

Neue ISO-Norm für Langzeitarchivierung von Fernerkundungsdaten

WOLFGANG KRESSE¹

Zusammenfassung: Vor allem auf Betreiben der NASA ist eine Norm für die Langzeitarchivierung von Fernerkundungsdaten entstanden. Die Norm hat den langen Titel „Preservation of digital data and metadata — Part 2: Content specifications for Earth observation data and derived digital products“. Diese Norm baut auf zwei Basisnormen auf, nämlich der „ISO 19165-1:2018, Geographic information — Preservation of digital data and metadata — Part 1: Fundamentals“ und der „ISO 14721:2012, Space data and information transfer systems — Open archival information system (OAIS) — Reference model“. Derzeit läuft die Abstimmung zum Draft International Standard, die zum Zeitpunkt der Jahrestagung abgeschlossen sein wird. Inhaltlich folgt die Norm weitgehend einem offenbar bei amerikanischen Stellen üblichen Arbeitsfluss: Konzeptuelle Vorbereitung, Modellierung, Implementierung, Durchführung und Nachbereitung.

1 Einleitung

Wohl jeder hat inzwischen schon einmal eine Situation erlebt, in der Daten verloren gegangen sind, weil der Datenträger nicht mehr lesbar war und gleichzeitig eine vorherige Sicherung versäumt worden war. Diese unerwünschte Situation kann nicht nur im privaten Bereich entstehen, sondern in viel größerem Maßstab auch im professionellen Umfeld. Geodaten haben den wahrscheinlich größten Umfang im Vergleich zu Datensätzen anderer Fachgebiete. Daher ist eine weitere Frage zu beantworten, die sonst seltener gestellt wird: Da die Datenmenge schneller als der verfügbare Speicherplatz wächst, muss ausgewählt werden, welche Daten langfristig erhalten werden können und welche anderen Daten wann wieder gelöscht werden dürfen.

Bei Gesprächen mit amtlichen Archivverwaltungen zeigt sich, dass dort die Geodaten mit ihren besonderen Eigenschaften in Umrissen bekannt sind, aber aufgrund der etwas aufwändigeren Archivierungslösung derzeit noch weitgehend ausgeklammert werden. Daher besteht dringender Handlungsbedarf, um für Geodaten den notwendigen langfristigen Schutz zu gewährleisten.

Seitens der ISO, der International Organization for Standardization, wurden in den letzten zehn Jahren eine Reihe von Normen veröffentlicht, die Grundsätze zur Langzeitarchivierung von Geodaten festschreiben. Diese Normen werden weiter unten vorgestellt. Daneben sind eine Reihe von Data Stores für große Geodatenmengen entstanden. Als Beispiel sei hier der „Copernicus Open Access Hub“ für Sentinel-Szenen der ESA genannt, in dem Ende Januar 2020 nach knapp fünf Jahren des Betriebs schon 23.200.000 Datenpakete mit in Summe etwa 20 PBytes frei zum Download zur Verfügung gestellt werden. Die europäischen Sentinel-Satelliten stellen derzeit weniger als zehn von insgesamt etwa 100 Satelliten mit bildgebenden Sensoren im Orbit dar.

¹ Hochschule Neubrandenburg, Brodaer Straße 2, D-17033 Neubrandenburg, E-Mail: kresse@hs-nb.de

2 Besonderheiten von Geodaten

Wie oben schon erwähnt erreichen die aktuellen Fragestellungen der Archivwelt nicht die drängenden Probleme der Geodaten-Community. Für die Archivare steht die Festlegung von stabilen Dokumentformaten im Vordergrund, allen voran das pdf-Format. Geodaten hingegen besitzen eine Reihe von Besonderheiten, die in der Archivwelt sonst eher selten eine Rolle spielen und daher unbekannt sind:

- Erfordernis einer verlustfreien Datenkomprimierung
- Zusammenfassung vieler Einzeldateien zu Kartenwerken
- Georeferenzierung und parallele Verarbeitung von Vektor- und Rasterdaten
- Beschreibung von Nachbarschaften über die Topologie
- Kartographische Ausgestaltung durch Signaturierung
- Ergänzungen durch Metadaten und Quicklooks

