

Jak pracovat s datovými specifikacemi INSPIRE

Miroslav Fanta
CENIA, oddělení geoinformatiky
poslední aktualizace: 17.7.2020
miroslav.fanta@cenia.cz

Cílem dokumentu je v maximální možné míře zjednodušit, zestručnit a shrnout základní informace potřebné pro přípravu a vytvoření dat, která budou vyhovovat požadavkům směrnice INSPIRE, aby pracovníci připravující data nemuseli procházet několikasetstránkové manuály a návody a často poměrně složitá schémata.

Úvod

Témata INSPIRE

Téma INSPIRE je základní jednotka, dle které se data poskytovaná do systému INSPIRE člení. Témata jsou rozdělena do tří skupin (Annexů).

Kód

tématu

ANNEX I.

RS	Coordinate reference systems	Souřadnicové referenční systémy
GG	Geographical grid systems	Systémy geografických souřadnicových sítí
GN	Geographical names	Zeměpisná jména
AU	Administrative units	Územní správní jednotky
AD	Addresses	Adresy
CP	Cadastral parcels	Parcely
TN	Transport networks	Dopravní sítě
HY	Hydrography	Vodstvo
PS	Protected sites	Chráněná území

ANNEX II.

EL	Elevation	Nadmořská výška
LC	Land cover	Krajinný pokryv
OI	Orthoimagery	Ortofotografické zobrazení
GE	Geology	Geologie

ANNEX III.

SU	Statistical units	Statistické jednotky
BU	Buildings	Budovy
SO	Soil	Půda
LU	Land use	Využití území
HH	Human health and safety	Lidské zdraví a bezpečnost
US	Utility and Government services	Veřejné služby a služby veřejné správy
EF	Environmental monitoring facilities	Zařízení pro sledování životního prostředí
PF	Production and industrial facilities	Výrobní a průmyslová zařízení

AF	Agricultural and aquaculture facilities	Zemědělská a akvakulturní zařízení
PD	Population distribution – demography	Rozložení obyvatelstva - demografie
AM	Area management / restriction / regulation zones and reporting units	Správní oblasti, chráněná pásma, regulovaná území a jednotky podávající hlášení
NZ	Natural risk zones	Oblasti ohrožené přírodními riziky
AC	Atmospheric conditions	Stav ovzduší
MF	Meteorological geographical features	Meteorologické geografické prvky
OF	Oceanographic geographical features	Oceánografické prvky
SR	Sea regions	Mořské oblasti
BR	Bio-geographical regions	Biogeografické regiony
HB	Habitats and biotopes	Stanoviště a biotopy
SD	Species distribution	Rozložení druhů
ER	Energy resources	Energetické zdroje
MR	Mineral resources	Nerostné suroviny

Výběr stránky tématu: <https://inspire.ec.europa.eu> -> INSPIRE Themes

<https://inspire.ec.europa.eu/Themes/Data-Specifications/2892>

Stránka obsahuje přehled témat INSPIRE ze všech Annexů. Po výběru daného tématu se dostaneme na **stránku tématu**, kde je v hlavní části stránky k dispozici řada odkazů vztahujících se k danému tématu.

Registr témat INSPIRE je zde: <https://inspire.ec.europa.eu/theme>

Aplikační schéma (Application schema)

Aplikační schéma je podmnožinou tématu INSPIRE. Téma může mít jedno nebo více aplikačních schémat. Například téma III/6 Utility and Government services (Veřejné služby a služby veřejné správy) má 8 Aplikačních schémat.

Na stránce každého tématu (viz výše) jsou aplikační témata uvedena v sekcích Data Models a Data Schema. Každé Aplikační schéma má svůj UML model, XSD schéma a Mapovací tabulku (viz níže).

