

Mehrsprachigkeit

Beschreibung

Der interoperable Austausch harmonisierter Geodaten erfolgt häufig über sprachliche und kulturelle Grenzen hinweg, aber auch über die Grenzen verschiedener Fachdomänen oder Fachgebiete. Das Interoperabilitätselement *Mehrsprachigkeit* soll Mechanismen identifizieren, um:

- den Austausch, die Interpretation und die Verwendung von Daten über Sprachgrenzen und Fachgebiete hinweg zu verbessern. Beispielsweise gibt es unterschiedliche Objektbezeichnungen bei „*Straßenabläufen*“ (*Entwässerungen von Straßen*) in verschiedenen Fachdomänen, die umgangssprachlich als *Gully* bezeichnet und in Fachdatenbeständen als *Sinkkasten* bzw. als *Anschlusspunkt des Kanalsystems* enthalten sind.
- nationale und internationale terminologische Abstimmungen zu ermöglichen. Die Festlegung von Terminologie (Begriffe und ihre Definition) in der GDI-DE wird in einem separaten Interoperabilitätselement geregelt.
- eine national und europaweit abgestimmte Interpretation, Erweiterung und Nutzung der INSPIRE Datenmodelle zu ermöglichen.

Mehrsprachigkeit betrifft insbesondere folgende Aspekte:

- Sprache von **Datenspezifikationen / Datenschemata**: Die einheitlichen Festlegungen von wesentlichen Begriffen und der Definitionen müssen bei Bedarf mehrsprachig vorgehalten werden. Dies betrifft insbesondere **Codelisten, Enumerationen, Objektartenkataloge** und **Register für Featurekonzepte**.
- Geographische Objekte haben häufig unterschiedliche **Namensgebungen** in verschiedenen Sprachen und mehrere lokale Bezeichnungen (Endonyme und Exonyme¹). In Ortsregistern (sog. Gazetteers) können in der Regel mehrere Bezeichnungen für einen Ort abgelegt werden. Diese Register werden beispielsweise auch bei der Datensuche zur automatischen Übersetzung der Suchbegriffe verwendet.
- **Metadaten**: Insbesondere Schlüsselwörter (Keywords) sollten mit einem entsprechenden multilingualen Thesaurus (kontrolliertes Vokabular, strukturierte Wortschatzsammlung) verknüpft werden. Thesauri ermöglichen die Überbrückung kultureller Unterschiede bzw. die Übersetzung von Fachtermini unterschiedlicher Fachgebiete. Ein typisches Beispiel zur Realisierung der Mehrsprachigkeit von Schlüsselwörtern in Metadaten ist im INSPIRE-Umfeld der GEMET Thesaurus (GEneral Multilingual Environmental Thesaurus). Er umfasst 27 Sprachen der EU und bildet damit einen Ausgangspunkt für die Nutzung von mehrsprachigen Metadaten. Bei Bedarf können beschreibende Texte in den Metadaten (wie z.B. der Titel, die Kurzbeschreibung oder der Zweck) multilingual vorgehalten werden.

¹ https://de.wikipedia.org/wiki/Exonym_und_Endonym

Aktueller Stand in INSPIRE und der GDI-DE

Für INSPIRE wird das Thema „multi-lingual text and cultural adaptability“ im Generic Conceptual Model D2.5 behandelt und entsprechend Vorgaben gemacht:

- Von den Mitgliedsstaaten wird nicht gefordert per se multilinguale Daten bereitzustellen oder sämtliche Daten in eine einheitliche Sprache zu übersetzen.
- Festlegungen für Ortsnamen (geographical names):
 - Die Verwendung von explizit vorgehaltenen Exonymen wird zur Übersetzung vorgeschrieben (z.B. München / Munich).
 - Eine direkte Übersetzung von Ortsnamen, Teilen von Ortsnamen ist nicht vorgesehen (z.B. die Übersetzung von Vorsilben *Klein-* oder *Neu-* bei Ortsnamen).
 - Die Anzahl der Namen in unterschiedlichen Sprachen für ein Objekt soll nicht begrenzt sein.
- Der Gebrauch von Freitexten als Datentyp in Geodaten und insbesondere in technischen Spezifikationen sollte minimiert werden. Nach Möglichkeit sind Klassifizierungen über Codelisten zu verwenden, die in einem INSPIRE Register geführt werden.
- Alle INSPIRE Register werden multilingual geführt (z.B. Feature Concept Dictionary²). Die zentralen Vokabulare stellen die Begrifflichkeiten in allen offiziellen Sprachen der EU bereit.