3 ISO-Normen

3.1 Überblick

Die ISO hat bisher drei Normen für die Langzeitarchivierung von Geodaten veröffentlicht, zwei davon speziell ausgerichtet auf Fernerkundungsdaten. Diese sind die

- ISO 14721:2012 „Space data and information transfer systems — Open archival information system (OAIS) — Reference model“, die
- ISO 19165-1:2018 „Geographic information — Preservation of digital data and metadata — Part 1: Fundamentals“ und die
- ISO/DIS 19165-2:2019-09 „Geographic information — Preservation of digital data and metadata — Part 2: Content specifications for Earth observation data and derived digital products“ (DIS = Draft International Standard).

3.2 ISO 14721:2012

Die Entwicklung der ISO 14721:2012 wurde hauptsächlich von Organisationen initiiert, die im Bereich der Fernerkundung tätig sind, unter anderem von der NASA und der ESA. Mit dieser Norm wurde eine Grundstruktur für Langzeitarchivierung festgeschrieben, auf der die weiteren Normen aufbauen können. Vor allem definiert die ISO 14721:2012 eine Kapselung der zu archivierenden Daten in einzureichende Datenpakete (SIP, Submission Information Package), zu archivierende Datenpakete (AIP, Archival Information Package) und wieder zu verteilende Datenpakete (DIP, Dissemination Information Package). Die Abb. 1 zeigt einen Überblick über dieses Konzept.

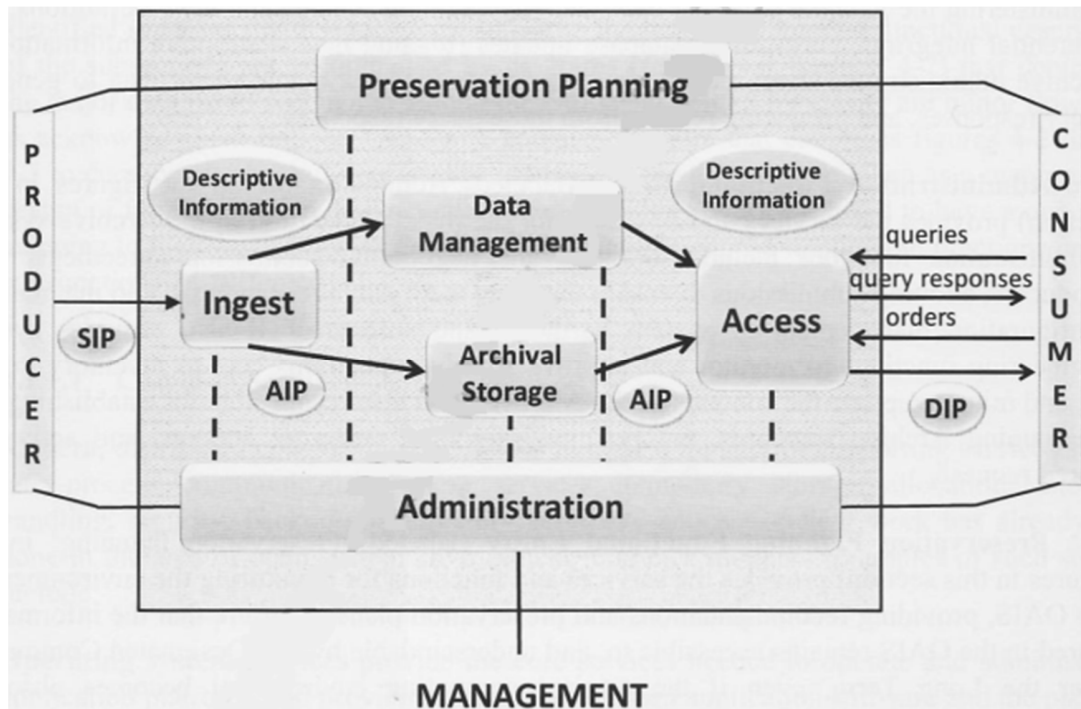


Abb. 1: Struktur und Arbeitsablauf für die Langzeitarchivierung nach ISO 14721:2012 (SIP = Submission Information Package, AIP = Archival Information Package, DIP = Dissimination Information Package)