Příklad: stránka tématu Veřejné služby a služby veřejné správy

(<https://inspire.ec.europa.eu/Themes/136/2892>) , část sekce Data Models, Aplikační schémata jsou podtržena červeně



[Data Models](#)

Administrative and Social Governmental Services

> [UML Model](#)

Common Utility Network Elements

> [UML Model](#)

Electricity Network

> [UML Model](#)

Environmental Management Facilities

> [UML Model](#)

Registr aplikačních schémat a jejich přiřazení k jednotlivým tématům INSPIRE je k dispozici zde: <https://inspire.ec.europa.eu/applicationschema>.

Číselníky (Code lists)

Číselníky, kódovníky, klasifikační třídy, předdefinované povolené hodnoty údaje (parametru). Při tvorbě výsledných souborů prostorových dat pro INSPIRE je žádoucí u parametrů obsahujících hodnoty z číselníků uvádět odkaz na tyto číselníky v registru číselníků INSPIRE (např. pro hodnotu air z číselníku Media to bude <http://inspire.ec.europa.eu/codelist/MediaValue/air>)

Registr všech číselníků INSPIRE včetně jejich přiřazení k aplikačním schémátům a tématům je k dispozici zde: <https://inspire.ec.europa.eu/codelist>

Výčty (Enumerations)

Podobný význam jako číselníky mají i výčty (Enumerations).

Registr výčtů INSPIRE včetně jejich přiřazení k aplikačním schémátům a tématům je k dispozici zde: <https://inspire.ec.europa.eu/enumeration>

Vybrané obecné pojmy používané v datových specifikacích

Multiplicity

násobnost údaje (parametru)

1 – povinný údaj (parametr), právě jeden výskyt

1..* – povinný údaj, minimálně jeden výskyt

0..1 – nepovinný údaj, maximálně jeden výskyt

0..* – nepovinný údaj, může být více výskytů

Voidable charakteristiky (parametry)

Voidable parametr nebo hodnota v případě datových specifikací INSPIRE znamená, že se kalkuluje s možností, že příslušný údaj či parametr není k dispozici ve zdrojové datové sadě připravované poskytovatelem dat nebo ho nejde z existujících dat odvodit.

„Povinnost“ poskytnout v datech INSPIRE určitý údaj (vyplnit parametr), je tedy třeba posuzovat dle požadované násobnosti a dle toho, zda je daný údaj voidable. Tj. je-li pro určitý parametr multiplicity 1 či 1..* a zároveň je tento parametr voidable, pak, pokud daný parametr (údaj) ve zdrojových datech existuje, stává se povinným. Pokud daný údaj ve zdrojových datech neexistuje, pak i při násobnosti 1 či 1..* daný parametr nemusí být vyplněn, resp. parametr se ponechá prázdný a doplní se

VoidReasonValue – zde jsou možné následující hodnoty:

- Unpopulated - nevyplněný
- Unknown – neznámý
- Withheld – skrytý (údaj existuje, ale z důvodu utajení není poskytnut)

