

Automatisierte Einbindung und Verarbeitung von Copernicus-Daten in ein Digitales Lager geotopografischer Daten

MARKUS MUERTH¹, MARCO ZULKOWSKI¹, NINA KRÜGER¹,
INES KOPPENHAGEN¹ & DANIEL HOLWEG¹

Zusammenfassung: Das Ziel von Geodateninfrastrukturen (GDI) in Verwaltung und Wirtschaft ist meist die unterbrechungsfreie Bereitstellung und Verarbeitung von Geodaten unterschiedlichster Formate und Quellen. Satellitendaten des EU Copernicus Programm stellen u.a. durch ihre hohe zeitliche Auflösung eine besondere Ergänzung des Datenarchivs dar. Ein Digitales Lager geotopographischer Daten ermöglicht die vollautomatische und regelbasierte Einbindung, Konfektionierung, Datenhaltung, Bereitstellung und Verarbeitung vieler Arten von Geodaten. Neben einer konkreten Umsetzung des Konzepts als Software, zeigen wir zwei typische Anwendungsfälle für die Einbindung von Sentinel-2 Daten auf, die mit Hilfe des Digitalen Lagers wiederkehrend automatisch ausgeführt werden können.

1 Einleitung

Im Rahmen des europäischen Erdbeobachtungsprogramms Copernicus wurden eine Reihe von Möglichkeiten geschaffen, die Nutzern einen leichten und freien Zugang zu Sentinel Satellitendaten und weiteren Copernicus-Diensten über Datenportale (z.B. ESA Science Hub) und Cloud-Plattformen (z.B. CODE-DE, AWS, Creodias) ermöglichen. Besonders seit dem Start von Sentinel-2A im Jahr 2015 entwickelte sich ein großes Interesse an Copernicus in verschiedenen potentiellen Nutzerkreisen. Andererseits entwickelten sich bereits seit den 1990er Jahren, durch den Einzug von Geoinformationssystemen in Wirtschaft und Verwaltung, server-basierte Geodateninfrastrukturen (GDI) zur Datenhaltung und Verarbeitung geotopografischer Daten. Bei diesen werden neue Daten entweder durch die Nutzer oder über vordefinierte Schnittstellen eingebunden. Moderne GDI erlauben darüber hinaus auch die Darbietung der Produkte an Abnehmer, z.B. über Portale und Web Dienste wie WMS oder WFS.

Aufgrund der oft großen Datenmengen und zum Teil komplexen Datenstrukturen von gerasterten Geodaten, gilt es, jetzt bestehende lokale Dateninfrastrukturen (z.B. für geotopographische Daten) mit den neueren, cloudbasierten Infrastrukturen (z.B. für Satellitendaten) zu verbinden. Dies kann durch eine regelbasierte, voll automatisierte Einbindung von Satellitendaten aus der Cloud in eine GDI geschehen. Beispielsweise da große Mengen an Daten schon in der GDI vorgehalten werden bzw. dort entsprechende Rechenkapazitäten und Algorithmen bereitstehen.

Nach einer kurzen Übersicht über die Problemstellung und den Lösungsansatz werden die Vorteile der Einbindung von Satellitendaten in ein Digitales Lager für Geodaten herausgestellt. Schließlich beleuchtet der Beitrag die effiziente Einbindung von Satellitendaten in ein Digitales Lager anhand eines Beispiels. In diesem Beispiel erfolgt die automatisierte regelbasierte Auswahl, Aufbereitung und Bereitstellung von Sentinel-2 Daten für das Digitale Lager durch den Dienst *Sentinel Archiver*.

¹ M.O.S.S. Computer Grafik Systeme GmbH, Hohenbrunner Weg 13, D-82024 Taufkirchen,
E-Mail: [mmuerth, dholweg]@moss.de

Das Modul *Sentinel Connector* als Teil der GDI-Komponente *novaFACTORY* registriert die bereitgestellten Daten in der GDI und koordiniert voreingestellte Verarbeitungsschritte. Schließlich werden die Satellitendaten im Digitalen Lager auf Basis der Software *novaFACTORY* konfektioni-ert und gelagert, sowie Anwendern und Veredelungs-Prozessen bereitgestellt. So können sich die Nutzer ganz auf die produktive Auswertung der Satellitendaten auch im Zusammenspiel mit anderen Fernerkundungsdaten wie z.B. Luftbildern und amtlichen, vektor-basierten Geodaten (wie z.B. ALKIS und ATKIS) fokussieren.

