastBoundLongitude	EASTBOUNDINGCOORDINATE	productUpperRightLong
	southBoundLatitude	SOUTHBOUNDINGCOORDINATE	productLowerLeftLat
	northBoundLatitude	NORTHBOUNDINGCOORDINATE	productUpperLeftLat
	centerLongtiude	-	sceneCenterLong
	centerLatitude	-	sceneCenterLat
	scenePath	WRS_PATH	scenePath
	sceneRow	WRS_ROW	sceneRow
	referenceSystemIdentifier	PROJECTION_PARAMETERS	earthModel+mapProjection
	cloudCoverPercentage	-	cloudCoverPercentage
	imageQualityCode	-	overallQuality
	processingLevel	DATA_TYPE	productLevel

column, the term spatialResolution describes the ability of the remote sensor to distinguish small details of an object, generally in meters, thereby making it a major determinant of image resolution. Hence, the spatialResolution is mapped to NadirDataResolution in the HDFEOS metadata column and pixelSpacing in the CCRSDA metadata column.

The terms scenePath and sceneRow are orbit parameters of the satellite in the Worldwide Reference System (WRS), just mapping to WRS PATH and WRS ROW in the HDF-EOS metadata column. The term imageQualityCode is a characteristic of a remote sensing image that measures the perceived image degradation, and has the same meaning as the overallQuality in the CCRSDA metadata column. The term processingLevel denotes the type of the remote sensing data, and is mapped to the DATA TYPE in the HDF-EOS metadata column and productLevel in the CCRSDA metadata column. TABLE I: The ISO 19115-2:2009-based uniform metadata format and transformational rules. ISO: International Standardization Organization; CCRSDA: China Center for Resources Satellite Data and Application; HDF-EOS: Hierarchical Data Format Earth Observation System.

B. Distributed multi-source remote sensing data integration

Distributed multi-source remote sensing data integration refers to the process of validating, inserting, updating, or deleting metadata in the main center metadata management system; it affects only the metadata for the distributed data providing sub-centers. The metadata management is mainly realized by the components of OODT, including the OODT crawler, OODT push-pull, and OODT file manager [24] (see Figure 2). In the main data center, the push-pull daemon will be launched automatically by using its daemon launcher at the defined time interval. The daemon will wrap one of two processes: (1) RemoteCrawler, or (2) ListRetriever. The RemoteCrawler process crawls remote sites for files in the distributed subcenters. Meanwhile, the RemoteCrawler process also automatically extracts metadata and transforms them into the ISO 19115-2:2009-based uniform metadata format. The ListRetriever retrieves known files from remote sites in the distributed sub-centers (that is, the path and file name to each file is known and has been specified in a property file, and a parser for that property file has been specified). After crawling or retrieval, the push-pull framework will be responsible for downloading remote content (pull), or accepting the delivery of remote content (push) to the main center for use by the LocalCrawler for ingestion into the file manager. Here, the remote content includes the metadata file and thumbnail of remote sensing data. It is worth mentioning that the Local-Crawler is developed in the main center, and is primarily responsible for crawling the local client system for files in the main center. The file manager component is responsible for tracking, ingesting, and moving metadata and thumbnails between a client system and a server system in the main center. Finally, the remote sensing metadata will be indexed by the SolrCloud, and their corresponding thumbnails will be archived in the file system. Both the RemoteCrawler and LocalCrawler have an incremental control mechanism in order to avoid duplicate data ingestion. In the intervals between crawling and data ingestion, the RemoteCrawler executes a Message Digest 5 (MD5) file verification process between the remote sites' files in the sub-center and the archived files in the main center. If the file has been archived in the main center, data ingestion will be stopped; otherwise, data ingestion continues. The LocalCrawler implements the second MD5 file verification process between the files in the client system (files from sub-centers downloaded to the main center) and the server system (archived files in the main center). If the files have been ingested and moved into the server system, the data ingestion will be stopped; otherwise, it continues.