Základní zdroje informací o datových specifikacích

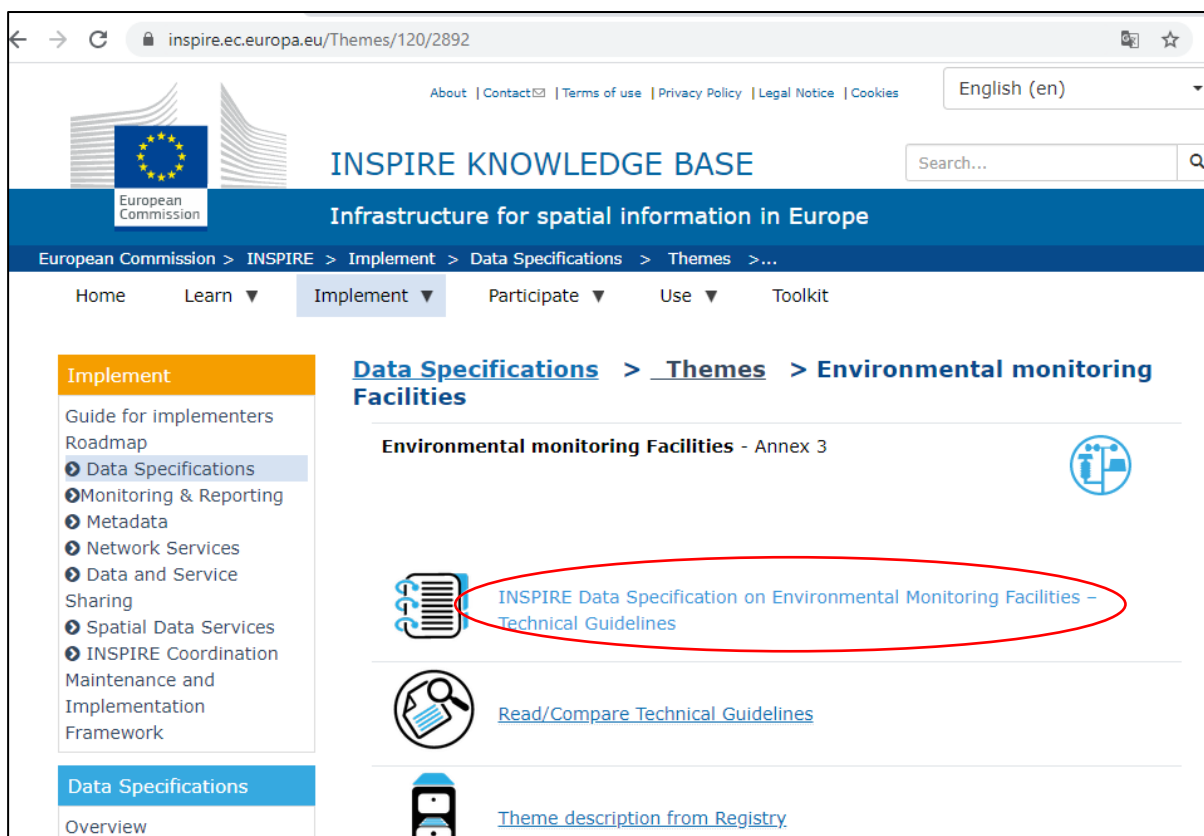
1. Dokument INSPIRE Data Specification on {název tématu} – Technical Guidelines

<https://inspire.ec.europa.eu/id/document/tg/{kód tématu}>

např: **INSPIRE Data Specification on Environmental monitoring facilities – Technical Guidelines**

<https://inspire.ec.europa.eu/id/document/tg/ef>

Po výběru daného tématu se dostaneme na stránku tématu a odkaz na dokument *INSPIRE Data Specification on ... - Technical guidelines* je vždy uveden jako první



Technical guidelines pro všechna témata jsou k dispozici rovněž zde:

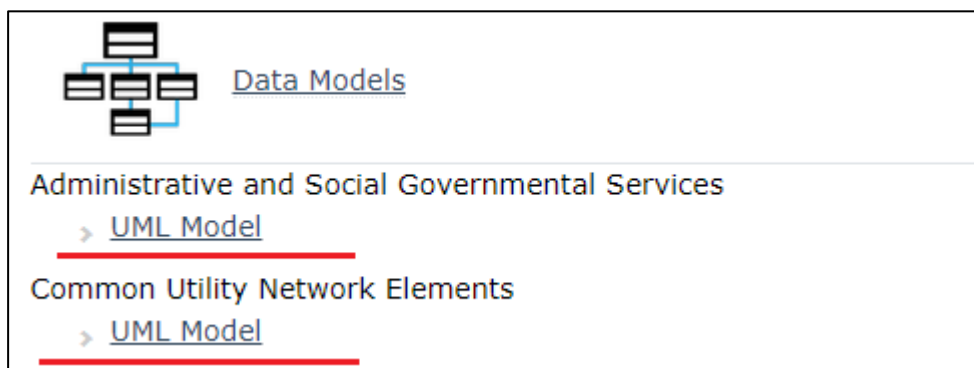
<https://inspire.ec.europa.eu/Technical-Guidelines/Data-Specifications/2892>

Technical guidelines či jeho jednotlivé části lze rovněž procházet interaktivně po kliknutí na odkaz Read / Compare Technical Guidelines (pod odkazem na vlastní Technical Guidelines), kde se následně v levém sloupečku zobrazí základní struktura uvedeného dokumentu a lze pro prohlížení vybírat jednotlivé části.