2 Regelbasierte Bereitstellung von Satellitendaten für geotopogra-phische Produktions- und Vertriebsprozesse

2.1 Problemstellung

In verschiedenen öffentlichen und wirtschaftlichen Bereichen werden bereits seit Jahrzehnten Geodatenbestände aufgebaut und verarbeitet, die insbesondere seit dem Start des Copernicus-Pro-gramms nun auch die Möglichkeit bieten, unterschiedlichste frei verfügbare Satellitendaten und darauf aufbauende Produkte für ihre Anwendungen zu nutzen. Dazu gehören insbesondere opti-sche Sensordaten und abgeleitete Produkte aus der Sentinel-2 Mission, die z.B. die Landnutzungs-kartierung (z.B. MALINOWSKI et al. 2020) oder das Monitoring von Wäldern (FASSNACHT et al. 2021) durch eine hohe Dichte der Zeiterie verbessern. Satellitendaten ermöglichen aber auch ver-schiedene Anwendungen in der Landschaftsplanung, z.B. das Ausloten von Potentialen für die Nutzung Erneuerbarer Energien im Projekt COP4EE (JONAS et al. 2019).

Da moderne Satellitensensoren eine große Menge an Daten bereitstellen, werden diese in zuneh-mendem Maße über Datenportale und Cloud-Plattformen auch zur Verarbeitung in der Cloud be-reitgestellt. Jedoch stellt die wiederkehrende Verwendung von Satellitendaten in operativen Pro-zessen Anforderungen, die z.T. über den Möglichkeiten dieser Plattformen liegen (siehe u.a. KLIEN et al. 2019).

- Die **Massendatenverarbeitung** von Satellitendaten mit z.T. noch größeren Mengen an lokal vorliegenden Daten (z.B. Luftbildern)
- Der **Datenschutz** von sensitiven Eingangsdaten muss gewährleistet werden.
- Eine „**Datenfusion**“ von Sensoren mit ähnlichen Eigenschaften (z.B. Sentinel-2 und Land-sat) oder von Daten aus unterschiedlichen Aufnahmetechniken (z.B. Sentinel-1 SAR, Sen-tinel-2 MSI oder Laserscanner Punktwolken) erfordert immer wiederkehrende Vorverar-beitung der Satellitendaten.
- Satellitendaten unterschiedlicher Anbieter werden über jeweils **unterschiedliche Reposi-torien** bzw. Plattformen bereitgestellt.
- **Kommerzielle Anwendungen** oder vom Nutzer selbst erstellte Algorithmen lassen sich oft nur aufwändig oder gar nicht in andere Infrastrukturen übertragen.
- Die automatische Ausführung von **datengetriebenen Workflows** für komplexe Arbeits-abläufe benötigt ein spezialisiertes Workflow Management System.

Jedoch erfüllen manche GDI Software Lösungen diese Anforderungen, z.B. solche die auf dem Konzept des Digitalen Lagers beruhen.

2.2 Lösungsansatz

Technische Basis der vorgestellten Lösung für eine operationelle Einbindung von Satellitendaten in eine GDI ist eine server-basierte Datenverwaltungs- und Prozessierungsumgebung, wie sie bereits vielfach in der Vermessungs- und Umweltverwaltung, sowie in der Planung, Verwaltung und Überwachung von raumwirksamen wirtschaftlichen Infrastrukturen eingesetzt wird. Zusammen mit Frontend-Anwendungen wie z.B. GIS Oberflächen für Desktop und Browser kann ein sogenanntes Digitales Lager für eine Vielzahl von geotopografischen Daten (3D, Punktwolken, Raster, Vektor), und auch räumlich verorteter Dokumente (Geodokumente) entstehen. Beispiele dafür sind Photographien, Verträge oder amtliche Dokumente wie grafische Nachweise und Zahlennachweise der Vermessung.