In der GDI-DE gibt es bisher keine übergreifenden, konkreten Vorgaben bzw. detaillierte Regelungen zum vereinheitlichten Umgang mit Mehrsprachigkeit in den Geodaten, Datenspezifikationen und Datenschemata. Im Dokument zu den Konventionen zu Metadaten (Version 1.2.0, 08-2017)³ wird die Zuordnung von INSPIRE Datenthemen als Schlüsselwort unter Verwendung des GEMET beschrieben. Weitere Aussagen oder Empfehlungen zur Mehrsprachigkeit, insbesondere bei den Metadatenelementen wie Kurzbeschreibung oder Titel werden nicht gemacht.

In den GDI der Länder wird teilweise, z. B. in Sachsen, die Zuweisung mindestens eines zutreffenden Schlüsselwortes aus dem GEMET empfohlen, unabhängig davon ob der Datensatz für INSPIRE relevant ist.

Beispiele für Mehrsprachigkeit in grenznahen oder grenzübergreifenden Daten sind im Crossdata (CILC3)-Projekt⁴ für den Bereich der Raumplanung zu finden.

² <http://inspire.ec.europa.eu/featureconcept>

³ <https://wiki.gdi->

[de.org/download/attachments/3344909/Architektur_GDI_DE_Konventionen_Metadaten_v1_2_0.pdf?version=1&modificationDate=1504012535773&api=v2](https://wiki.gdi-de.org/download/attachments/3344909/Architektur_GDI_DE_Konventionen_Metadaten_v1_2_0.pdf?version=1&modificationDate=1504012535773&api=v2)

⁴ <http://www.cross-data.eu/>

Bewertung und Handlungsbedarf

Das Thema Mehrsprachigkeit in der GDI-DE ist auch unter der Annahme relevant, dass der Anwendungsbereich Deutschland und die Sprache Deutsch ist. Eine standardmäßige Übersetzung von Daten oder Datenschemata in andere Sprachen wird nicht benötigt. Eine Übersetzung sollte immer dann erfolgen, wenn sie durch entsprechende internationale Anwendungsfälle gefordert ist oder wenn die Daten sprach- und fachübergreifend genutzt oder angeboten werden.

Festlegungen:

- Aus technischer Sicht sind ISO 19110 konforme Objektartenkataloge (engl. feature catalogues) und ISO 19126 konforme Verzeichnisse und Register für Featurekonzepte (engl. feature concept dictionaries) multilingual, so dass an dieser Stelle keine weiteren Anforderungen gestellt werden.
- Für erweiterte INSPIRE Datenmodelle wird eine englische Übersetzung der Namen der wesentlichen Objektklassen, Definitionen und Codelisten empfohlen, um eine internationale Weiternutzung zu erleichtern.
- Bereitstellung von Anwendungsschemata sollte in der Sprache erfolgen, in der auch die Daten vorliegen.
- Anwendungsschemata (Bezeichnungen der Datenelemente, Objektartendefinitionen) sollen jeweils einheitlich in einer Sprache vorliegen.

Handlungsbedarf:

Nr.	Maßnahmen	<u>Federführer, Akteure</u>	Dokumente und Referenzen	Status
10.1	Allgemeine Festlegungen oder Empfehlungen zur Mehrsprachigkeit (in Keywords und Textfeldern wie z.B. Titel, Kurzbeschreibung oder Zweck) bzw. der/den zu verwendenden Sprache(n) und der Verwendung von Thesauri in den Metadaten	<u>AK Metadaten</u>	Konventionen zu Metadaten (Version 1.2.0, 08-2017)	
10.2	Prüfung, ob die zentralen Komponenten (insb. Registry) der GDI-DE Mehrsprachigkeit ausreichend unterstützen	<u>AK Architektur</u> AG Geodaten		
10.3	Proof of Concept: Aus Sicht der AG Geodaten eignet sich das Thema zur Vertiefung in einem Pilotprojekt. Hier sollte insbesondere die Bereitstellung mehrsprachiger Datenschemata und Codelisten sowie das Zusammenspiel zwischen der jeweiligen Registry (zur Realisierung der Mehrsprachigkeit) und den Softwarelösungen, die eine mehrsprachige Umgebung anbieten, untersucht werden.	<u>AG Geodaten</u> (Erstellung der Projektskizze, Projektkoordination) AK Architektur AK Metadaten GDI-DE Betriebsstelle		Abstimmung zwischen den Akteuren in 2019