Tento projekt je financován se státní podporou Technologické agentury ČR v rámci Programu BETA2.

www.tacr.cz

Výzkum užitečný pro společnost.

MINISTERSTVO VNITRA ČESKÉ REPUBLIKY

Terminologické a věcné vymezení objektů NaSaPO

Výsledek V02 – metodika (Nmet)

Konečný uživatel výsledků: **Ministerstvo vnitra České republiky**

Náměstí Hrdinů 1634/3

140 21 Praha 4 – Nusle

Název projektu: NaSaPO – Národní sada prostorových objektů

Číslo projektu: **TITIMV706**

Řešitel projektu: Západočeská univerzita v Plzni

Další řešitelé: Výzkumný ústav geodetický, topografický a kartografický, v. v. i.

GEOREAL, spol. s r. o.

T-MAPY, spol. s r. o.

Doba řešení: 1. 1. 2020 – 31. 12. 2021

Důvěrnost a dostupnost: veřejně nedostupný

Informace o autorském týmu:

Ing. Radek Makovec
(Výzkumný ústav geodetický, topografický a kartografický, v. v. i.)

Ing. Jaroslav Zemek, CSc.
(Výzkumný ústav geodetický, topografický a kartografický, v. v. i.)

RNDr. Eva Sovjáková
(Výzkumný ústav geodetický, topografický a kartografický, v. v. i.)

Doc. Ing. Václav Čada, CSc.
(Západočeská univerzita v Plzni, Fakulta aplikovaných věd, katedra geomatiky)

Doc. Ing. Mgr. Otakar Čerba, Ph.D.
(Západočeská univerzita v Plzni, Fakulta aplikovaných věd, katedra geomatiky)

Další informace o projektu:

Zpracování terminologického a věcného vymezení objektů NaSaPO nejen pro účely projektu NaSaPO a podporu Strategie rozvoje prostorových informací v ČR do roku 2020 (GeoInfoStrategie).

Tento dokument vznikl v rámci řešení programu BETA2 projektu č. TITIMV706 s názvem „NaSaPO – Národní sada prostorových objektů“ s finanční podporou TA ČR.

Program veřejných zakázek v aplikovaném výzkumu a inovacích pro potřeby státní správy BETA2 byl schválen usnesením vlády České republiky č. 278 ze dne 30. 3. 2016 a je zaměřen na podporu aplikovaného výzkumu a inovací pro potřeby orgánů státní správy. Poskytovatelem finančních prostředků je Technologická agentura ČR.

Tento projekt je financován se státní podporou Technologické agentury ČR v rámci Programu BETA2.
www.tacr.cz
Výzkum užitečný pro společnost.

MINISTERSTVO VNITRA ČESKÉ REPUBLIKY

Obsah

I. část – Terminologické vymezení a konceptuální formalismus 6

1 Úvod 7

2 Termíny a zkratky 8

3 Terminologické a věcné vymezení objektů NaSaPO 9

3.1 Přístup k terminologii 10

3.2 Standardy a normy 11

3.2.1 Úrovně norem 13

3.2.2 Vytvoření standardu a normy 15

3.2.3 Proč potřebujeme normy 16

3.3 Terminologie ISO norem geografické informace 19

3.3.1 Jev, abstrakce a koncept 20

3.3.2 Základní terminologie v konceptuálním modelu 23

3.3.2.1 Vztah k terminologii specifikace UML 24

3.3.2.2 Vztah k terminologii v projektech INSPIRE 24

3.3.3 Obor úvah ISO norem geografické informace 25

3.3.3.1 Referenční model 29

3.3.3.2 Interoperabilita 32

3.3.3.3 Čtyřvrstvá architektura ISO norem geografické informace 35

3.4 Podkladové terminologické slovníky 39

3.4.1 Slovník pojmů GeoInfoStrategie 39

3.4.2 Terminologický slovník zeměměřictví a katastru nemovitostí 39

3.4.3 Terminologický výkladový slovník pro potřeby realizace Akčního plánu Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020 40

3.4.4 Terminologie ISO norem geografické informace v gesci ISO/TC211 40

3.4.5 Další terminologické slovníky 41

3.5 Principy tvorby mezioborových slovníků 42

3.5.1 Principy křížového mapování mezi doménovými (oborovými) slovníky 43

3.5.2 Metodika mezidoménového křížového mapování 44

3.5.3 Registr položek mezidoménového křížového mapování 46

4 Metadatové struktury a výměnné formáty pro třídy objektů NaSaPO 47

4.1 Základní metadata 49

4.1.1 Rozšíření pro získávání a zpracování 52

4.1.2 Výměnný formát metadat 52

4.1.3 Specifikace datového produktu 53

4.2 Katalog vzhledů jevů a pojmové slovníky vzhledů jevů 57
4.2.1 Katalog vzhledů jevů 58
4.2.2 Pojmové slovníky 60
4.3 Kvalita 63
4.3.1 Požadavky na kvalitu dat 64
4.3.2 Prvky kvality dat 66
4.3.3 Hodnocení kvality 67
4.3.3.1 Hodnocení kvality dat 68
4.3.3.2 Hodnocení kvality služeb nad prostorovými daty 69
4.3.4 Formy výsledného hodnocení 69
4.3.5 Publikování výsledků hodnocení kvality 70
4.4 Implementace metadat 71
4.4.1 Rozšíření metadat do komunitního profilu 72
4.5 Garance za datový produkt 73
4.5.1 Všeobecné požadavky na garantovaná prostorová data 74
4.5.2 Certifikace prostorových dat 75
4.5.3 Abstraktní sestava zkoušek kvality dat 75
4.5.4 Akreditační a certifikační instituce pro oblasti ICT 76
4.6 Archivace prostorových dat 77

5 Klasifikace datových modelů 78
5.1 Klasifikace digitálních modelů v rámci NIPI 80

6. Závěr terminologického a věcného vymezení objektů NaSaPO 83

Seznam příloh 84

Literatura 85

II. část – Metodika pro budování otevřené a rozšířené ontologické struktury (NMet – certifikovaná metodika) 87

7 Metodika pro budování otevřené a rozšířené ontologické struktury 88
7.1 Úvod do metodiky pro tvorbu ontologie 88
7.2 Metodika pro tvorbu ontologie 90
7.3 Specifikace ontologie 90
7.4 Jádrová ontologie 91
7.5 Příklad konceptu ontologie – Budova 96

Literatura 98

Tento projekt je financován se státní podporou
Technologické agentury ČR v rámci
Programu BETA2.

www.tacr.cz
Výzkum užitečný pro společnost.

MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

I. část

Terminologické vymezení a konceptuální formalismus

Zpracovali: Ing. Radek Makovec, VÚGTK, v. v. i.
Ing. Jaroslav Zemek, CSc., VÚGTK, v. v. i.
RNDr. Eva Sovjáková, VÚGTK, v. v. i.

1 Úvod

Vlastní řešení projektu NaSaPO je založeno na využití ISO norem geografické informace. Úvodní část tohoto dokumentu je proto věnována těmto normám a zdůvodnění, proč je potřebné a výhodné je aplikovat na řešení a provozování NaSaPO i při rozvoji Národní infrastruktury pro prostorové informace. Pro správné a jednotné chápání a používání terminologie ISO norem, ale i dalších standardů, které většinou z ISO norem vychází, jsou v dokumentu podrobně vysvětleny základní termíny, např. objekt, jev, vzhled jevu, abstrakce a jiné, které jsou často i odborníky v oblasti prostorových informací zaměňovány, nebo jim je přisuzován ambivalentní, mnohdy až kolidující význam.

Tento výsledek řešení sestává ze dvou částí, z nichž první je věnována principům konceptuálního formalismu a popisu jeho praktického využívání při konceptuálním modelování opírajícím se o co nejširší aplikování norem ISO řady 19100. Podává základní informace o obsahu a způsobu aplikování těchto norem při navrhování, výrobě, distribuci a archivaci prostorových informací. V úvodu první část seznamuje se systémem jejich tvorby, s druhy norem i s jejich postavením v české legislativě. Velký prostor je dále věnován terminologickým otázkám, definicím základních a nejfrekventovanějších pojmů i metodám a principům tvorby pojmových a mezioborových terminologických slovníků. Jsou prezentovány výsledky rozsáhlé rešerše a stručné vyhodnocení veřejně dostupných terminologických slovníků z oblasti geografické informace. K hlavním slovníkům jsou poskytnuty také údaje k jejich zpřístupnění.

Jednotlivé termíny a jejich definice byly převzaty z relevantních norem ISO a jejich význam, pravidla a možné způsoby použití jsou podrobně vysvětleny a popsány s využitím četných UML schémat. Řešitelé byli vedeni snahou podat ucelený přehled o postupu při modelování geografických informací a vysvětlit role a vzájemnou provázanost mezinárodních norem při zajišťování interoperability výsledných prostorových informací. Z procesů tvorby datových sad je pozornost věnována zejména dvěma úzce vzájemně souvisejícím tématům – požadavkům na data ve specifikacích datových produktů a metadatům. S ohledem na požadavek poskytovat NaSaPO jako státem garantovaná data je komplexně a poměrně velmi detailně zpracovaná problematika kvality prostorových dat, včetně jejich certifikace.