Bei den Arbeiten zur Norm ISO 19165-1:2018 fiel auf, dass die ISO 14721:2012 ohne Abstimmung mit den Normen des ISO/TC 211 „Geographic information / Geomatics“ entwickelt worden war. Beispielsweise waren und sind viele Definitionen nicht mit denen der Normen der ISO 19100-Serie vereinbar. Bei weiteren Arbeiten fiel auch auf, dass die Autoren der ISO 14721:2012 ihre zweifellos gute, aber eben etwas eigenständige Norm, verbissen verteidigten. Als Folge davon war der Entwicklungsspielraum der ISO 19165-1 stärker eingeschränkt als vorher angenommen.

3.3 ISO 19165-1:2018

Für den Autor begann die Arbeit an der Norm Ende 2013. Damals wurde er am Rande einer ISO/TC 211-Plenarsitzung vom Chairman des ISO/TC 211, damals Olaf Østensen, und dem Normungsverantwortlichen der AdV, Markus Seifert, bedrängt, die Projektleitung zu übernehmen. Glücklicherweise stellte sich bald heraus, dass die Aufgabe von mehreren Seiten erkannt und aktiv angegangen wurde, so dass eine ausreichend breite Expertise für die Entwicklung der Norm zur Verfügung stand. Insbesondere ist dem jungen Nachwuchswissenschaftler Joan Masó aus Spanien (Katalonien) zu verdanken, dass durch seine Vorarbeiten für das Open Geospatial Consortium (OGC) und deren Übernahme in die ISO-Norm nach kurzer Zeit ein stabiles Datenmodell zur Verfügung stand. Ein Beispiel für ein Klassendiagramm zeigt die Abb. 2.