Die Integration von Erdbeobachtungsdaten in solch eine GDI sollte durch die Anbindung eines Dienstes erfolgen, der in der Cloud-Umgebung, in der Satellitendaten zur Verfügung gestellt werden, installiert ist. Die Raum-Zeit Komponente und die Art der importierten Daten können so von der lokalen GDI über den Dienst in der Cloud gesteuert werden. Es entsteht bei den Nutzern der lokalen GDI die volle Kontrolle über die verwendeten Daten. Auch ermöglicht es ihnen einen regelbasierten, automatischen Import und die Weiterverarbeitung der Daten mit bekannten Werkzeugen.

2.3 Methodik

2.3.1 Das Digitale Lager

Mit dem Konzept des Digitalen Lager für Geodaten werden Grundgedanken der Lagerhaltung aus der industriellen Produktion auf Geodaten übertragen. Damit soll eine zuverlässige, effiziente und ressourcenschonende Umgebung für die Produktion und Bereitstellung von Geodatenbeständen geschaffen werden. Hierbei liegt der Schwerpunkt auf der bedarfsgerechten Bereitstellung von Daten und der darauf aufbauenden Prozesse, unabhängig davon, ob es sich dabei um interne Verarbeitungs- oder externe Bereitstellungsprozesse finaler Datenprodukte handelt. (siehe auch M.O.S.S., 2019)

Zu den wichtigsten Aufgaben des Digitalen Lagers gehört dabei die **Konfektionierung** der Geodaten über klar definierte Datenformate, Kachelungen, Metadaten und Lagerorte. Wie in einem physischen Lager werden alle Informationen über die Daten in einer zentralen Datenbank gespeichert. Zusätzliche können Datenflüsse und Verarbeitungsschritte so festgelegt werden, dass sie wiederholbar und automatisiert ablaufen können.

Dies ermöglicht eine automatisierte **Veredelung** der Daten, die über die reine, einstufige Satellitendatenverarbeitung (wie z.B. Indexberechnungen) hinausgeht. Dazu gehören:

- Verschneidung z.B. mit stark volumenbildenden Daten wie Luftbilder und Vektordaten, die bei den Nutzern von GDIs meist bereits lokal vorliegen,
- Anbindung von frei definierbaren oder vorhandenen Methoden durch Workflow Management (z.B. kaskadierende und mehrstufige Bearbeitungsverfahren),
- Kundenprozesse die auf Anwender-eigenen oder kommerziellen Algorithmen (z.B. Change Detection mit Künstlicher Intelligenz) beruhen, und die Ergebnisse auf Kundeneigenen Endpunkten veröffentlichen.