In addition, there is also the DaemonManager, in which the DaemonLauncher will register each daemon it creates. The DaemonManager ensures that no two Daemons are ever running at the same time. If a daemon is running when another requests permission to run, permission will be denied and the daemon will be added to the wait queue until the current running daemon and all other daemons ahead of it in the queue complete their tasks [30].

IV. Experiment and Analysis

In order to verify the availability of our proposed solution, a virtual multidata center environment was set up based on the OpenStack cloud computing framework. The main data center was composed of three Linux virtual machines. All of the three machines were developed with the SolrCloud environment, responsible for metadata index and retrieval. One of them was developed with OODT system framework, responsible for data ingestion and thumbnail archiving. The distributed sub-center was composed of eight Linux virtual machines, corresponding to eight satellite data centers. Each machine was mounted with a one-terabyte (TB) cloud drive so as to provide image storage space. In addition, all the machines in the main and sub centers were configured with 4 gigabytes (GBs) of RAM and 2 virtual processor cores. The framework of the virtual multi-data center environment is shown in Figure 3.

The experimental images of the distributed integration test mainly include Landsat 8 OLI TIRS, Landsat 7 ETM+, Landsat 5 TM, Landsat 15 MSS, Aster L1T, CEBERS-1/2 CCD, HJ-1A/B CCD, HJ-1A HSI, and FY- 3A/B VIRR images, which were freely downloaded from the USGS (https://earthexplorer.usgs.gov/) , NSMC (http://satellite.nsmc.org.cn/portalsite/default.aspx) and CCRSDA (http://www.cresda.com/CN) websites. A total of 3380 files were downloaded. These images were distributed in the eight sub-centers according to data type. The total number of our experimental images are shown in Table II.

The distributed data integration experiment mainly includes remote sensing data polling, metadata extraction, thumbnail generation, file transferring, thumbnail archiving, metadata indexing, and other processes. The experimental results are primarily with respect to the already-crawled data volume and total time consumption from the RemoteCrawler launch to metadata being indexed by SolrCloud/Lucene. Because no two push-pull daemons ever run concurrently, the distributed data integration experiment was carried out one sub-center at a time. The experiment procedures and results are shown in Table III.

As can be seen in Table III, the number of main center-integrated remote sensing images is equal to the total number of each sub-center's stored images. That is to say, there is no information lost during the process of data integration. Moreover, our designed ISO 19115-2:2009-based uniform metadata model includes all fields of integration by participating remote sensing metadata, and the SolrCloud indexed metadata can also maintain the metadata information of each remote sensing image perfectly. As for the transfer rate, it mainly depends on the window size for the OODT-push-pull component. In our experiment, the window size was set at 1024 bytes, and the average transfer rate is between 9.8 and 13.8 MB/s. This is enough to satisfy the demands of metadata and thumbnail transfer across a distributed data center spatial infrastructure. Therefore, the experimental results showed that our OODT-based distributed remote sensing data integration was feasible.

V. Conclusions

In view of the current issues of remote sensing data integration, we proposed an OODT-based data integration framework. Specifically, aiming at heterogeneous features of multi-source remote sensing data, we proposed an ISO 19115-2:2009-based metadata transform method to achieve unity of metadata format in the distributed sub-centers. In order to achieve efficient, stable, secure and usable remote sensing data integration across a distributed data center spatial infrastructure, we