V závěru této I. části je stručně prezentována Národní infrastruktura pro prostorové informace a provedena klasifikace v ní používaných datových modelů.

2 Termíny a zkratky

Termíny a zkratky v projektu NaSaPO využívané napříč výstupy projektu NaSaPO jsou uvedeny v samostatné příloze A: Terminologický výkladový slovník projektu NaSaPO.

Jednotlivé dokumenty, resp. výstupy projektu NaSaPO, obsahují další specifické termíny a zkratky v nich použité. Pro výstup V02 byly vybrány následující termíny a jejich definice:

2.1 fakt *(fact)*

stav skutečnosti

[ISO 16354:2013]

2.2 definice *(definition)*

vyjádření pojmu pomocí deskriptivní výpovědi, která slouží k jeho odlišení od souvisejících pojmů

[ISO 16354:2013]

2.3 pojem/koncept *(concept)*

1. znalostní jednotka vytvořená jedinečnou kombinací aspektů a/nebo komponent (upraveno podle ISO 1087-1:2000);
2. společné znaky mezi jednotlivými věcmi definované jedním nebo více omezeními popisujícími limity pro zahrnutí jednotlivých věcí odpovídajících danému pojmu

[ISO 16354:2013]

2.4 termín *(term)*

1. verbální označení pojmu nebo jednotlivé věci v určitém oboru (upraveno podle ISO 1087-1:2000);
2. role fyzického řetězce znaků s možnými mezerami nebo fyzického zvuku s pomlčkami, které se používají k označení pojmu (např. druhu fyzického objektu nebo aspektu) nebo k označení jednotlivé věci v konkrétním jazyce (systému kódování) a konkrétní jazykové komunitě

[ISO 16354:2013]

2.5 zkratka *(abbreviation)*

označení vytvořené vypuštěním slov nebo písmen z delší podoby pojmenování a označující též pojem

[ISO 1087-1:2000]

2.6 označení *(designation; designator)*

vyjádření pojmu znakem, který ho označuje (pojmenovává)

POZNÁMKA: V terminologické práci se rozlišují tři druhy označení: značky, vlastní názvy a termíny.

[ISO 1087-1:2000]

2.7 lexikální jazyk *(lexical language)*

jazyk, jehož syntaxe je vyjádřena symboly definovanými jako řetězce znaků (B.36)

[ISO 19101-1:2014]

2.8 **aspekt** *(aspect)*

pojem, jehož prostřednictvím je možno vnímat existenci a vzezření určité věci, který nemůže existovat bez existence svého nositele a který je buď vnitřní (inherentní) neoddělitelnou součástí nahlížení na svého nositele, nebo jeho rolí

[ISO 16354:2013]

3 Terminologické a věcné vymezení objektů NaSaPO

Termíny a zkratky včetně jejich definic jsou jedním z hlavních předmětů zájmu tohoto dokumentu. Správně nastavená a použitá terminologie, jednotné pojetí používaných pojmů je základem k porozumění veškerých potenciálních uživatelů NaSaPO, kterými primárně jsou kraje, obce, správci technické infrastruktury, správci dopravní infrastruktury, stavební úřady, ministerstva, Český úřad zeměměřický a katastrální, Zeměměřický úřad, Státní pozemkový úřad, Český telekomunikační úřad, Český statistický úřad, orgány krizového řízení a složky Integrovaného záchranného systému (Zdravotnická záchranná služba, Policie ČR, Hasičský záchranný sbor ČR), vysoké školy a výzkumné organizace, projekční a geodetické kanceláře, developeři, stavebníci včetně dalších uživatelů z oblasti odborné i laické veřejnosti.

Zcela zásadní je nalezení terminologického konsensu s odborníky především z oblasti informačních technologií, navrhování a řízení staveb, správci dopravní a technické infrastruktury.

Při řešení aplikací nad geografickými daty je používáno mnoho geografických termínů, které se mohou překrývat se stejnými termíny z jiných oborů, majícími však často zcela nebo zčásti odlišný význam. Odlišné a zkreslené chápání nebo vůbec nepochopení přenášeného sdělení vlivem používání nesprávných pojmů tak může mít i v komunitě řešitelů a uživatelů aplikací nad geografickými daty velmi negativní důsledky. Jedním z prvních úkolů vedení projektu bylo logicky zadání vypracovat jednotnou a závaznou terminologii. Protože vlastní koncepce a řešení projektu NaSaPO budou založeny na mezinárodních normách ISO řady 19100 pro geografickou informaci, bude využita celá řada těchto navzájem integrovaných norem, přičemž s ohledem na terminologii je třeba zmínit především normy pro definování referenčního rámce, vytvoření konceptuálních modelů a schémat zpracovaných s využitím zvoleného jazyka konceptuálního schématu (ČSN EN ISO 101-1:2018 Geografická informace – Referenční model – Část 1: Základy, ČSN P ISO/TS 19103:2006 Geografická informace – Jazyk konceptuálního schématu, ČSN EN ISO 19109:2020 Geografická informace – Pravidla pro tvorbu aplikačních schémat, v ní především část věnovaná Obecnému modelu vzhledu jevu – GFM.)

Základy terminologie NaSaPO budou postaveny na důsledném využití ČSN P ISO/TS 19104:2010 Geografická informace – Terminologie. Tato norma zaručuje, že je ve všech normách této řady pro daný pojem (koncept) použitý jeden termín. Tento velmi potřebný princip bude realizován i v tomto projektu a bude tak dosaženo terminologické konzistence u všech jeho řešitelů a uživatelů. Obsahem této normy je i metodický návod pro sběr a údržbu termínů z oboru geomatiky, včetně jejich repository. V projektu budou aplikovány i zásady pro sestavování definic termínů uvedené v této normě. V její příloze B je pak uvedeno několik set do češtiny přeložených závazných základních termínů z oboru geografická informace, včetně jejich definic, které jsou v pravidelném půlročním cyklu doplňovány a revidovány. V každé použité ISO

normě jsou uvedeny i další termíny a definice ze segmentu geografické informace, o kterém norma pojednává. Také tyto termíny se svým obsahem uvedeným v definicích jsou pro projekt NaSaPO závazné.

Geografická informace se často nachází na místě, kde se kříží s jinými odbornými disciplínami. Lze očekávat, že to bude i velmi častým případem při řešení NaSaPO, a to pravděpodobně nejvíce s odvětvím stavebnictví a dopravy. Logicky se tedy budou setkávat i domény jejich terminologií a bude docházet k překrývání termínů nebo základních pojmů. Termíny mohou mít několik významů v závislosti na kontextu a na určitý pojem může být naopak odkazováno několika různými termíny, přičemž každý z nich může vyjadřovat jiné konotace a mít jiný důraz. Metodiku, jak překonat tyto terminologické problémy vyplývající z přijetí technických slovníků různých odborných odvětví geografickou komunitou, a principy tvorby mezioborových slovníků poskytuje norma ČSN EN ISO 19146:2018 Geografická informace – Mezidoménové slovníky. Tato norma se nezabývá ontologií ani taxonomií geografické informace. Ontologie, jako explicitní specifikace konceptu geografické informace, je předmětem skupiny norem ISO řady 19150-x, a taxonomie, jako věda a praxe klasifikace, je řešena částí normy ISO 19144-1:2009 Geographic information – Classification systems – Part 1: Classification system structure. Těmto záležitostem se bude podrobněji věnovat část II.

3.1 Přístup k terminologii

Terminologické potřeby aplikace nad prostorovými informacemi jsou zajištěny z několika pramenů. Nejdůležitějším pramenem je globálně (mezinárodně) založený seznam odborných termínů, které jsou nové a současně jsou klíčovými termíny rozvoje technických specifikací, pro který je zaveden název ISO normy geografické informace.

Tento zdroj je pro globální použití založen v angličtině, překlady do francouzštiny a ruštiny garantuje sekretariát ISO, překlady do jazyků dalších evropských států zabezpečují k tomu určené národní agentury, v České republice je to v pověření ÚNMZ Česká agentura pro standardizaci (ČAS).

Je nutné si uvědomit, že toto je klíčový pramen terminologie pro tu část projektových prací, které budou na mezinárodních normách postaveny. Avšak normy jsou nezávazné, takže je umožněno využít v lokálním jazyce odborný žargon (pojmy/koncepty), příp. přezdívky, anglicky označované jako „alias“. Platí ovšem podmínka, že zvolené „alias“ není v konfliktu s klíčovými termíny modelování v jiných oblastech standardizace, např. SW, databází, komunikačních termínů pro adresování, UML, WWW, interoperability, atd. Zejména v souvislosti s řešením NaSaPO bude velmi důležité správně pochopit a používat především termín „objekt“, který se vyskytuje jako samostatný nebo s četnými adjektivy v celé řadě vzájemně prolnutých oblastí, např. v konceptuálním modelování podle ISO norem geografické informace, v INSPIRE, v UML a mnoha dalších.