Vlastní struktura požadovaných dat je v Technical guidelines dokumentech uvedena v části „Data content and structure“, v rámci níž se nachází podkapitola(y) „Application schema...“ a v ní popis a vlastnosti jednotlivých atributů prostorových dat („Feature catalogue“).

2. UML model

Odkaz na UML model pro každé Aplikační schéma je k dispozici na stránce daného tématu v sekci Data Models.



UML modely byly vytvořeny v software Enterprise Architects a pro jejich správnou interpretaci a porozumění je žádoucí mít s těmito modely určitou zkušenost.

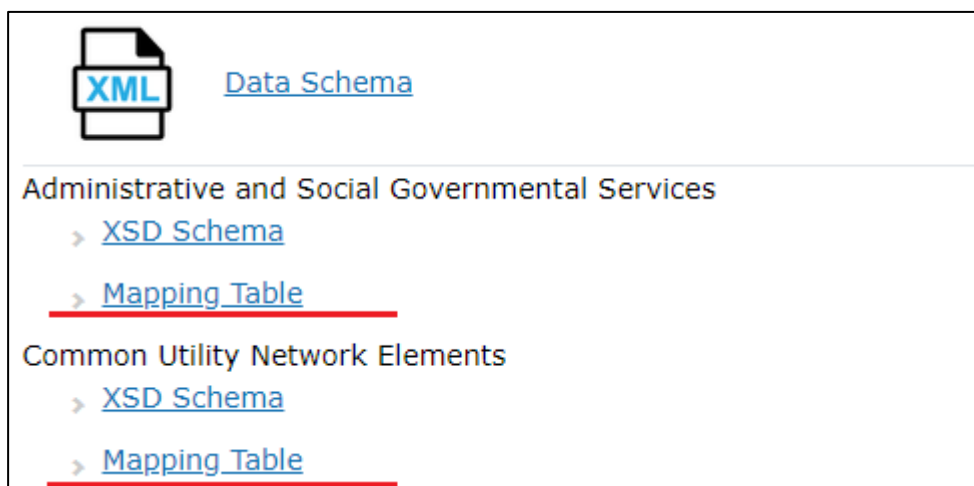
Komplexní UML model INSPIRE pro všechna témata je k dispozici zde:

<https://inspire.ec.europa.eu/data-model/approved/r4618/html/> . Např. Annex III -> Environmental Monitoring Facilities -> <<applicationSchema>> Environmental Monitoring Facilities -> grafické schéma Environmental Monitoring Facility, odkazy na jednotlivé kódovníky atd.

3. Mapovací tabulka (Mapping table)

Mapovací tabulka je primárně určena jako pomůcka pro specifikaci transformace připravovaných dat do struktury požadované v INSPIRE (vazba mezi atributy existujícími v připravovaných datech a atributy požadovanými v INSPIRE), ale po drobné úpravě rovněž dobře poslouží jako kompaktní a snadno pochopitelný přehled jaké atributy jsou pro dané téma a datovou sadu požadovány, zda jsou tyto atributy povinné apod.

Odkaz na mapovací tabulku pro každé aplikační schéma je k dispozici na stránce daného tématu v sekci Data Schema.



Na odkaz „Mapping Table“ klikněte pravým tlačítkem myši, z kontextového menu zvolte možnost „Ulož jako...“ a uložte soubor jako XML dokument, následně otevřete v Excelu a přeuložte jako xlsx soubor. Dále je vhodné takto získanou tabulku formálně upravit (smazat prázdné sloupce, upravit šířku vyplněných sloupců + výšku řádků, ve vhodných případech doplnit hodnoty číselníků apod.)