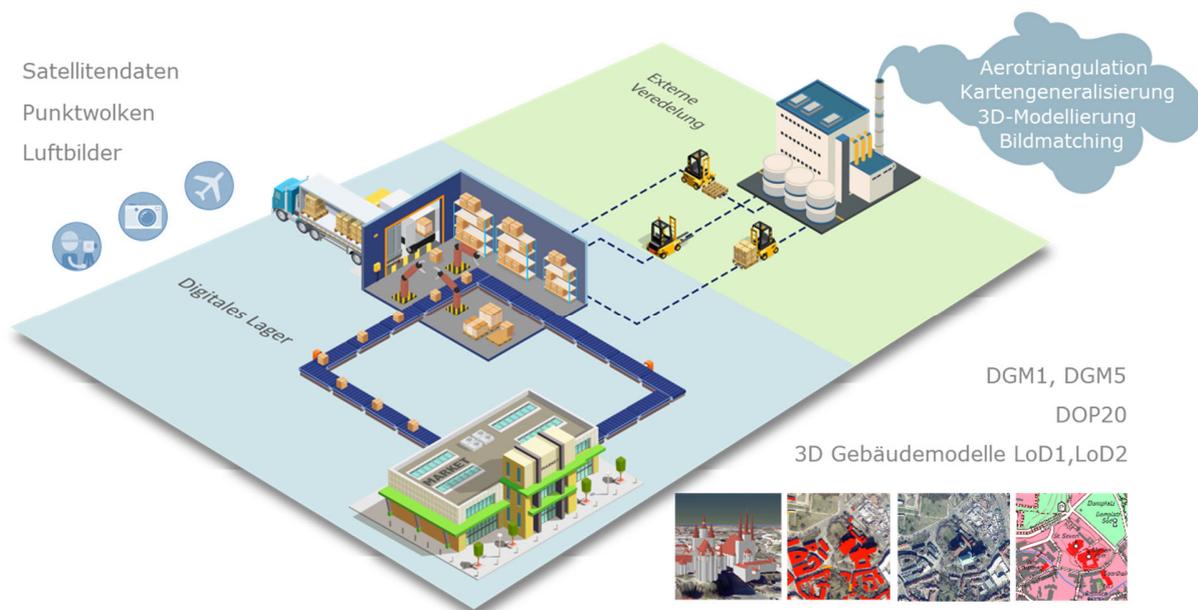


Abb. 1: Konzeptionelle Darstellung eines Digitalen Lagers

Die Vorteile einer automatisierten Veredelung können nur unter bestimmten Voraussetzungen für die Nutzer in Wert gesetzt werden. Dazu gehören:

- Automatische, **regelbasierte Einbindung** aller volumenbildenden bzw. zeitlich variablen Daten in das Digitale Lager
- Bereitstellung von Produkten oder Zwischenprodukten über **regelbasierte Dienste**, insbesondere bei zeitlich variablen Eingangsdaten
- Die Workflows zur Veredelung der Daten über größere Zeiträume oder Gebiete müssen in einer **skalierenden Rechenumgebung** ausführbar sein.

Eine moderne GDI, die das Konzept des Digitalen Lagers aufgreift, stellt genau diese Funktionalitäten bereit. So können z.B. aktuelle Themen in der Fernerkundung wie die automatische Fusion von Daten unterschiedlicher Satellitensensoren oder jede andere Art von Erdbeobachtungsdaten über eine GDI abgedeckt werden. Schließlich sollte ein Digitales Lager auch erlauben, diese Datenprodukte als „Pakete“ (Dateien), webbasierte Dienste (z.B. WMS etc.) und über Portale nach außen bereitzustellen.

2.3.2 Regelbasierter Zugriff auf Satellitendaten

Um Satellitendaten (beispielsweise von Copernicus) in einem Digitalen Lager bereitzustellen, muss die integrierte **Bereitstellungsfunktion** mehrere Bedingungen erfüllen:

- Die Satellitendaten müssen über voreingestellte oder dynamische Raum- und Zeitparameter automatisch ausgewählt werden.
- Idealerweise erlaubt die Schnittstelle zwischen Digitalem Lager und Datenquelle die Auswahl bestimmter Kanäle und vorverarbeiteter Satellitendatenprodukte eines Sensors.
- Die Daten müssen in einheitliche Datenformate übertragen werden und ihre Metadaten auf eine klar strukturierte Weise im Digitalen Lager zur Verfügung gestellt werden.

- Wenn nötig, muss eine Vorverarbeitung der Daten (z.B. Transformation oder zeitliche Aggregation) automatisiert ohne Interaktion mit dem Nutzer durchgeführt werden.

Werden diese Voraussetzungen erfüllt, kann eine Vielzahl von Satellitendaten in automatisierte Workflows eingebunden werden. Dazu können auch Verarbeitungsabläufe gehören, die bei einem bestimmten Auslöser auf das eingebundene Digitale Lager zugreifen, um in kürzester Zeit beispielsweise Kartierungen von Hochwasser- oder Sturmschäden zu erzeugen.

Die von der GDI gesteuerte Auswahl und Aufbereitung einzelner Satellitendaten und ihrer Metadaten sollte jedoch direkt in der Umgebung ausgeführt werden in der diese archiviert sind. Dazu bietet sich an, einen Dienst in der Cloud zu installieren, da dieser dann performant auf das gesamte Satellitendaten-Archiv (z.B. in einer Cloud Plattform) zugreifen kann.