V opačném pohledu, pokud někdo chce své práce připravené v lokálním jazyce přenést na úroveň mezinárodní spolupráce, konkrétně v oboru geografické informace/geomatika, představit a porovnat dosažené výsledky na evropské úrovni nebo globálně, pak je to možné jen pomocí odsouhlasených termínů ISO/TC 211. Jednotlivci možná někdy nabudou přesvědčení, že v zahraničí jejich výkladu rozumí, otázkou

však je, co za anglickými termíny s „českým“ obsahem zahraniční experti skutečně chápali. Obdobný problém vzniká i při čtení slovenských odborných textů k projektu INSPIRE.

Jak již bylo řečeno, texty mezinárodních norem musí respektovat také příbuzné obory standardizace, zejména ty, které jsou typově zatříděny do skupiny „Rozvoj informačních technologií, včetně geografické informace“. V seznamech odborných termínů geografické informace tak nesmí být použit žádný z klíčových termínů těchto souvisejících oborů. Pokud se v textu mezinárodní normy pro geografickou informaci tyto výrazy vyskytují, referují o obecných IT modelech a tato skutečnost je u definice tohoto termínu uvedena (např. termíny prvek, třída, balíček v UML). V případě stejných termínů s odlišným významem je nutné uvádět vždy také obor, pro který každá z definic platí, případně její zdroj.

Na evropské úrovni platí stejné zásady. V evropských směrnicích, návodech a technických specifikacích však často nejsou dodržovány, protože jejich překlad primárně zajišťují překladatelky Evropské komise a není zde pro český jazyk zajištěna terminologická jazyková kontrola ze strany ČAS.

Již v dřívějších pracích zadávaných ze strany MV ČR bylo zjištěno, že pro národní úroveň může terminologické sjednocení v oblasti prostorové informace čerpat z řady terminologických slovníků (a tezaurů). Pro vlastní obor aplikací prostorové informace je primárním Terminologický slovník zeměměřictví a katastru nemovitostí (více viz čl. 3.4.2), jehož gestorem je ČÚZK a který je vytvořen a spravován VÚGTK. Dalším, avšak nepříliš rozsáhlým, výsledkem úsilí o národní terminologické sjednocení bylo pořízení Terminologického výkladového slovníku pro potřeby realizace Akčního plánu Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020.

Vybrané termíny pro problematiku NaSaPO jsou uvedeny v příloze A – Terminologický výkladový slovník projektu NaSaPO.

3.2 Standardy a normy

Standardizace má v českém prostředí tradici trvající již po mnoho staletí. Technická normalizace byla na národní úrovni zorganizována na začátku 20. století. V současnosti představuje poměrně velmi široký okruh vzájemně propojených činností. Ten zahrnuje normalizaci (včetně technické), metrologii (cejchování) a (státní) zkušebnictví a ověřování jakosti prakticky ve všech technických oblastech. Především z důvodů zajištění datové a technologické interoperability, ale i z mnoha dalších závažných důvodů je naprosto nezbytné, aby byly i vývoj, řízení a využívání NaSaPO založeny a velmi těsně svázány po celou dobu její životnosti s využíváním standardů a norem.

Je proto užitečné se alespoň stručně seznámit s problematikou standardizace a normami, jejich vývojem a klíčovou úlohou a objasnit způsoby jejich vytváření a využívání.

Standardizace *je proces, který je zásadní součástí obchodního procesu. Specifikuje aktuálně nejlepší, univerzálně dosažitelné vlastnosti produktů, a to při zohlednění reálných požadavků jejich spotřebitelů či uživatelů. Výsledkem standardizace jsou standardy. [10]*

Standard *specifikuje nabídku produktu, jeho technický návrh i obsah a popisuje využití výrobku. Poskytuje pravidla a směrnice a definuje charakteristiky umožňující propojení mezi systémy, daty, lidmi, technickými*

a programovými prostředky a postupy ve firemním, národním, regionálním nebo globálním rozsahu. Standard upřesňuje pravidla a návody k použití produktu, stanoví kritéria posuzování a soubor standardizovaných zkoušek, kterými se ověřuje a zaručuje, že vše, co je použito k realizaci obchodního kontraktu, je pro daná kritéria splněno. [10]

Typy standardů:

- **Standardy de iure (standardy podle práva)** – standardy dosažené dohodou, vytvořené a přijaté v souladu s platnou legislativou. Tyto standardy jsou vytvářeny mezinárodními, regionálními nebo státními organizacemi pověřenými vývojem a přijímáním norem.
- **Standardy de facto (praktické standardy)** – jsou definovány jako skutečnost, která nastala na základě praxe nebo vývojem a používáním technik, které se staly dominantními na trhu. Tyto standardy mohou být neoficiálními, nebo se časem stát povinnými (de jure). Standardy de facto existují, protože byly na trhu první, praxe je přijala a ve svém zájmu je používá dobrovolně. Příkladem de facto standardu je série dokumentů Request for Comments (RFC), které specifikují mnoho protokolů používaných na Internetu – např. široce používaný HTTP protokol sloužící pro přenos webových stránek a dalších dat. Dalšími příklady je klávesnice QWERTY nebo formát papíru A4 používaný v úřední korespondenci.
- **Otevřené standardy** – standardy veřejně přístupné včetně jejich kompletní dokumentace.

Technická normalizace *(normalizace) je činnost, kterou se zavádějí ustanovení pro všeobecné a opakované použití, zaměřená na dosažení optimálního stupně uspořádání v dané souvislosti.*

Technická norma *(norma) je dokument vytvořený na základě konsenzu a schválený uznaným orgánem, poskytující pravidla pro všeobecné a opakované používání.*

Konsenzus *je všeobecný souhlas charakterizovaný tím, že žádná z důležitých zainteresovaných stran nemá odůvodněné námitky vůči podstatným sporným otázkám, a vedený snahou vzít v úvahu stanoviska všech zainteresovaných stran a smírně vyřešit protichůdné argumenty.*

[ČSN EN 45020]

Pod termínem normalizace jsou chápány všechny činnosti související s tvorbou, vydáváním a distribucí technických norem. Tyto činnosti jsou v současné době vykonávány Českou agenturou pro standardizaci (ČAS), která byla zřízena jako státní příspěvková organizace Úřadem pro technickou normalizaci, metrologii a státní zkušebnictví (ÚNMZ) podle § 5 zákona č. 22/1997 Sb., o technických požadavcích na výrobky, ve znění zákona č. 265/2017 Sb. a od něhož oblast normalizace od 1. 1. 2018 převzala. ÚNMZ je správním úřadem a organizační složkou státu v resortu Ministerstva průmyslu a obchodu, zajišťujícím úkoly vyplývající ze zákonů ČR, které upravují technickou normalizaci, metrologii a státní zkušebnictví, a úkoly v oblasti harmonizace technických předpisů a norem v rámci členství ČR v EU. ÚNMZ je členem evropských normalizačních organizací (CEN, CENELEC, ETSI) a mezinárodních normalizačních organizací (ISO, IEC). Normy vytvořené na mezinárodní úrovni jsou přebírány na národní úroveň, resp. evropskou úroveň. Členské země musí příslušnou evropskou normu zavést do 6 měsíců do svých národních norem. Výsledkem normalizace na národní úrovni jsou technické normy označované ČSN.

(Podrobněji o ČAS a normalizaci viz např. http://www.agentura-cas.cz/node/88)

POZNÁMKA 1: Z výše uvedených definic je zřejmé, že se v případě termínů standard a norma i standardizace a normalizace jedná o termíny s obdobným významem. V této dokumentaci bude termín standard používán pro dokumenty vydávané na základě obecné shody organizacemi, které sdružují experty v dané oblasti, a termín norma bude označovat technickou normu vydanou schváleným uznaným orgánem (ÚNMZ, CEN, ISO).

POZNÁMKA 2: Obecně jsou dále termíny standard a norma odlišeny z pohledu typu výstupů. Jako „standard" mohou být označovány libovolné dokumenty (specifikace, doporučení, přepisy, atd.). Jako „norma" jsou označovány pouze technické normy, které mají stanovenou formu.

Specifickou oblastí standardizace v oblasti vojenské a bezpečnostní techniky a postupů je tvorba, vydávání a distribuce obranných standardů NATO. K tomuto účelu byl zřízen Úř OSK SOJ, jehož působnost není omezena jen na vojenské a bezpečnostní složky, ale je dána pro veškeré státní zakázky. V oblasti standardizace tento úřad vydává České obranné standardy, které jsou zpracovávány na základě standardů NATO označovaných STANAG. Část těchto ČOS se zabývá i vojenskými geografickými informacemi, které jsou v gesci pracovní skupiny NATO IGeo WG (Defence Geospatial Information Working Group).

Veřejnoprávní a soukromoprávní instituce mohou vyžadovat povinné používání norem, zejména u veřejných zakázek. Normy se tak stávají smluvně závazné (de iure). Ve vyspělém právním státě jsou technické normy uplatňovány i v soudní praxi. Přesto, že jsou normy nezávazné, jsou v soudním řízení používány jako základní měřítko pro posuzování technického řešení. Splnění normy je tak chápáno jako nezbytné základní minimum pro zajištění bezpečnostních požadavků provozu technických výrobků při vstupu na volný trh, resp. splnění minima pro ochranu života, zdraví a majetku ve vztahu k užívání výrobků.