Všechny mapovací tabulky jsou dostupné ze stránky Data Specifications > Data Models (<https://inspire.ec.europa.eu/Data-Models/Data-Specifications/2892>) -> Mapping tables – dostaneme se na stránky <https://inspire.ec.europa.eu/data-model/approved/r4618/mapping/>, resp. <https://inspire.ec.europa.eu/data-model/approved/r4618-ir/mapping/>, přičemž platí, že stránka .../r4618/... obsahuje mapovací tabulky pro datové modely s rozšířeními a stránka .../r4618-ir/... pouze mapovací tabulky pro datové modely bez rozšíření.

Např. stránka .../r4618/... obsahuje "Soil Mapping Table" i "SoilContamination Mapping Table", ale stránka .../r4618-ir/... pouze "Soil Mapping Table".

Příklad:

Upravená mapovací tabulka pro téma Environmental Monitoring Facilities, Aplikační schéma Environmental Monitoring Facilities, typ objektu Environmental monitoring facility (část).

Type	Documentation	Attribute / Association role / Constraint	Attribute / Association role / Constraint documentation	Values / Enumerations	Multi plicity	Voidable / Non-Voidable
EnvironmentalMonitoringFacility Supertypes: AbstractMonitoringFeature AbstractMonitoringObject	environmental monitoring facility A georeferenced object directly collecting or processing data about objects whose properties (e.g. physical, chemical, biological or other aspects of environmental conditions) are repeatedly observed or measured. An environmental monitoring facility can also host other environmental monitoring facilities. NOTE 1: An EnvironmentalMonitoringFacility is not a facility in the common INSPIRE sense realised by the Generic Conceptual Model class ActivityComplex. NOTE 2: Laboratories are not EnvironmentalMonitoringFacilities from an INSPIRE perspective as the exact location of the laboratory does not add further information to the measurement. The methodology used in the laboratory should be provided with observational data.	inspireId	External object identifier.	Identifier	1	
		name	Plain text denotation of the AbstractMonitoringObject.	CharacterString	0..*	voidable
		additionalDescription	Plain text description of additional information not fitting in other attributes.	CharacterString	0..1	voidable
		mediaMonitored	Monitored environmental medium.	MediaValue * air * biota * landscape * sediment * soil/ground * waste * water	1..*	
		legalBackground	The legal context, in which the management and regulation of the AbstractMonitoringObject is defined.	LegislationCitation	0..*	voidable
		responsibleParty	Responsible party for the AbstractMonitoringObject.	RelatedParty	0..*	voidable
		geometry	Geometry associated to the AbstractMonitoringObject. For mobile facilities the geometry represents the area the facility is expected to measure in.	GM_Object	0..1	
		onlineResource	A link to an external document providing further information on the AbstractMonitoringObject.	URL	0..*	voidable
		purpose	Reason for which the AbstractMonitoringObject has been generated.	PurposeOfCollectionValue	0..*	voidable
		observingCapability	A link pointing to the explicit capability of an AbstractMonitoringObject. This provides a clean link between the observed property, the procedure used as well	ObservingCapability	0..*	
		broader	A link pointing to a broader AbstractMonitoringObject (a higher level in a hierarchical structure). The association has additional properties as defined in the association	AbstractMonitoringObject	0..1	voidable

Takto upravená tabulka dává patrně nejlepší a nejjednodušší přehled o tom, jaký je popis, kterého atributu, která pole jsou povinná, která jsou voidable, jaká je jejich násobnost, která mají být vyplněna hodnotou z číselníku apod.