www.czasopisma.pan.pl PAN www.journals.pan.pl
POLSKA AKADEMIA NAUK

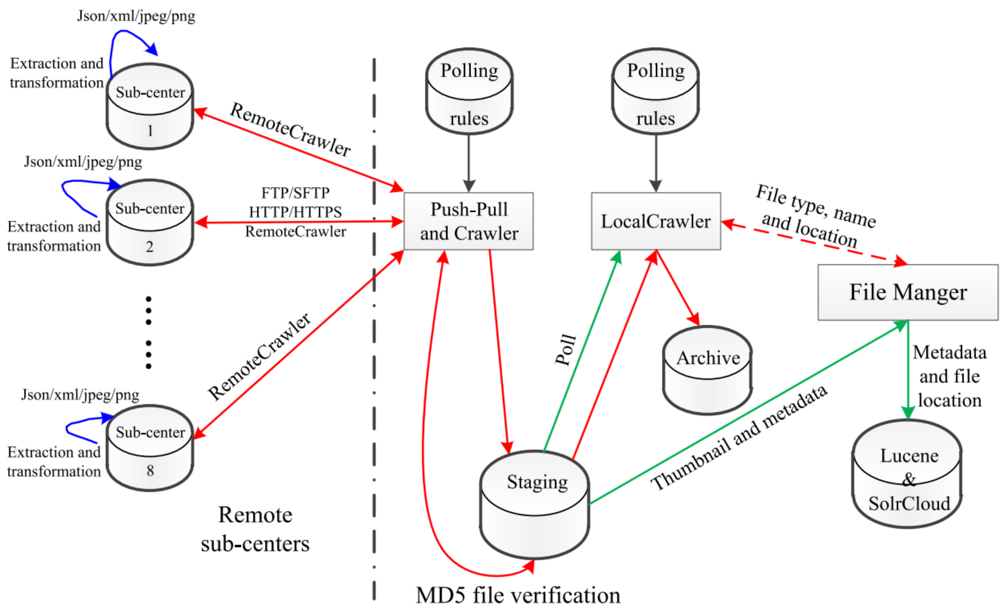

Fig. 2. The process of distributed data integration

TABLE II
A SUMMARY OF THE EXPERIMENTAL IMAGES

Sub-Center	Satellite	Data Type	Volume of Images	Image Format
1	Landsat 8	OLI_TIRS	310	GeoTIFF
2	HJ-1A	HSI	350	HDF5
2	CEBERS-1/2	CCD	270	GeoTIFF
3	Landsat 7	ETM+	450	GeoTIFF
4	Landsat1-5	MSS	260	GeoTIFF
5	HJ-1A/B	CCD	710	GeoTIFF
6	Landsat 5	TM	430	GeoTIFF
7	FY-3A/B	VIRR	450	HDF5
8	Aster	L1T	150	HDF4

TABLE III
EXPERIMENTAL RESULTS OF DISTRIBUTED DATA INTEGRATION

Satellite	Data Type	Volume of Images Stored in Sub-Center	Volume of Images Integrated by Main Center	Average Transfer Rate(MB/s)
Landsat 8	OLI_TIRS	310	310	9.8
HJ-1A	HSI	350	350	10.1
CEBERS-1/2	CCD	270	270	11.7
Landsat 7	ETM+	450	450	10.5
Landsat1-5	MSS	260	260	12.8
HJ-1A/B	CCD	710	710	9.9
Landsat 5	TM	430	430	13.8
FY-3A/B	VIRR	450	450	11.2
Aster	L1T	150	150	10.8

adopted the OODT framework based on its stable, efficient, and easy-to-expand features, to implement remote sensing data polling, thumbnail generation, file transfer, thumbnail archiving, metadata storage, etc. In addition, in order to verify the availability of our proposed program, a series of distributed data integration experiments was carried out. The results showed that our proposed distributed data integration program was effective and provided superior capabilities.