V českém právním systému lze již v současnosti odkazovat na normy. V závazných právních předpisech se stanoví rámcové požadavky, ke kterým jsou určeny technické normy. Tyto určené normy doporučují, jak těmto základním požadavkům vyhovět při realizaci konkrétního technického řešení. Splnění požadavků norem je předpokladem splnění požadavků právního předpisu.

3.2.1 Úrovně norem

Globální úroveň

Na globální úrovni vydávají normy ISO a IEC (International Electrotechnical Commission) a ITU (International Telecommunication Union). Pro geografickou informaci je vrcholovým centrem pro vydávání norem ISO (International Organization for Standardization) se sídlem v Ženevě. Jednotlivé oblasti, pro které jsou vydávány mezinárodní normy ISO, jsou zastřešeny příslušnou odbornou technicko-normalizační komisí. Geografická informace a geomatika je předmětem činnosti i Technicko-normalizační komise ISO/TC 211. Tato oblast je zařazena do klasifikace: ICS: 35.240.70 – IT applications in science (including digital geographic information). Na globální úrovni jsou stanoveny tři povinné jazyky (angličtina, francouzština a ruština). Vydání normy se označuje „mezinárodní norma ISO".

POZNÁMKA: ISO je obchodní značka, nikoliv zkratka.

Regionální úroveň

Globální úroveň je rozdělena do šesti regionů, přičemž na regionální úrovni ČR patří do působnosti CEN, CENELEC a ETSI, vydávajících evropské a evropské harmonizované normy. CEN je smluvní organizací ISO, avšak platí zvláštní pravidla (viz tzv. Vídeňská dohoda), zjednodušující vzájemné přejímání mezinárodních a evropských norem. Tato dohoda umožňuje také vzájemnou výpomoc při vydávání norem. ISO navrhovanou normu předá do CEN k hlasování ve své územní působnosti a schválenou normu přejímá následně ISO. Pokud je v rámci CEN převzata mezinárodní norma ISO, je k jejímu popisu doplněno „mezinárodní a evropská norma EN ISO“. Je třeba akceptovat, že pro evropskou normu nelze obecně použít termín „mezinárodní“. Zajímavé je, že slovně vyjádřené stupně „mezinárodní“ a „evropská“ se píší v opačném sledu, než jsou řazeny písmenné kódy EN a ISO. Normy nesoucí označení ISO EN musí všechny členské státy CEN do šesti měsíců zavést do národních souborů norem. Evropské normy jsou povinně vydávány v jazycích angličtina, francouzština, němčina.

POZNÁMKA: Adjektivum „harmonizovaná“ znamená, že to původně byla mezinárodní norma ISO.

Národní úroveň

Národní pověřené organizace (v ČR jsou jimi ÚNMZ/ČAS a Úř OSK SOJ) jsou oprávněny stanovit své regulace pro zavádění mezinárodních a evropských norem. Jako instituce členské země EU a NATO musí však respektovat i evropskou legislativu a závazné postupy NATO Agentury pro standardizaci.

Normativní činnost se v ČR řídí především ustanoveními zákona č. 22/1997 Sb., o technických požadavcích na výrobky.

Použití termínu norma je v češtině zavedeno v § 4 zákona č. 22/1997 Sb., který definuje českou technickou normu jako *„dokument schválený Úřadem pro technickou normalizaci, metrologii a státní zkušebnictví (dále jen „Úřad“), pro opakované nebo stálé použití, vytvořený podle zákona a označený písmenným označením ČSN, jehož vydání bylo oznámeno ve Věstníku Úřadu pro technickou normalizaci, metrologii a státní zkušebnictví (dále jen „Věstník Úřadu“). Česká technická norma není obecně závazná.“* Název česká technická norma a písmenné označení ČSN nesmějí být použity k označení jiných dokumentů. Zákon také zavádí definici:

„Soustavu českých technických norem tvoří

- *a) původní české technické normy,*
- *b) evropské či mezinárodní normy přejaté překladem a*
- *c) evropské či mezinárodní normy přejaté v původním jazyce nebo přejaté schválením k přímému používání.“*

Na národní úrovni ČR jsou k dispozici dostatečně účinné prostředky, které mohou učinit ze zásadně nezávazných norem dokumenty závazné. V oboru ICT a geoprostorových informací jsou však tyto prostředky dosud velmi málo využívány.

3.2.2 Vytvoření standardu a normy

Aby standard mohl aspirovat na globální platnost, je třeba získat pro jeho vytvoření co nejširší mezinárodní tým zkušených expertů. Práci expertního týmu je nutné řídit (dohodnout konceptuální rámec a časový plán prací). Na závěr jsou standardy důkladně posouzeny a schváleny většinovým systémem členů, kteří mají hlasovací právo. Před dokončením je standard nabídnut organizacím vydávajícím normy. Někdy je tvorba standardu přímo těmito organizacemi vyvolána formou zadaného projektu. Standardy můžeme tedy označit jako podpůrné výsledky skupiny expertů k normalizaci.

Úroveň	Organizace	Podpora	Poznámka
globální	ISO, IEC, ITU-T	OSN	Nezávislé, nevládní, neziskové normalizační organizace. Členy mohou být vlády, neziskové nebo veřejné obchodní organizace.
regionální	CEN, CENELEC, ETSI	EK	Nařízení Evropského parlamentu a Rady Evropy (EU) č. 1025/2012, o evropské normalizaci
národní	ÚNMZ / ČAS	MPO	Zákon č. 22/1997 Sb., o technických požadavcích na výrobky, v platném znění

Tabulka 1: Organizace vydávající normy

Standardy a normy relevantní pro integrovaný management geoprostorové informace lze s ohledem na širší kontext interoperability systémů, na vazby k dalšímu strategickému směrování a na vývoj a iniciativy e-Government, rozčlenit do několika typů:

- Standardy a normy k obecným účelům ICT – umožňující interoperabilitu mezi různými přístroji, sítěmi, senzory a systémy (např. HTML, IPv6, JPEG, UML, jsou vytvářeny organizacemi, jako jsou např. W3C, ITU, OASIS, IEEE, OMG…)
- Standardy a normy specifických domén – zabývající se obecnými geografickými aplikacemi a technologiemi a aplikacemi geografických informací (např. pro poštovní adresy ISO 19160-1, vojenské geografické informace DIGEST, námořní navigaci S-57, vyvíjené např. organizacemi UPU, MAS, IHO, ICAO, ISO, OGC, DGI WG…)
- Standardy a normy všeobecné geografické informace – představují většinu standardů a norem zabývajících se specificky geografickou informací. Jsou tvořeny především organizacemi ISO/TC 211 a CEN, OGC, DGI WG a lze je dále členit na skupiny:
    - zaměřené na obsah – datové modely, definice a reprezentace dat, jejich kvalita, obecná architektura (jsou představovány většinou norem ISO řady 19100);
    - zaměřené na další aspekty geoprostorové informace, jako jsou popisy dat (metadata), postupy jejich sběru apod.;
    - zaměřené na geoprostorové technologie a geoprostorové služby (WMS, WFS).

Detailnější členění norem ISO pro geografickou informaci, popisy jejich obsahu, vzájemných vazeb, možností a způsobů jejich využívání lze nalézt například v [10] a zejména v [12].

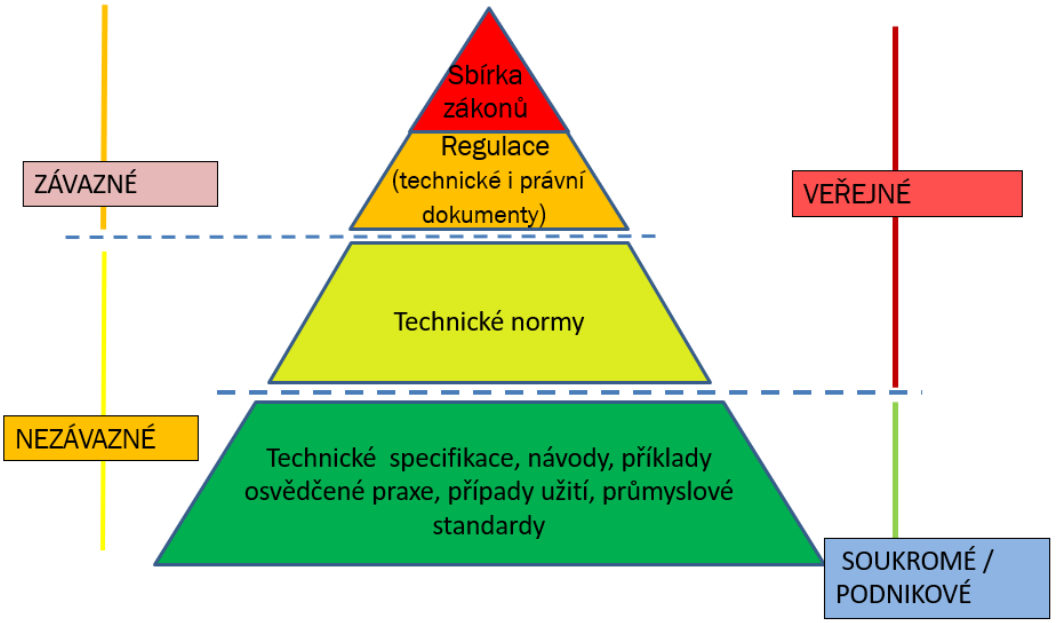
3.2.3 Proč potřebujeme normy

Hovořit v této době stále ještě o klíčovém významu normalizace a snášet argumenty pro tento základní předpoklad úspěšného a perspektivního vývoje a výroby v jakékoliv technické oblasti vyspělé společnosti se může jevit už zbytečné. Nízká úroveň znalostí, malá míra využívání ISO norem řady 19100 u českých projektantů, výrobců a uživatelů geoprostorových dat i nezájem o tyto normy, a někdy dokonce i averze k nim, však svědčí o opaku. Není účelem této analýzy zkoumat příčiny tohoto stavu, patří k nim však bezesporu vysoké náklady na pořízení potřebného souboru norem a nároky na jejich zvládnutí do úrovně praktické implementace.