4. Katalog prvků (Feature catalog)

Dostupný ze stránky Data Specifications > Data Models (<https://inspire.ec.europa.eu/Data-Models/Data-Specifications/2892>) -> FC

nebo přímo na následujících odkazech:

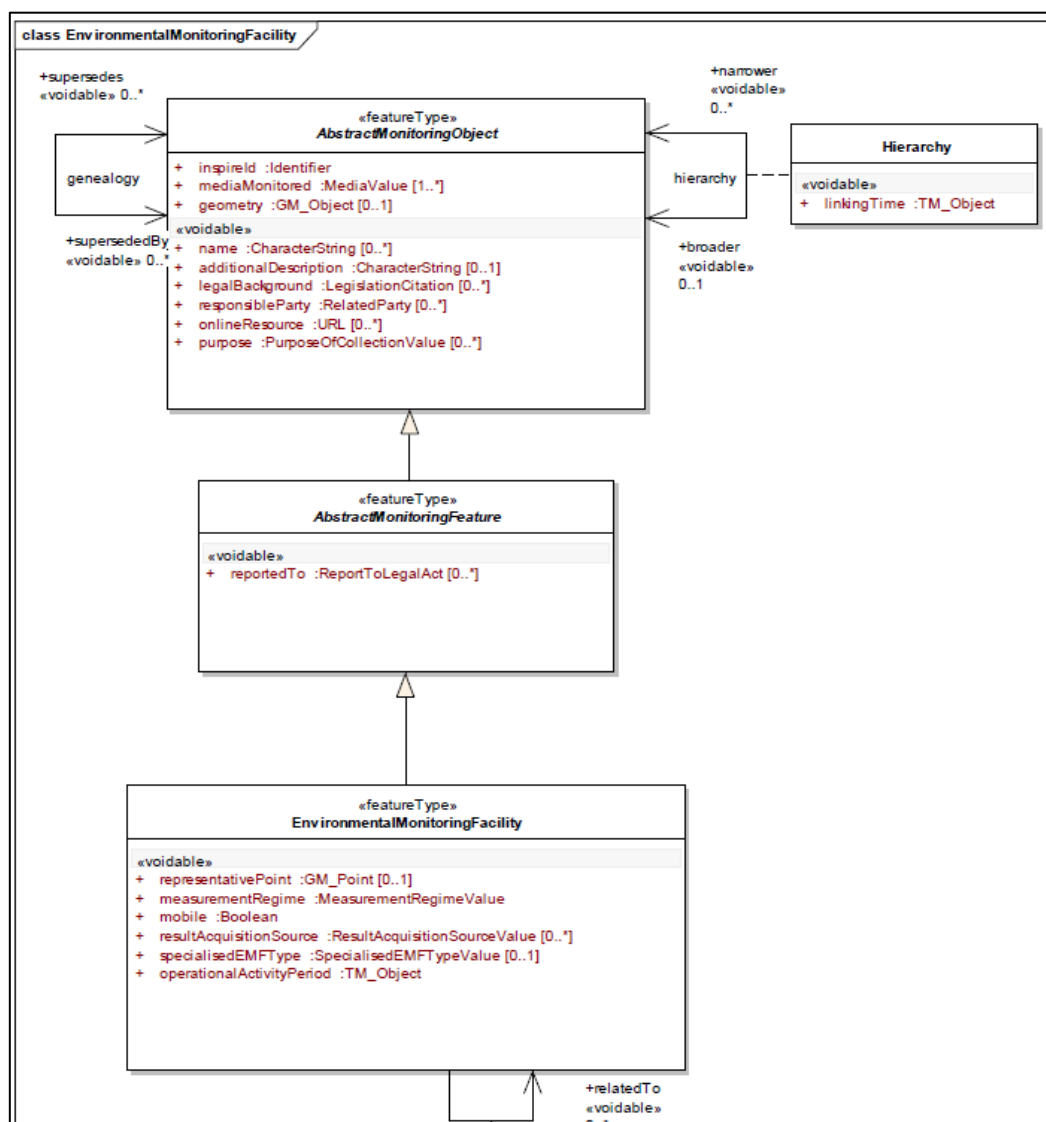
Katalog prvků pro datové modely s rozšířeními: <https://inspire.ec.europa.eu/data-model/approved/r4618/fc/>

Katalog prvků pro datové modely bez rozšíření: <https://inspire.ec.europa.eu/data-model/approved/r4618-ir/fc/>

Katalog prvků je členěn dle abecedně seřazených Aplikačních schémat. Poskytuje analogické informace jako mapovací tabulky, obsahuje i popis položek kódovníku, odkazy na strukturu identifikátorů apod.