2.4 Beispielhafte Umsetzung des Digitalen Lagers

Eine Möglichkeit der Umsetzung des Konzepts Digitales Lager für geotopographische Daten bietet die modulare und erweiterbare GDI-Komponente *novaFACTORY* (M.O.S.S. 2019). *novaFACTORY* erlaubt die Einbindung, Verwaltung und Bereitstellung einer Vielzahl von Geodaten unterschiedlichster Formate und Quellen und wird hierfür bereits bei einer großen Anzahl an öffentlichen Institutionen und Betrieben genutzt. Zudem ist eine weitere Hauptanwendung von *novaFACTORY* die wiederholbare, vollautomatische Steuerung von Workflows für den Import, den Export und die Verarbeitung von Geodaten.

Die Satellitendaten für die regelbasierte Einbindung in *novaFACTORY* stellt hierbei der cloudbasierte Dienst *Sentinel Archiver* bereit, der die automatisierte regelbasierte Auswahl, Aufbereitung und Bereitstellung von Sentinel-2 Daten für das Digitale Lager durchführt. Er stellt die nötigen Metadaten bereit und bereinigt nicht mehr benötigte oder aktualisierte Daten automatisch. Anwender profitieren von der hohen Frequenz der Datenaktualisierung ohne regelmäßig Kapazitäten dafür zu binden.

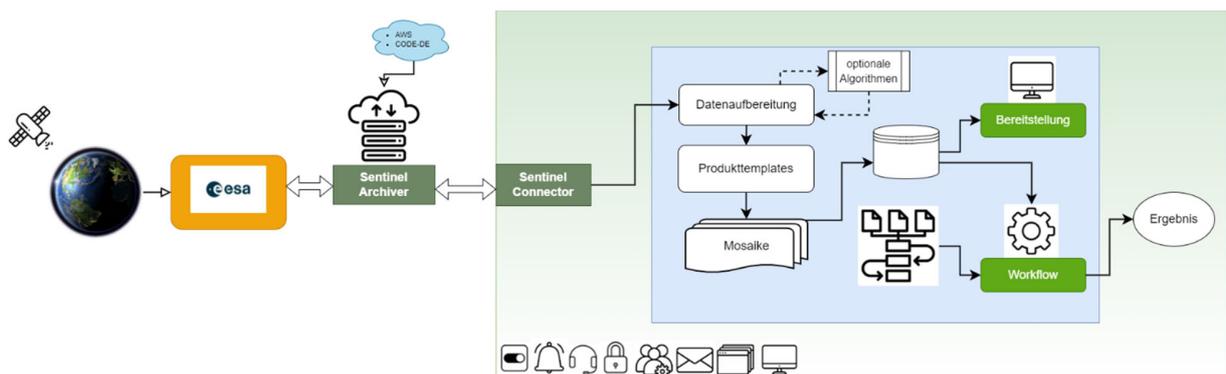


Abb. 2: Workflow zur Einbindung von Sentinel-2 Daten in das Digitale Lager

Das Modul *Sentinel Connector* als Teil der GDI-Komponente *novaFACTORY* registriert schließlich die bereitgestellten Sentinel-2 Daten in der GDI, koordiniert über sogenannte Produktemplates voreingestellte Verarbeitungsschritte (z.B. die Generierung von wolkenfreien Mosaiken und Index Karten) und benachrichtigt die Nutzer bei Bereitstellung dieser Produkte.

Sind die vorverarbeiteten Sentinel-2 Daten schließlich im Digitalen Lager eingebunden, können sie den eigentlichen Nutzern über Download-Portale und Kartendienste (z.B. WMS) zur weiteren Verwendung dargeboten werden. Zudem stehen sie dann regelbasierten Veredelungsprozessen als Input zur Verfügung, so z.B. Change Detection (CD) Algorithmen (Abb. 3). Im Thüringer Landesamt für Bodenmanagement und Geoinformation (TLBG) wird die Aktualisierung des ATKIS Basis-DLM (Digitales Landschaftsmodell) nach ADV, 2019 bereits operationell mit Hilfe der *novaFACTORY* Datenhaltungs- und Workflow-Funktionalitäten durchgeführt. Der vom Anwender definierte Workflow benötigt sowohl Sentinel-2 als auch Digitale Orthophotos (DOPs) und ALKIS/ATKIS Daten die zeitgleich in *novaFACTORY* vorgehalten und automatisch regelbasiert bereitgestellt werden. Solch ein Workflow wird in der GDI skalierbar verarbeitet und das Ergebnis wieder im Digitalen Lager zur weiteren Verwendung verfügbar gemacht.