Investovat do zvládnutí ISO norem geografické informace a jejich implementace v potřebném rozsahu je však naprosto nezbytné a je jednou z kritických podmínek úspěšného vyřešení velmi naléhavé perspektivní problematiky NaSaPO.

Implementace ISO norem řady 19100 má nesporně klíčový význam pro vývoj a provozování NaSaPO a je potřeba akcentovat především následující přednosti a přínosy jejich využívání:

- jsou podmínkou pro domácí i zahraniční obchodování s geografickými daty i jejich výměnu;
- zajišťují vnitřní konzistenci a interoperabilitu geografických dat, umožňující jejich sdílení a opakované užívání k více účelům;
- minimalizují překážky pro sdílení a integrování dat k podpoře vládní politiky, k řídícím a rozhodovacím procesům a pro poskytování veřejných služeb;
- zvyšují schopnosti sdílet geografická data v místním, regionálním a globálním měřítku;
- umožňují rychle mobilizovat nové zdroje dat a technologii a zabránit závislosti technologie na jejím dodavateli (tzv. vendor lock-in);
- usnadňují a urychlují správné a jednoznačné pochopení obsahu a struktury geografických dat projektanty a uživateli v jednotlivých resortech a mezi nimi, ve veřejné správě a v soukromém sektoru i ze strany občanů;
- zefektivňují produkci a management životního cyklu geografických dat;
- podstatně snižují pracovní úsilí, šetří čas a náklady jak při prvotním vývoji a implementaci technologií a produkci geografických dat, tak i při jejich opětovném použití a to i pro jiné účely, než pro které byly původně vyprodukovány;
- zlepšují návratnost investic a s přidáním minimálního úsilí přispívají k rozvoji budoucích systémů s významnými novými schopnostmi.



Obrázek 1: Členění standardů
(Zpracováno podle: Using and referencing ISO and IEC standards to support public policy, ISO/IEC Switzerland, 2015)

Právní předpisy jsou vytvářeny OVM. Standardy a normy nejsou obvykle vytvářené veřejnou správou, a to ani v případě, že OVM má zmocnění „k vydávání standardů“. Vyplývá to z § 2 písm. h) zákona č. 22/1997 Sb., o technických požadavcích na výrobky, v platném znění. Pak ovšem nelze takový „standard“ zařadit do rámce zabezpečování obchodních požadavků a nelze ho použít v zadání veřejných zakázek. Z pohledu standardizace se jedná o prázdné dokumenty, bez jakékoliv váhy.

Příklady z běžné praxe:

- Standard výměnného formátu…
- Norma ministerstva … pro pořizování a výměnu fotogrammetrických výsledků…
- Standard Krajského úřadu Valašského království…
- Standardy ISVS…
- Standard SIS k územní identifikaci.

Kontrola technických právních předpisů, technických dokumentů a technických norem, pokud by vznikaly v procesech VS, je zajištěna pověřeným Úřadem pro normalizaci, metrologii a státní zkušebnictví podle zákona č. 22/1997 Sb., o technických požadavcích na výrobky, v platném znění. Ve svém § 2 písm. h) specifikuje:

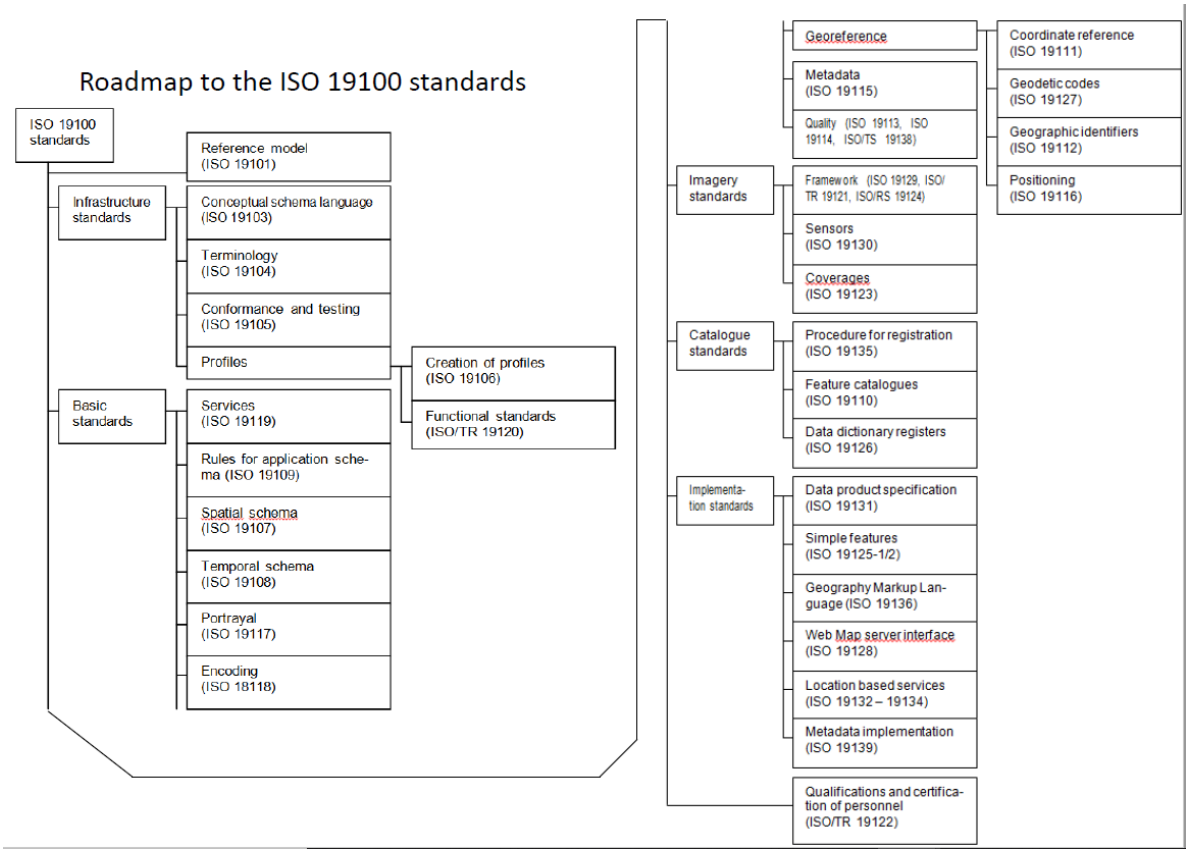
Pro účely tohoto zákona se rozumí technickými požadavky na výrobek

1. *technická specifikace obsažená v právním předpisu, technickém dokumentu nebo technické normě, která stanoví požadované charakteristiky výrobku, jakými jsou úroveň jakosti, užitné vlastnosti, bezpečnost a rozměry, včetně požadavků na jeho název, pod kterým je prodáván, úpravu názvosloví, symbolů, zkoušení výrobku a zkušebních metod, požadavky na balení, označování výrobku nebo*

opatřování štítkem, postupy posuzování shody výrobku s právními předpisy nebo s technickými normami, výrobní metody a procesy mající vliv na charakteristiky výrobků,

2. *jiné požadavky nezbytné z důvodů ochrany oprávněného zájmu nebo ochrany spotřebitele, které se týkají životního cyklu výrobku poté, co je uveden na trh, popřípadě do provozu, např. podmínky používání, recyklace, opětovného použití nebo zneškodnění výrobku, pokud takové podmínky mohou významně ovlivnit složení nebo povahu výrobku nebo jeho uvedení na trh, popřípadě do provozu.*

V současné době je v přímé odpovědnosti technického výboru ISO/TC 211 – Geographic information/Geomatics 75 publikovaných a 24 vyvíjených norem k digitální geografické informaci. Jejich úplný přehled a podrobné informace o působnosti, struktuře a personálním složení pracovních skupin tohoto technického výboru a jejich činnostech lze nalézt na https://www.iso.org/committee/54904.html. Seznam ISO norem geografické informace a informace o jejich využití v jednotlivých úrovních modelování a v procesech celého životního cyklu geografické informace jsou uvedeny v příloze C – Přehled ISO norem geografické informace. Nejdůležitější z těchto norem, které jsou setříděny do skupin podle účelu a tématu, jsou uvedeny na následujícím obrázku 2.