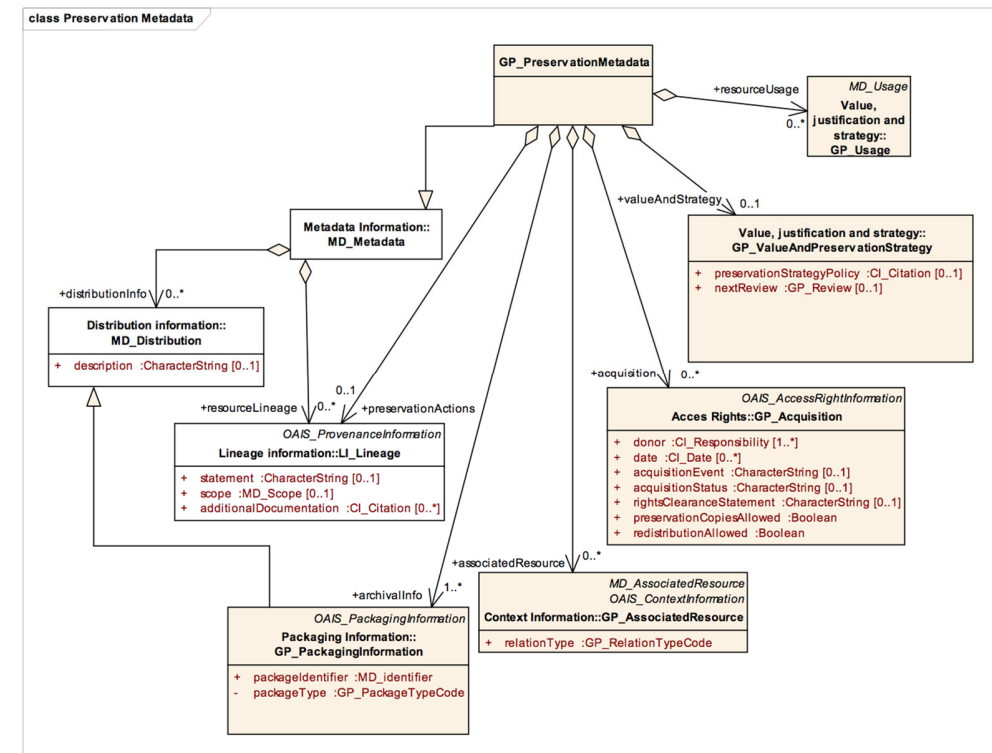


Abb. 2: Klasse GP_PreservationMetadata der ISO 19165-1:2018 mit Unterklassen

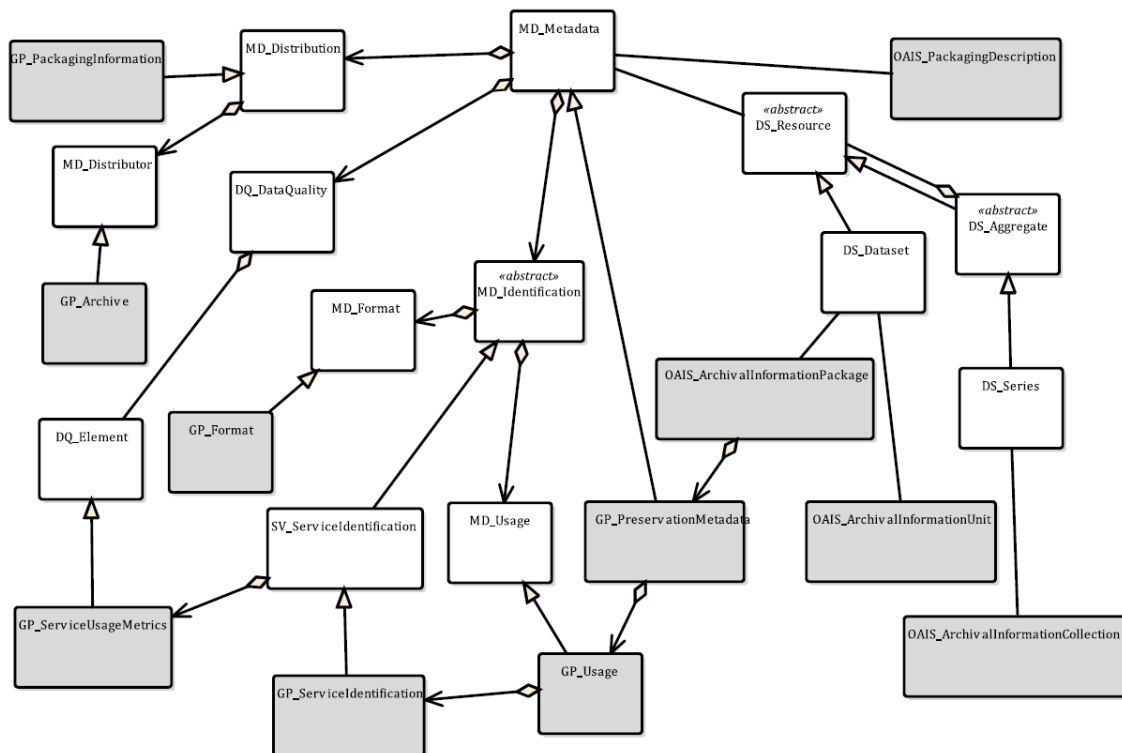


Abb. 3: Klasse MD_Metadata (ISO 19115-1:2014) mit Unterklassen. Grau: Klassen der ISO 19165-2:2018 (GP) und der ISO 14721:2012 (OAIS)

Die Klassen der ISO 19165-1:2018 sind als Spezialisierung der Metadatennorm ISO 19115-1:2014 modelliert ist (Abb. 3).