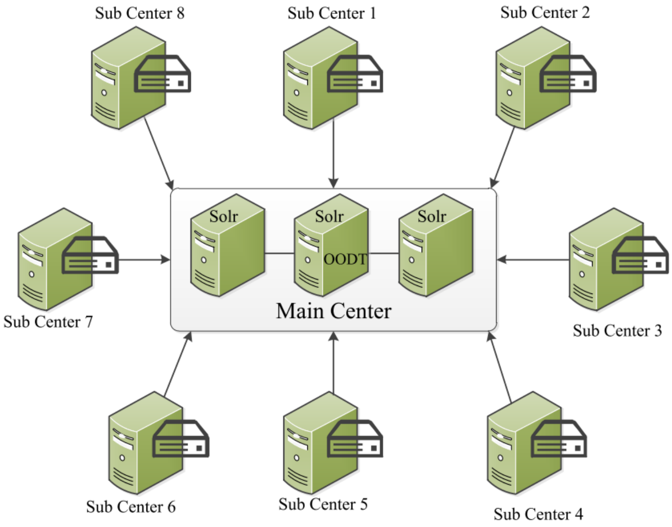

Fig. 3. The framework of the virtual multi-center data environment

However, the unified metadata conversion rule was pre-configured, and the metadata transformation was done manually. This was convenient and easy to operate, but less efficient. In particular, with an increase of data types, a great burden would be brought to data integration. Future studies based on deep learning algorithms using semantic matching and unified format conversion of remote sensing metadata will be performed.

REFERENCES

[1] Xiang Li and Lingling Wang. On the study of fusion techniques for bad geological remote sensing image. J. Ambient Intelligence and Humanized Computing, 6(1):141-149, 2015. https://doi.org/10.1007/s12652-015-0255-1.

[2] Robert Jeansoulin. Review of forty years of technological changes in geomatics toward the big data paradigm. ISPRS International Journal of Geo-Information, 5(9):155, 2016. https://doi.org/10.3390/ijgi5090155.

[3] A. Lowe, D.; Mitchell. Status report on NASA's earth observing data and information system (eosdis). In The 42nd Meeting of the Working Group on Information Systems Services, pages Frascati, Italy, 19-22 September, 2016. https://doi.org/10.5334/dsj-2019-040.

[4] China's fy satellite data center. Available online: http: // satellite. cma. gov. cn/ portalsite/ default. aspx , (accessed on 25 August 2017).

[5] China center for resources satellite data and application. Available online: http: // www. cresda. com/ CN/ sjfw/ zxsj/ index. shtml , (accessed on 25 August 2017).

[6] Jining Yan and Lizhe Wang. Suitability evaluation for products generation from multisource remote sensing data. Remote Sensing, 8(12):995, 2016. https://doi.org/10.3390/rs8120995.

[7] Minggang Dou, Jingying Chen, Dan Chen, Xiaodao Chen, Ze Deng, Xuguang Zhang, Kai Xu, and Jian Wang. Modeling and simulation for natural disaster contingency planning driven by high-resolution remote sensing images. Future Generation Computer Systems, 37(C):367-377, 2014. https://doi.org/10.1016/j.future.2013.12.018.

[8] Dongyao Wu, Liming Zhu, Xiwei Xu, Sherif Sakr, Daniel Sun, and Qinghua Lu. Building pipelines for heterogeneous execution environments for big data processing. IEEE Software, 33(2):60-67, 2016. https://doi.org/10.1109/MS.2016.35.

[9] Chaowei Yang, Qunying Huang, Zhenlong Li, Kai Liu, and Fei Hu. Big data and cloud computing: innovation opportunities and challenges. International Journal of Digital Earth, 10(1):13-53, 2017. https://doi.org/10.1080/17538947.2016.1239771.

[10] Xin Luo, Maocai Wang, Guangming Dai, and Xiaoyu Chen. A novel technique to compute the revisit time of satellites and its application in remote sensing satellite optimization design. International Journal of Aerospace Engineering,2017,(2017-01-31), 2017(6):1-9, 2017. https://doi.org/10.1155/2017/6469439.

[11] Andrew Mitchell, Hampapuram Ramapriyan, and Dawn Lowe. Evolution of web services in eosdis — search and order metadata registry (echo). In Geoscience and Remote Sensing Symposium,2009 IEEE International,igarss, pages 371-374, 2010. https://doi.org/10.1109/IGARSS.2009.5417653.