Obrázek 2: Skupiny hlavních ISO norem geografické informace – stav 2016 (převzato z [18])

3.3 Terminologie ISO norem geografické informace

Geografická informace se zabývá vzhledy jevů (feature) jako abstrakcemi jevů reálného světa se zvláštním zřetelem k jejich poloze vztažené k zemskému tělesu. Pro definování reálných aplikací užití geografické informace slouží řada ISO norem 19100. Odborná terminologie těchto norem je velice rozsáhlá vzhledem k jejich oboru zájmu, který je pevně provázán především s terminologií oboru informačních technologií (IT). Právě kvůli složitosti a rozsahu terminologické působnosti je důležité nejen vést terminologii systémově, podle stanovených postupů a pravidel, ale i ji v komunikaci a dokumentaci důsledně dodržovat.

Seznam základních odborných termínů ISO norem řady 19100 geografické informace (ISO 19104:2016 Geographic information – Terminology) vznikl z popudu orgánů veřejné moci těch států, které se aktivně zúčastňují práce na vývoji a připomínkování norem v anglickém originále a kromě jiného platí poplatky za možnost úzce spolupracovat, a tím ovlivňovat směr vývoje mezinárodních norem ISO pro geografickou informaci. Na rozdíl od států, které nejsou přihlášeny k „participaci“, tzv. P-členství v příslušném oboru standardizace, spolupracující státy získávají v desetiletém předstihu informace, kterými tématy se bude standardizace zabývat, a mají možnost upozornit na potřebu revize některé z norem. ČR je P-členem v technicko-normalizační komisi ISO/TC 211 Geographic information/Geomatics. Například Slovenská republika toto postavení nemá a využívá zprostředkovaně termínů z překladů směrnic, návodů, technických specifikací velkých evropských projektů s mezinárodní účastí a také z překladů mezinárodních a evropsky harmonizovaných norem pořízených ČAS, tj. Českou republikou. Jiný příklad je Slovinsko, které se P-členem stalo v roce 2019 a odborné termíny dosud využívá zprostředkovaně přes projekt INSPIRE.

Seznam asi 500 odborných termínů z oblasti geografické informace je publikován včetně definic a uvedení jejich zdrojů v Technické specifikaci ISO/TS 19104:2016. Seznam je veden a udržován sekretariátem ISO/TC 211, dvakrát ročně je provedena revize a jsou vyřazeny zastaralé termíny. Nové termíny jsou publikovány v terminologické části každé individuální nové nebo revidované normy.

Ucelený souhrn normalizované terminologie používané v normách ISO řady 19100 poskytuje i její elektronicky vedené úložiště. V něm je navíc pečlivě zaznamenáváno, v které mezinárodní normě je termín zaveden poprvé a v kolika dalších mezinárodních normách je využit v povinné kapitole každé normy: „Terms and definition“. Pokud je odborný termín vázán na matematický vzorec, je v seznamu možné najít odborný výklad nebo matematickou definici.

V této části je nutné upozornit na platnou normu ČSN P ISO/TS 19104 Geografická informace – Terminologie z roku 2010, která obsahuje české překlady základních odborných termínů norem ISO geografické informace.

V češtině byla tato norma ISO 19104:2008 Geographic Information – Terminology vydána překladem jako prozatimní technická specifikace ČSN P ISO/TS 19104:2010. S ohledem na dynamický vývoj teorie a oblastí aplikací geografické informace je sice náročné, ale pro sémantickou interoperabilitu naprosto klíčové zajistit konzistenci a jednotné používaní a chápání odborných termínů. Je tedy logické, že tato technická specifikace a uvedené elektronické úložiště termínů patří k nejvíce používaným zdrojům při řešení terminologických problémů.

Vybrané termíny z ČSN P ISO/TS 19104:2010 Geografická informace – Terminologie použité pro projekt NaSaPO jsou uvedeny v příloze A – Terminologický výkladový slovník projektu NaSaPO.

3.3.1 Jev, abstrakce a koncept

Jak bylo uvedeno výše, *„geografická informace se zabývá vzhledy jevů (feature) jako abstrakcemi jevů reálného světa se zvláštním zřetelem k jejich poloze vztažené k zemskému tělesu“*.

V této části autoři postupují velice opatrně. Cílem tohoto dokumentu není seznámení s teoretickými základy a obecnými definicemi pojmů používaných při modelování informací. Ty lze získat studiem poměrně velkého množství dostupné odborné literatury, např. http://www.jaist.ac.jp/~bjorner/course/c5.pdf. Je však potřeba tyto obecné teorie jevů a konceptů uplatnit pro obor geografické informace, uvést hlavní pojmy, které jsou základem pro modelování geografické informace, a vysvětlit je tak, aby softwarový inženýr jasně, jednoznačně a správně pochopil tyto klíčové pojmy a nezaměňoval je.

Jev (fenomén) můžeme volně chápat jako nějakou fyzickou věc, kterou lidé mohou vnímat nebo kterou umělá technologie založená na přírodních vědách může měřit, a nebo jako lidský artefakt, tj. něco pozorované při vědeckém výzkumu nebo experimentu, co není v reálném světě přirozeně přítomné, ale vyskytuje se v důsledku přípravného nebo vyšetřovacího postupu.

Abstrakce (mentální činnost) je proces, který hraje v modelování klíčovou roli. Jeho podstatou je odvození obecnějšího konceptu z více reálných nebo konkrétních konceptů. Jako příklad lze uvést dálnici, která je zvláštním druhem silniční komunikace. V procesu abstrakce oddělujeme od vnímaného jevu to, co považujeme za podružné, nepodstatné, a zdůrazňujeme to, co považujeme za zásadní, podstatné, charakteristické. Úroveň abstrakce představuje množství detailů zachycených v modelu a nakolik příznačné jsou tyto detaily pro konkrétní implementaci. V abstrakci tak modely mohou sahat od definic základních schémat postupu v modelování, přes na platformě závislé specifikace implementace, až k definicím konceptů (pojmů). Vybrat správnou úroveň abstrakce pro model je obtížným úkolem. Nízká úroveň abstrakce představuje pořízení velkého množství detailů, které vytvoří na implementaci závislý model za cenu jeho menšího obecného využití. Vysoká úroveň abstrakce použitá pro konceptuální modely zajistí na sémantické úrovni interoperabilitu s ostatními doménami, ale může být zase obtížnější vzájemně přiřadit (mapovat) konkrétní pojmy mezi těmito doménami.

Podrobnější vysvětlení lze nalézt v Příloze A normy ČSN EN ISO 19126:2012.

Abstrakce[1] *je definována jako akt vymezení nebo oddělené uvážení společných vlastností nebo charakteristik rozdílných objektů, utvoření pojmu nebo ideje takovým aktem a interpretace jména třídy jako výsledek utvoření pojmu. V tomto smyslu může být abstrakcí jevů reálného světa pouze typ vzhledů jevů; instance vzhledu jevu je instancí třídy objektů, z níž byl vlastní pojem abstrahován.*

Pojem[1] *je abstraktní idea zahrnující podstatné atributy třídy nebo logických druhů nebo naopak vyjádření významu obecného termínu. Typ vzhledů jevů je formou pojmů.*

[ČSN EN ISO 19126:2012]

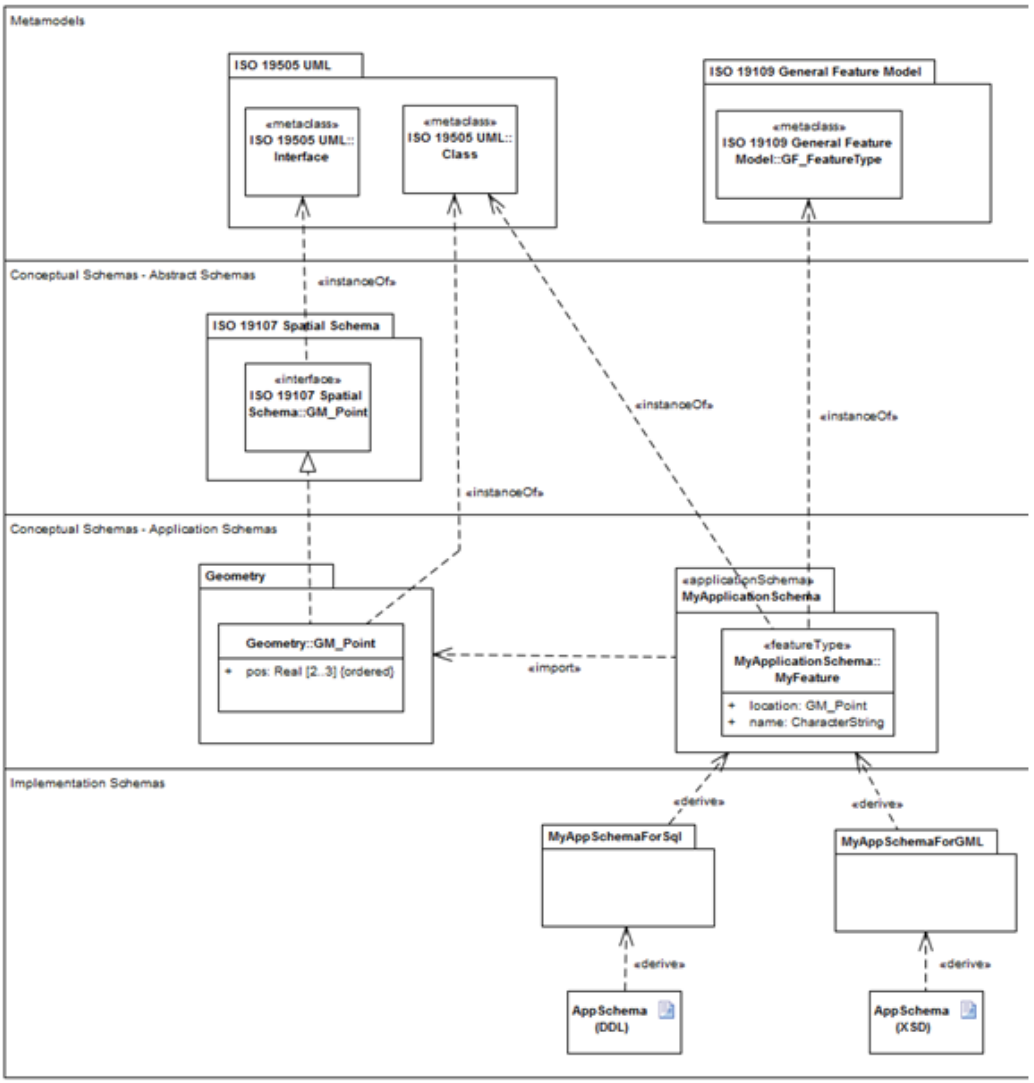
Užitečná diskuse k abstrakci, jejím úrovním a různým aspektům je vedena na platformě: https://github.com/ISO-TC211/UML-Best-Practices/wiki/Level-of-abstraction.

