



Informationen zur Migration der Daten

Metadatenworkshop 2021
Henry Rotzoll (DVZ)

27. September 2021

- Vorgehen bei den Fehlerkorrekturen
- Zahlen der Migration
- Offene Fehler
- Behandlungen von Datensätzen vor der Validierung
- Warnungskorrekturen ausgeführt nach der Validierung
- Abschlussbemerkungen

Vorgehen bei den Fehlerkorrekturen

Metadatenexport aus dem alten GeoMIS.MV

Import der Datensätze in die neue Datenbankstruktur mittels des offiziellen **con terra-Migrationstools**

Export der XML-Daten aus der Datenbank und Anwendung der Fehlerkorrekturen mit Hilfe der **GDI-DE Testsuite** (Version: 1.0.27)

Import der korrigierten XML Metadaten Dateien wieder mit Hilfe des con terra Migrationstools in die endgültige Datenbank

- 1641 Datensätze gesamt migriert (Stand: 13.09.2021)
 - Davon 1183 Datensätze inspireidentifiziert, 458 Datensätze nicht
- 416 Datensätze validierten vor der Korrektur
 - Davon 322 Datensätze inspireidentifiziert, 94 Datensätze nicht
- 910 Datensätze validierten nach der Korrektur
 - Davon 565 Datensätze inspireidentifiziert, 345 Datensätze nicht
- 315 Datensätze hatte noch (bekannte) Fehler
 - Davon 296 Datensätze inspireidentifiziert, 19 Datensätze nicht

- OPERATES_ON_REFERENCE: FEHLER: Ein Dienst-Metadatensatz sollte mindestens eine Verbindung zu einer ebenfalls mit Metadaten beschriebenen Datenquelle besitzen. Laut ([INS VO MD], B.1.6; [ISO 19119]) kennzeichnet das Element operatesOn den bereitgestellten Datensatz durch den eindeutigen Ressourcenidentifikator (URI) des Datensatzes (siehe 3.1). Gemäß [INS TG MD] soll die Referenz dabei jedoch auf ein *MD_DataIdentification*-Objekt eines Daten- Metadatensatzes verweisen
 - 250 INSPIRE, 26 GDI-DE = 276 x Gesamt
- SCHLUESSELWORT_GEMET_FEHLT: Prüfung der Existenz eines Schlüsselwortes aus dem Thesaurus "GEMET - INSPIRE themes" im Element *keyword*
 - 14 INSPIRE, 0 GDI-DE = 14 x Gesamt
- KEYWORD_MISSING: FEHLER: Ein Dienst-Metadatensatz muss immer mindestens ein Schlüsselwort aus dem Teil D (Ziffer 4) der [INS VO MD] stammt ([INS VO MD], Teil B, 3, [INS TG MD], 4.1.2.2, Requirement 3.4) enthalten.
 - 8 INSPIRE, 0 GDI-DE = 8x Gesamt
- EXISTENCE_IDENTIFICATIONINFO_ELEMENT : Prüfung auf Existenz eines *identificationInfo*-Element mit eindeutigem Ressourcenidentifikator.
 - 1x INSPIRE, 1x GDI-DE = 2x Gesamt

- Im Dateisystem direkt auf der XML Datei durchgeführt
- Keine Zahlen, wie oft es durchgeführt wurde, da simples suchen und ersetzen von Zeichenketten

- GmlNamespaceApISOFileTreatment
 - Namespace URL von gml von <http://www.opengis.net/gml> auf <http://www.opengis.net/gml/3.2> geändert und auf den AP ISO:2007 Standard angepasst
- SchemaLocationApISOFileTreatment
 - Namespace URLs von gmd, gmx und srv auf den AP ISO:2007 Standard angepasst
- InternalUrlsHttpsFileTreatment
 - Interne GeoMIS Urls auf von http auf https umgestellt
- SpatialScopeFileTreatment
 - Spatial Scope Keywörter korrigiert
- PreviewPicturesFileTreatment
 - Preview Bilder Urls angepasst

- ReferenceSystemsTreatment
 - Referenzsystemknoten auf die neue URI Form umgestellt
 - bei 1597 Datensätzen durchgeführt, bei 44 Datensätzen nicht durchgeführt
- GeographicIdentifierTreatment
 - Regionalschlüssel von MV ergänzt, wo nötig
 - bei 1514 Datensätzen durchgeführt, bei 127 Datensätzen nicht durchgeführt
- BoundingBoxTreatment
 - Einhaltung von 2 Nachkommastellen bei den Bounding Boxen
 - bei 226 Datensätzen durchgeführt, bei 1415 Datensätzen nicht durchgeführt
- CI_RoleCodeCodeListValueTreatment
 - Überall „Ansprechpartner“ (pointOfContact) bei den Rollen gesetzt, wo nicht schon geschehen
 - bei 608 Datensätzen durchgeführt, bei 1033 Datensätzen nicht durchgeführt

- **ExchangeCodeListUriWarning**
 - Url
http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/codelist/ML_gmxCodelists.xml
zu <http://standards.iso.org/iso/19139/resources/gmxCodelists.xml> geändert. Ein möglich nachfolgender Hash wird bewahrt.
 - Bei 691 Datensätzen durchgeführt, bei 950 Datensätzen nicht durchgeführt
- **ExchangeCodeListUriLanguageCodeWarning**
 - Url http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodelists
zu <http://www.loc.gov/standards/iso639-2/> geändert.
 - Bei 0 Datensätzen durchgeführt, bei 1641 Datensätzen nicht durchgeführt
- **SetServiceTypeLocalNameCodeSpaceWarning**
 - serviceType/LocalName codeSpace Attribut auf “SetServiceTypeLocalNameCodeSpaceWarning” gesetzt
 - Bei 0 Datensätzen durchgeführt, bei 1641 Datensätzen nicht durchgeführt
- **ConformanceResultWarning**
 - Konformitätsergebnis von “unknown” auf “false” gesetzt
 - Bei 421 Datensätzen durchgeführt, bei 1220 Datensätzen nicht durchgeführt

- Alle sinnvoll automatisiert durchzuführenden Änderungen wurden vorgenommen
- Validierung immer mit Hilfe der GDI-DE Testsuite
- Inhaltliche Änderungen müssen die Metadatenpfleger selbst ändern
- Für jeden Metadatenpfleger sind im Zuge der Migration detaillierte Logs erstellt worden, was an seinen Daten geändert wurde