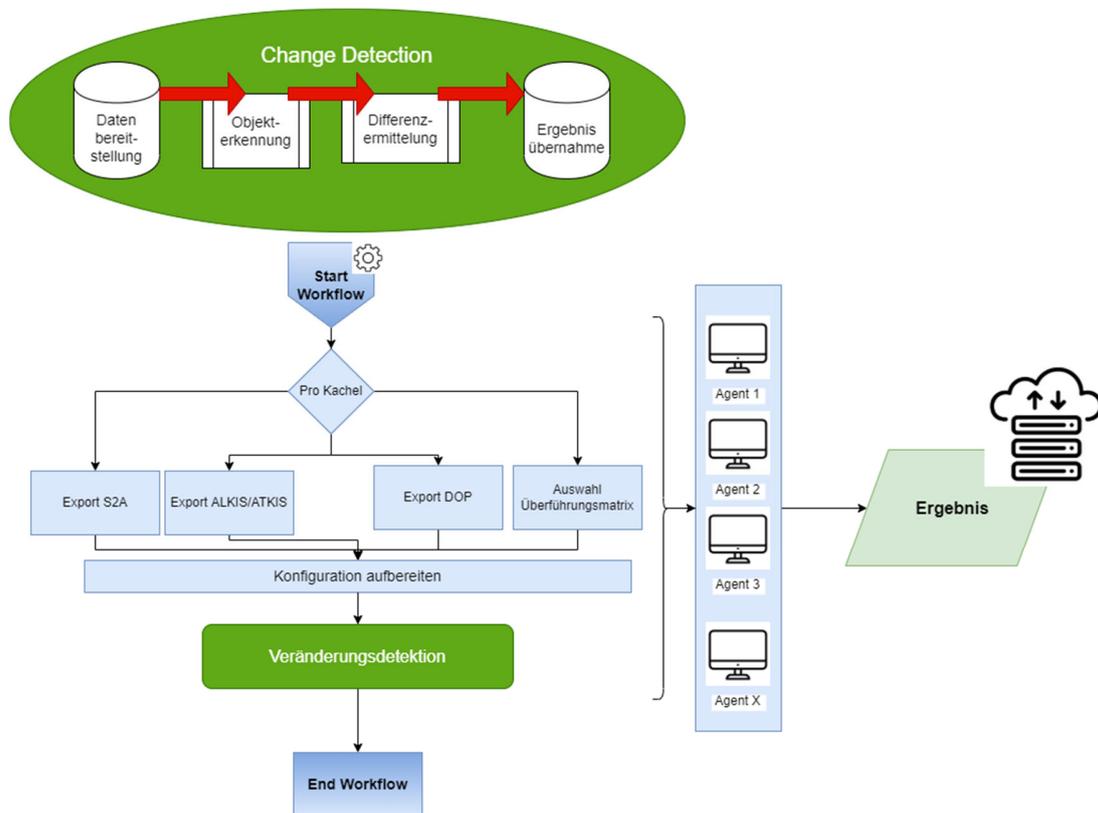


Abb. 3: Beispielhafter Change Detection Workflow

3 Fazit & Ausblick

Die Verarbeitung von Copernicus-Daten in der Cloud stellt für viele nutzerspezifische Fragestellungen immer noch eine Hürde in technischer Hinsicht dar. Beispielsweise benötigt die Erstellung von Waldschadenskarten oder die Change Detection mit Satelliten- und Luftbildern sowie vektorbasierten Karten häufig noch eine große Menge an lokal vorliegenden Daten und/oder basiert auf spezifischen Auswertungsmethoden. Daher ist die Einbindung von Satellitenprodukten (auch aus

unterschiedlichen Quellen) in eine lokale GDI bei operationellen Verarbeitungs- und Bereitstellungsprozessen für komplexe Anwendungen im Vorteil.