Při vyhledávání informací atributů požadovaných pro jednotlivé typy prostorových objektů je třeba brát do úvahy hierarchii zakomponovanou v katalogu prvků – tedy např. u samotného objektu EnvironmentalMonitoringFacility na první pohled nenalezneme očekávaný atribut InspireId. Zjistíme ale, že tento objekt je podtypem položky AbstractMonitoringFeature a ta je zase podtypem pro AbstractMonitoringObject, kde se již InspireId vyskytuje. Z tohoto důvodu se zdá používání mapovací tabulky příjemnější.

Hierarchické vyjádření výše zmíněných vazeb EnvironmentalMonitoringFacility -> AbstractMonitoringFeature -> AbstractMonitoringObject v UML modelu:



5. Presentace CENIE (vytvořila L. Rejentová)

<https://geoportal.gov.cz/web/guest/legislativa-eu/>

Pro vybraná témata je připravena prezentace popisující strukturu dat pro dané téma a přehled parametrů.

Na výše uvedené stránce rozkliknout „Datové specifikace“. Zde jsou přímo odkazy na datové specifikace jednotlivých témat INSPIRE (viz výše „1. Dokument INSPIRE Data Specification on {název tématu} – Technical Guidelines“). Prezentace CENIE jsou připraveny pro témata, u kterých je za názvem tématu uvedeno v závorce slovo „schéma“, pod kterým je odkaz na uvedenou prezentaci. Tento odkaz je doplněn i do následujícího přehledu. Jedná se o následující témata:

Příloha (Annex) I.

I.9 Chráněná území ([schéma](#))

Příloha (Annex) II.

II.2 Krajinový pokryv ([schéma](#))

II.4 Geologie ([schéma](#))

Příloha (Annex) III.

III.3 Půda ([schéma](#))

III.4 Využití území ([schéma](#))

III.5 Lidské zdraví a bezpečnost ([schéma](#))

III.6 Veřejné služby a služby veřejné správy ([schéma](#))

III.7 Zařízení pro sledování životního prostředí ([schéma](#))

III.8 Výrobní a průmyslová zařízení ([schéma](#))

III.9 Zemědělská a akvakulturní zařízení ([schéma](#))

III.11 Správní oblasti, chráněná pásma, regulovaná území a jednotky podávající hlášení ([schéma](#))

III.12 Oblasti ohrožené přírodními riziky ([schéma](#))

III.13 Stav ovzduší ([schéma](#))

III.14 Zeměpisné meteorologické prvky ([schéma](#)) (prezentace společná pro témata III.13 a III.14)

III.20 Energetické zdroje ([schéma](#))

III.21 Nerostné suroviny ([schéma](#))

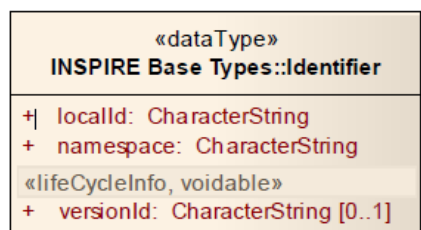
Identifikátory

InspireId: Jedinečný identifikátor v rámci INSPIRE

InspireId (Inspire identifier) Je definován jako externí identifikátor objektu, tj. jako identifikátor publikovaný poskytovatelem dat s tím, že může být použit třetími stranami (externími aplikacemi) pro identifikaci prostorového objektu („to reference the spatial object“).