[12] Available online: http://oodt.apache.org/. Oodt. (accessed on 25 January 2017).

[13] Chris A. Mattmann, Daniel J. Crichton, Nenad Medvidovic, and Steve Hughes. A software architecture-based framework for highly distributed and data intensive scientific applications. In International Conference on Software Engineering, pages 721-730, 2006. https://doi.org/10.1145/1134285.1134400.

[14] Chris A. Mattmann, Dana Freeborn, Dan Crichton, Brian Foster, Andrew Hart, David Woollard, Sean Hardman, Paul Ramirez, Sean Kelly, and Albert Y. Chang. A reusable process control system framework for the orbiting carbon observatory and npp sounder peate missions. In IEEE International Conference on Space Mission Challenges for Information Technology, 2009. Smc-It, pages 165-172, 2009. https://doi.org/10.1109/SMC-IT.2009.27.

[15] Dennis C. Reuter, Cathleen M. Richardson, Fernando A. Pellerano, James R. Irons, Richard G. Allen, Martha Anderson, Murzy D. Jhabvala, Allen W. Lunsford, Matthew Montanaro, and Ramsey L. Smith. The thermal infrared sensor (tirs) on Landsat 8: Design overview and prelaunch characterization. Remote Sensing, 7(1):1135-1153, 2015. https://doi.org/10.3390/rs70101135.

[16] Khandelwal, Sumit and Goyal, Rohit. Effect of vegetation and urbanization over land surface temperature: case study of Jaipur City. EARSeL Symposium, pages 177-183, 2010.

[17] Yaxing Wei, Liping Di, Baohua Zhao, Guangxuan Liao, and Aijun Chen. Transformation of hdf-eos metadata from the ecs model to iso 19115- based xml. Computers & Geosciences, 33(2):238-247, 2007. https://doi.org/10.1016/j.cageo.2006.06.006.

[18] Mahaxay, Manithaphone and Arunpraparut, Wanchai and Trisurat, Yongyut and Tangtham, Nipon. Modis: An alternative for updating land use and land cover in large river basin. Thai J. For, 33(3):34-47, 2014.

[19] Bo Zhong, Yuhuan Zhang, Tengteng Du, Aixia Yang, Wenbo Lv, and Qinhuo Liu. Cross-calibration of hj-1/ccd over a desert site using Landsat etm + imagery and Aster gdem product. IEEE Transactions on Geoscience & Remote Sensing, 52(11):7247-7263, 2014. https://doi.org/10.1109/TGRS.2014.2310233.

[20] Ranjeet Devarakonda, Giriprakash Palanisamy, Bruce E. Wilson, and James M. Green. Mercury: reusable metadata management, data discovery and access system. Earth Science Informatics, 3(1-2):87-94, 2010. https://doi.org/10.1007/s12145-010-0050-7.

[21] Nengcheng Chen and Chuli Hu. A sharable and interoperable metamodel for atmospheric satellite sensors and observations. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 5(5):1519-1530, 2012. https://doi.org/10.1109/JSTARS.2012.2198616.

[22] Peng Yue, Jianya Gong, and Liping Di. Augmenting geospatial data provenance through metadata tracking in geospatial service chaining. Computers Geosciences, 36(3):270-281, 2010. https://doi.org/10.1016/j.cageo.2009.09.002.

[23] Gilman, Jason Arthur and Shum, Dana. Making metadata better with cmr and mmt. NASA, 2016. https://ntrs.nasa.gov/search.jsp?R= 20160009277.

[24] Ann B. Burgess and Chris A. Mattmann. Automatically classifying and interpreting polar datasets with Apache tika. In IEEE International Conference on Information Reuse and Integration, pages 863-867, 2014. https://doi.org/10.1109/IRI.2014.7051982.