Sehr empfindlich auf die Neuentwicklung vom ISO/TC 211 reagierten die Väter der ISO 14721:2012. Nach einer etwa ein-Jahr-langen Diskussionen beinhaltete die Kompromisslinie eine Weiterentwicklung der SIP, AIP und DIP zu neuen spezialisierten so genannten GeoSIP, GeoAIP und GeoDIP, und den expliziten Hinweis darauf, dass die neue ISO 19165-1 die Festlegungen der ISO 14721:2012 nicht infrage stelle. Diese Abstimmung war der schwierigste Teil der Entwicklung der neuen ISO 19165-1:2018.

Die ISO 19165-1:2018 äußert sich auch zur Frage der physischen Datenspeicherung. Hierbei sollen Daten und Metadaten getrennt abgelegt werden. Maßgeblich ist die ISO/IEC 29500-2:2012 „Open Packaging Convention“. Diese Norm verlangt die zip-Kompression in Verbindung mit xml. Das Ergebnis ist dem tar-Format ähnlich.

3.4 ISO/DIS 19165-2

Die ISO/DIS 19165-2 ist, wie der Name mitteilt, noch nicht veröffentlicht, aber fast fertig. „DIS“ bedeutet „Draft International Standard“. Die Projektleitung obliegt Hampapuram Ramapriyan, einem US-Amerikaner mit indischer Herkunft, der die Interessen vor allem der NASA im ISO/TC 211 vertritt.

Diese Norm ist eine Implementierungsnorm der ISO 19165-1 für Attribute, die in der Erdbeobachtung benötigt werden. Erdbeobachtung kann von Satelliten, Flugzeugen oder Vor-Ort-Sensoren (in situ) geleistet werden. Die Norm unterteilt die gelieferten Daten in fünf Kategorien: Konzeptuelle Vorbereitung (Mission Concept Stage), Modellierung (Mission Definition Stage), Implementierung (Mission Implementation Stage), Durchführung (Mission Operations Stage) und Nachbereitung (Post Mission Stage). Die Inhalte lassen sich unter anderem mit den Stichworten Planung, Georeferenzierung, Kalibrierung, Datenqualität, Visualisierung und Datenhistorie beschreiben.

Die Tabellen 1 – 5 zeigen alle in dieser Norm definierten Attribute, die in den meisten Fällen eine Protokollierung der Abläufe darstellen. Um die Übersetzung als Fehlerquelle auszuschließen, werden in der Tabelle die originalen englischen Begriffe aus der Norm verwendet.

Tab. 1: Konzeptuelle Vorbereitung (Mission Concept Stage)

Datenelement	Datentyp	Plattform
Preservation metadata	Documentation	All
Scientific / applications scenario, data producer and user communities	Documentation	All
Mission requirements document	Documentation	Satellite mission
Mission operation plan	Documentation	All
Mission cost and schedule	Documentation	All

Tab. 2: Modellierung (Mission Definition Stage)

Datenelement	Datentyp	Plattform
Preservation metadata	Documentation	All
Mission requirements specifications	Documentation	All
Space or aircraft to ground segment ICDs	Documentation	Primarily satellite missions
Sensor / instrument / platform requirements	Documentation	All
Sensor / instrument processing characteristics	Documentation, data record	All
Sensor / instrument qualification process	Documentation, data record	All
Pre-launch / preoperational calibration and characterization plan	Documentation, data record	Satellite missions
Ground / ocean calibration reference and scientific base	Documentation, data record	Satellite missions
Processing algorithms and data format specification	Documentation	Satellite missions, others, if available
Data product specifications	Documentation	All
Data management plan (DMP)	Documentation	All
Mission cost and schedule	Documentation	All

Tab. 3: Implementierung (Mission Implementation Stage)

Datenelement	Datentyp	Plattform
Preservation metadata	Documentation	All
Mission design	Documentation	All
Detailed space or aircraft to ground segment operations concept and implementation	Documentation	Satellite and aircraft missions
Updated DMP	Documentation	All
On board processing	Documentation	Satellite missions
Sensor / instrument design and Implementation	Documentation	Mostly satellite missions
Calibration and validation	Documentation	Satellite and aircraft missions
Ground processor design, algorithm implementation and supporting information for data processing.	Documentation	Satellite missions
Technical notes scientific papers	Documentation	All
Data format specifications	Documentation	All
Supporting information for processing	Documentation, data records	All
Qualification process	Documentation, data records	All