[1] Webster’s Third New International Dictionary, Encyclopaedia Britannica, Inc., 1981

V normě ISO 19103:2015 jsou popsány následující čtyři hlavní úrovně abstrakce:

1. Metamodely (nejabstraktnější metamodely, jako jsou Obecný model vzhledu jevu (GFM) z ISO 19109:2015 a UML Metamodel z ISO/IEC 19505:2012)
2. Konceptuální schéma – Abstraktní schéma (jádrové modely se základními pojmy, tj. geometrie a topologie z IS0 19107:2017)
3. Konceptuální schéma – Aplikační schéma (stále ještě konceptuální modely, ale specifické pro aplikace, tj. Silniční komunikace, Budovy)
4. Implementační schémata (schémata pro konkrétní implementace, tj. GML aplikační schémata (XSD))

Následující obrázek 3 ilustruje různé úrovně abstrakce.



Obrázek 3: Čtyři úrovně abstrakce (převzato z OGC Domain modelling cookbook)

Koncept je pojem, pod kterým volně rozumíme abstrakci (mentální konstrukci) pojatou lidmi, která zachycuje podstatu obvykle třídy jevů. Pojmem „koncept“ máme tedy na mysli abstrakci souboru podobných jevů do jednoho konceptu. Vzhledem k výše uvedenému chápeme popis konkrétního konceptu jako jeho konkretizaci identifikací, která stanoví vztah mezi konkrétním konceptem a alespoň jedním jevem, který je zamýšlen tímto konkrétním konceptem.

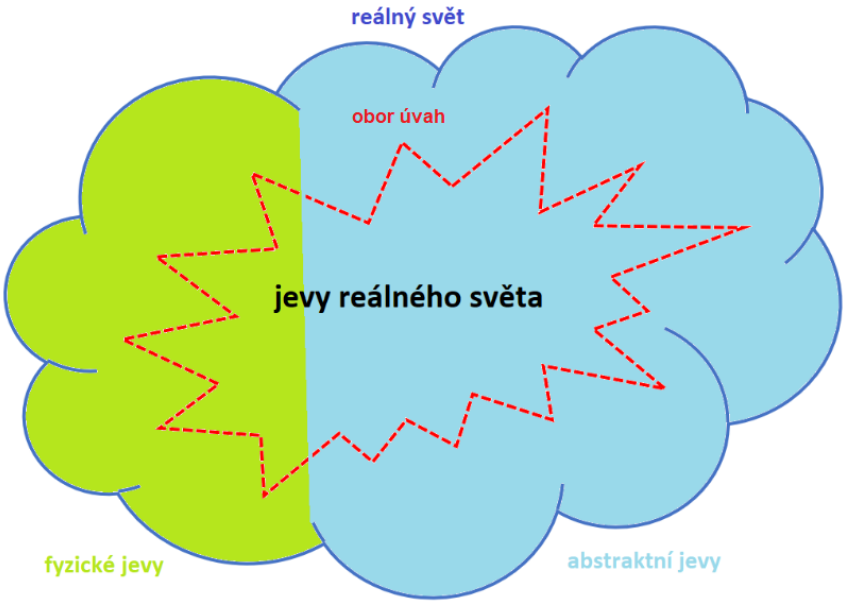
PŘÍKLAD: Pokud „povýšíme“ naše uvažování od konkrétního mostu, tj. od konkrétního jevu, na zástupce množiny mostů, pak jsme „abstrahovali“ od konkrétního jevu do konkrétní koncepce mostu. Když „povýšíme“ naše úvahy ještě dále, například od těch, které se zabývají mosty, tunely nebo budovami, tj. ze souboru konkrétních konceptů, a považujeme jejich instance za stavby, potom považujeme stavby za abstraktní koncept.

V Obecném modelu vzhledu jevu je podle ČSN EN ISO 19109:2020 k označení abstraktního nadtypu vzhledu jevu v metatřídě FeatureType používán povinný atribut isAbstract:

- ➔ *isAbstract – Booleovský atribut (povinný). Pokud má hodnotu true, typ vzhledu jevu působí jako abstraktní nadtyp.*

[ČSN EN ISO 19109:2020]

Pro ucelení přístupu modelování geografické informace je nutné uvést, že jakékoliv měření jevů reálného světa není exaktní, protože všechna realizovaná měření poskytují výsledky s nějakou přesností. Nikdy nelze měřit jev reálného světa absolutně přesně. S reálným světem tedy nejsme schopni pracovat exaktně, narozdíl od abstraktních konceptů, které jsou exaktní, jsou-li stanoveny s nulovou vnitřní vágností jejich významu. Je tedy nezbytné přesně stanovit vazby významu konceptu s lidskou psychikou, resp. ontologii, která zajistí, že každý odborník využívající koncepty bez pochyb ví, co znamenají.



Obrázek 4: Jevy reálného světa

- **jev** – jakýkoliv úkaz, který lze pozorovat, nebo vyvodit ze zobrazení operace nad daty.

Jevy reálného světa lze v principu rozdělit na fyzické a abstraktní (mentální).

  - **fyzický jev (physical phenomenon)** – předmět, událost nebo situace, která existuje nebo se děje a kterou lze vnímat lidskými smysly nebo měřit. Příkladem fyzického jevu je např. budova výzkumného ústavu, vedoucí řešitelského týmu, výzkumná zpráva apod.

    POZNÁMKA: Termín je synonymem k termínu fyzický objekt (viz příloha A).

  - **abstraktní jev (abstract phenomenon)** – pojem, který nereprezentuje fyzickou existenci, existuje jako myšlenka, která se nevztahuje k fyzickému objektu, reálné události nebo činnosti. Jako příklad lze uvést resort ČÚZK, etnickou skupinu, územní statistickou jednotku atd.

- **abstrakce** – proces potlačení nevýznamných podrobností pro vytvoření zjednodušeného modelu (konceptuální model) nebo výsledek takového procesu.
- **pojem/koncept** – jednotka poznání vytvořená jedinečnou kombinací charakteristik.
- **obor úvah/univerzum diskurzu** – náhled na reálný nebo hypotetický svět, který zahrnuje vše, co je předmětem zájmu.

V konceptuální fázi řešení projektu je nutné věnovat zvláštní pozornost především termínu **objekt**, který je obsažen i v samotném názvu předmětu řešení NaSaPO. Je to obecný termín s velmi vysokou četností využívání v komunikaci napříč technickými obory a v lidské činnosti, reprezentující cokoliv hmotného či nehmotného (např. osoba, věc, myšlenka, představa atd.), nač je zaměřeno vnímání, myšlení či jednání. Jeho přesný význam, ostatně jako každého polysémantického termínu, je nutno chápat v kontextu a v souvislostech s prostředím, kde jsou použity. Pokud by mohla nastat nejednoznačnost v chápání významu termínu, bude uvedeno, ve které z fází (architektonické vrstvě) modelování je tento vícevýznamový termín použit. Termínu objekt a jeho možnému různému významu se podrobněji věnují i přílohy A a E.