Der hier vorgestellte Ansatz zielt deshalb auf die Verknüpfung von cloudbasierten Copernicus Diensten mit einer modernen Lösungsarchitektur für die digitale Lagerhaltung geotopografischer Daten. Ein weiterer Vorteil der automatischen Einbindung der Satellitendaten in eine lokale GDI ist, dass den Nutzern bereits bekannte Anwendungsoberflächen, Weiterverarbeitungsschritte und (auch kommerzielle) Analysetools zur Verfügung stehen. Der Zeit- und damit Kostenaufwand reduziert sich so erheblich gegenüber einer Übertragung anderer Geodaten in die Cloud und der Neugenerierung und Bedienung von komplexen, datengetriebenen Workflows in Cloud Plattformen.

Darüber hinaus ist ein möglicher zusätzlicher Vorteil der Einbindung von Satellitendaten in eine GDI der einheitliche Zugriff auf Daten verschiedener Sensoren. Diese können durch leicht zu automatisierende Vorverarbeitungsschritte in der GDI in einheitlichen Geometrien, Zeitabschnitten sowie einer bestimmten Kanalauswahl bereitgestellt werden. Das erlaubt dann, diese Daten ohne detaillierte Kenntnis ihrer Formate zu höherwertigen Informationsprodukten zu verarbeiten. Diese Form der sogenannten Datenfusion ist ohne entsprechende Infrastruktur nur zeitaufwändig durch Erstellung von manuellen Vorverarbeitungsschritten möglich.

4 Literaturverzeichnis

- ADV (ARBEITSGEMEINSCHAFT DER VERMESSUNGSVERWALTUNGEN DER LÄNDER DER BUNDESREPUBLIK DEUTSCHLAND), 2019: Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok). Gesamtkonzept vom 01.06.2019.
- FASSNACHT, E.F., POBLETE-OLIVARES, J., RIVERO, L., LOPATIN, L., CEBALLOS-COMISSO, A. & GALLEGUILLOS, M., 2021: Using Sentinel-2 and canopy height models to derive a landscape-level biomass map covering multiple vegetation types. *International Journal of Applied Earth Observation and Geoinformation*, **94**, 102236, <https://doi.org/10.1016/j.jag.2020.102236>.
- JONAS, F., WIEDEMANN, W., FIEDLER, C., JETTER, F., SCHMID, T., HOLWEG, D., KLÖBER, A., RICHTER, S., CHRIST, I., LESSING, R. & WALTHER, C., 2019: COP4EE - Entwicklung von Methoden und Diensten auf Basis von Satellitenbilddaten im Kontext von Copernicus zur Unterstützung des Einsatzes von erneuerbaren Energien. Abschlussbericht, DELPHI InformationsMusterManagement GmbH.
- KLIEN, E. M., KRÄMER, M., HOLWEG, D. & WILLKOMM, P., 2019: Cloud4EO - Schnelle und effiziente Verarbeitung von Erdbeobachtungsdaten. Abschlussbericht, Fraunhofer Institut für Geographische Datenverarbeitung.
- MALINOWSKI, R., LEWIŃSKI, S., RYBICKI, M., GROMNY, E., JENEROWICZ, M., KRUPIŃSKI, M., NOWAKOWSKI, A., WOJTKOWSKI, C., KRUPIŃSKI, M., KRÄTZSCHMAR, E., SCHAUER, P., 2020: Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery. *Remote Sensing*, **12**(21), 3523, <https://doi.org/10.3390/rs12213523>.
- M.O.S.S., 2019: Umsetzung eines Digitalen Lagers für Geodaten mit novaFACTORY. M.O.S.S. Computer Grafik Systeme GmbH, <https://t1p.de/nFdigitaleslager20>, letzter Zugriff am 11.01.2022.