Tab. 4: Durchführung (Mission Operations Stage)

Datenelement	Datentyp	Plattform
Preservation metadata	Documentation	All
Mission data access and service Requirements document and user handbook	Documentation	Satellite missions
Sensor ground segment operations plan	Documentation	Satellite missions
Mission operations acquisition plans and reports	Documentation, data records	Satellite missions
Raw / level 0	Data records	All
Level 1	Data records	All
Level 2+	Data records	All
Browse images	Data records	All
Ancillary data	Data records	Mostly satellite and aircraft missions
Auxiliary data	Data records	Mostly satellite and aircraft missions
Calibration and validation data	Documentation, data records	All
Quality parameters	Documentation, data records	All
Metadata	Documentation, data records	All
Level 0 consolidation	Software code	Satellite missions
Data processing software	Documentation, software code	Satellite and aircraft missions
Quality control software	Software code	Satellite missions
Science data tools	Software code	Satellite missions
Visualization tools	Software code	Satellite missions
Value-added software	Software code	Satellite missions
Product qualification and quality assurance monitoring reports	Software code	Satellite missions
Sensor / instrument evolution and history records	Documentation	All
Referred publications and papers	Documentation	All
Tandem and / or combined campaigns, comparisons	Documentation	All
Cross-campaign, cross-comparisons and cross-calibration activities documentation and data	Documentation, data records	All
Data access policy	Documentation	All

Tab. 5: Nachbereitung (Post Mission Stage)

Datenelement	Datentyp	Plattform
Preservation metadata	Documentation	All
Data consolidation & reprocessing strategy, implementation plans, and consolidated / reprocessed data: Processing	Documentation	Satellite missions
Data consolidation & reprocessing strategy, implementation plans, and consolidated / reprocessed data: Ancillary, Auxiliary	Documentation, data records	Satellite missions
Data consolidation & reprocessing strategy, implementation plans, and consolidated / reprocessed data: QA	Documentation, data records	All
Data consolidation & reprocessing strategy, implementation plans, and consolidated / reprocessed data: L0, L1, L2+	Data records	All
Data consolidation & reprocessing strategy, implementation plans and consolidated / reprocessed data: Metadata	Data	All
Referred publications and papers	Documentation	All
Historical data access policy	Documentation	Satellite missions
Historical mission user handbook	Documentation	Satellite missions

4 Schlussbetrachtung

Die drei Normen zur Langzeitarchivierung zeigen, dass in Archivierungsfragen zu Geodaten die ISO noch nicht zu einer ganz einheitlichen Linie gefunden hat. Während die ISO 19165-1:2018 generisch aufgebaut ist und dem gemeinsamen Datenmodell des ISO/TC 211 folgt (Abb. 3), sind die ISO 14721:2012 und die ISO 19165-2 eher auf die Bedürfnisse von Weltraumorganisationen wie NASA und ESA zugeschnitten. Die ISO 14721:2012 ist dabei gänzlich ohne Bezug zu ISO/TC 211 entstanden. Es bleibt zu hoffen, dass weitere Implementierungen auf Basis der ISO 19165-1:2018 entwickelt werden und dass nach dem Sammeln von Erfahrungen vielleicht eine weiter verbesserte neue Version der ISO 19165-1 entstehen kann.

5 Literaturverzeichnis

- ISO 14721:2012, 2012: „Space data and information transfer systems — Open archival information system (OAIS) — Reference model“. iso.org
- ISO 19165-1:2018, 2018: „Geographic information — Preservation of digital data and metadata — Part 1: Fundamentals“. ISO/TC 211, iso.org
- ISO/DIS 19165-2:2019-09, 2019: „Geographic information — Preservation of digital data and metadata — Part 2: Content specifications for Earth observation data and derived digital products“. ISO/TC 211, iso.org