Základní podklad: INSPIRE Generic Conceptual Model D2.5 v3.4, kapitola 14

https://inspire.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc3.pdf



Unikátní identifikátor prostorového objektu v rámci INSPIRE se skládá ze dvou povinných částí a jedné části nepovinné.

Povinné části identifikátoru InspireId:

namespace – identifikuje zdroj dat (datovou sadu). Překládá se jako „Jmenný prostor“. Namespace je specifikován poskytovatelem dat a měl by být registrován v **INSPIRE External Object Identifier Namespaces Register**; namespace je specifikováno jako URI (Uniform Resource Identifier; někdy používán rovněž termín Unique Resource Identifier). Namespace se typicky skládá ze dvou částí – první identifikuje poskytovatele v rámci země a druhá rozlišuje datové zdroje poskytované a spravované jedním poskytovatelem. První 2 písmena poslední části namespace (za posledním lomítkem, pokud se lomítko v namespace vyskytuje) má být zkratka země. Jako namespace je v některých příkladech uváděna URL adresa.

Příklady:

- FR.IGNF.BDCARTO (BDCARTO je datový produkt (datová sada))
- _EGGR.ERM (_EGGR = organizace EuroGeographics, ERM = datový produkt Euro Regional Map)
- <http://gdi.uba.de/arcgis/rest/services/inspire/DE.UBA.AQD>
- CZ-00025712-CUZK_HY

localId – identifikuje konkrétní objekt v této sadě, je unikátní v rámci namespace; může obsahovat příponu rozlišující typ objektu či hierarchii. Je rozhodně výhodné, pokud první 2 znaky localId obsahují rovněž zkratku země, i když tento požadavek není výslovně uveden.

příklady:

- IT123456.SITE (SITE = přípona localId rozlišující typ objektu)
- SKUEV0319
- ReportingHeader_B_Zones (první 2 znaky nejsou zkratka země)

Nepovinná část identifikátoru InspireId:

versionId, - časový údaj, případně doplněný dalšími informacemi (proto má tento parametr datový typ character a nikoliv date)

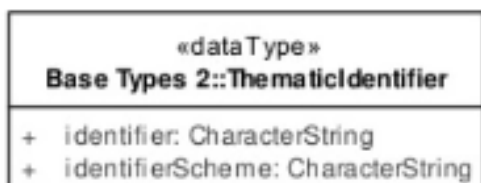
Příklady:

- prelim-2014-v2
- 2015-01--01

ThematicId

Kromě InspireId se lze v INSPIRE setkat rovněž s pojmem **ThematicId** (resp. **thematic identifier**). Tento je definován jako „real-world phenomenon“, tj. jev reálného světa. Jako příklady se uvádějí kódy NUTS nebo ICAO kódy pro identifikaci letišť (ICAO = Mezinárodní organizace pro civilní letectví). Používá se rovněž pro zajištění provázanosti objektů reportovaných v rámci INSPIRE s objekty reportovanými v rámci jiných reportovacích povinností.

Thematic identifier se skládá z částí Identifier a IdentifierScheme. Násobnost ThematicId je 0..*. ThematicId může být shodný s localId nebo s jeho částí.



Rozdíl mezi InspireId a ThematicId je specifikován následovně:

*„The key difference between the **inspireId** and **thematicId** is that the **inspireId** shall be a persistent, unique identifier that can be used in external datasets to reference to the spatial object by any third party. Whereas the **thematicId** is a descriptive unique object identifier assigned to the spatial object defined in an information community. “*

Životní cyklus prostorových objektů

Pro specifikaci životního cyklu se používají atributy:

beginLifespanVersion

endLifespanVersion

Jedná se o časové údaje. Je logické, že hodnota beginLifespanVersion musí být starší nežli endLifespanVersion. Pokud je údaj o životním cyklu pro danou datovou sadu požadován, a přitom není k dispozici, všechny objekty v této datové sadě budou mít v příslušných attributech prázdnou hodnotu a jako důvod bude uvedeno „unpopulated“.