3.3.2 Základní terminologie v konceptuálním modelu

Úvodem této části je nezbytné zdůraznit, že modelování geografické informace podle norem ISO musí být založeno na objektově orientovaných přístupech. V procesu konceptuálního modelování jsou používány tedy i odpovídající termíny a postupy a tyto jsou zakotveny ve skupině relevantních ISO norem geografické informace. Klíčovým a hodně diskutovaným je termín „**feature**“, který je v rámci ČSN překládán jako „vzhled jevu“. Pregnantně je volba tohoto českého ekvivalentu velmi frekventovaného anglického termínu „feature“ zdůvodněna Ing. Janem Neumannem, CSc. [20], který se při jeho zpracování opíral především o rozbor předního teoretika v předmětné oblasti Prof. RNDr. PhDr. Zdeňka Neubauera, CSc. [22]

Vzhledem k výše uvedenému a tomu, že při rozpadu jevu získáme opět jev (ať už fyzický nebo mentální), lze pro lepší pochopení významu termínu „feature“ níže uvedenou definici interpretovat takto: **Vzhled jevu (feature) je abstrakce jevu reálného světa nebo jeho významné části ve zvoleném konceptu.**

- **vzhled jevu (feature)** – abstrakce jevu reálného světa
  POZNÁMKA: Vzhled jevu se může vyskytovat jako typ nebo jako instance (tj. data sebraná k reprezentaci vzhledu jevu ve shodě se specifikovaným typem vzhledu jevu). Proto výraz „typ

vzhledu jevu“ nebo „instance vzhledu jevu“ by se měl použít tehdy, když je míněna pouze jedna z těchto alternativ.

- **objekt** – entita s jednoznačně definovanou hranicí[2] a identitou, která zapouzdřuje stav a chování
- **třída** – popis množiny objektů, které sdílejí stejné atributy, operace, metody, vztahy a sémantiku
- **prostorový objekt** – objekt používaný k reprezentování prostorových charakteristik vzhledu jevu

3.3.2.1 Vztah k terminologii specifikace UML

Na vyvíjené aplikaci nebo více aplikacích závislý, verbálně, volnou formou popsaný konceptuální model oboru úvah je nezbytné vyjádřit i v přehledném, jednoznačném a srozumitelném tvaru, nazývaném konceptuální schéma. K tomuto strukturovanému popisu je používán jazyk konceptuálního schématu (CSL) a v doméně geografických informací byl jako standardní vybrán UML – Jednotný modelovací jazyk (UML – Unified Modeling Language). Jeho jádrem jsou následující dva klíčové pojmy „objekt“ a „třída“.
K objasnění jiných významů těchto dvou termínů v jiné úrovni modelování (tj. při tvorbě aplikačního schématu) jsou uvedeny jejich definice v jazyce UML:

- **objekt (object)** – *objekt je instancí, která pochází ze třídy. V metamodelu je objekt podtřídou instance a pochází z alespoň jedné třídy. Množina tříd může být dynamicky modifikována, což znamená, že množina rysů objektu se může během své životnosti změnit. [11]*
- **třída (class)** – *popis množiny objektů, které sdílejí stejné atributy, operace, metody, vztahy a sémantiku. Třída může použít množinu rozhraní k určení kolekcí operací, které poskytuje svému prostředí. V metamodelu třída popisuje množinu objektů sdílejících kolekci vlastností, včetně operací, atributů a metod, které jsou společné pro množinu objektů. [11]*

3.3.2.2 Vztah k terminologii v projektech INSPIRE

Při modelování podle Směrnice INSPIRE jsou také z reálného světa vybírány množiny jevů a je zvolen objektově-orientovaný přístup. I když se řešení projektu INSPIRE rovněž opírá o ISO normy geografické informace, ne vždy byla dodržena vzájemná koherence termínů, což v praxi pochopitelně přináší řadu problémů. Případné odchylky od důsledného používání jednotící terminologie ISO norem nelze zakázat. V každém případě však vlastní terminologie a používání stejných termínů s jiným významem komplikují komunikaci uvnitř řešitelského týmu a zejména brání správnému pochopení vypracovaných řešení a sdílení dat.

Obecné termíny a jejich definice, které jsou určeny k jednotnému používání v dokumentech Směrnice INSPIRE a Prováděcí pravidla INSPIRE, jsou shrnuty a publikovány ve společném mnohojazyčném terminologickém slovníku INSPIRE Glossary. Je vlastněn Evropskou komisí a udržován Společným výzkumným střediskem jako jeden z registrů INSPIRE. Je dostupný z http://inspire.ec.europa.eu/glossary.

Následující výpis několika hlavních termínů pro modelování geografické informace a jejich definic z registru INSPIRE Glossary demonstruje odlišnosti od terminologie norem ISO řady 19100. Je z něj zřejmá určitá

[2] **hranice** – množina, která reprezentuje mez entity

obtížnost orientace v termínech projektu INSPIRE, které byly vytvořeny nebo převzaty z ISO norem, ale byl jim v některých případech přiřazen jiný význam.

Vybrané hlavní termíny a definice používané v projektu INSPIRE:

- **třída (class)** – popis množiny objektů sdílejících stejné vlastnosti, omezení a sémantiky [UML 2.1.2 – upraveno]
- **entita (entity)** – jev reálného světa
- **vzhled jevu (feature)** – abstrakce jevů reálného světa [ISO 19101-1:2014]
  POZNÁMKA: Termín „(geografický) vzhled jevu“ (geographic) feature, jak je používán v ISO normách řady 19100, v dalších specifikacích, jako je IHO S-57, a v dokumentu Glossary, je synonymem k termínu „prostorový objekt“ (spatial object), jak je používán v dokumentu Glossary. Termín „prostorový objekt“ (spatial object) je používán také v mezinárodních normách ISO řady 19100. Bohužel však s odlišným významem, což logicky přináší řadu problémů. Prostorový objekt v normách ISO řady 19100 je termínem používaným k reprezentování prostorových charakteristik vzhledů jevů – geometrie nebo topologie (viz definice v ISO/FDIS 19107).
- **objekt (object)** – v dokumentu Glossary je to termín použitý jako synonymum s termínem prostorový objekt
- **prostorový objekt (spatial object)** – abstraktní reprezentace jevu reálného světa vztaženého ke konkrétnímu místu nebo geografické oblasti [Směrnice INSPIRE]
  POZNÁMKA: Mělo by být vzato na vědomí, že termín má v normách ISO řady 19100 jiný význam než v dokumentech INSPIRE. Je také synonymem s termínem „(geografický) vzhled jevu“, jak je používán v normách ISO řady 19100.
- **prostorová věc (spatial thing)** – cokoliv s prostorovým rozsahem, tj. velikostí, tvarem nebo polohou [WGS84 Geo Positioning: an RDF vocabulary]

PŘÍKLAD: lidé, náměstí, koule na bowling, stejně tak i abstraktní areály, jako jsou statistické územní jednotky apod.

POZNÁMKA: Prostorové věci jsou ty podmnožiny jevů reálného světa, které se vztahují k místu.

3.3.3 Obor úvah ISO norem geografické informace

Podrobné vysvětlení zásadního termínu obor úvah, spolu s dalšími pojmy, které vytvářejí základ pro abstrakci jevů reálného světa v ISO normách pro modelování geografické informace, lze nalézt v příloze E normy ČSN EN ISO 19101-1:2018 Geografická informace – Referenční model – Část 1: Základy, odkud byla do terminologické normy ČSN P ISO/TS 19104:2009 převzata i jeho následující krátká, ale velmi výstižná a srozumitelná definice:

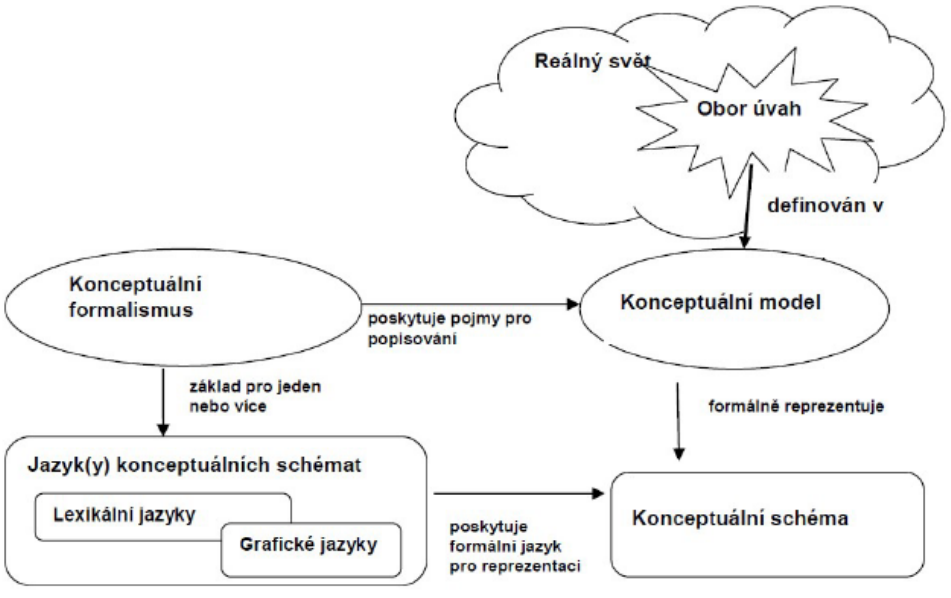
- **obor úvah, univerzum diskurzu (universe of discourse)** – náhled na reálný nebo hypotetický svět, který zahrnuje vše, co je předmětem zájmu. [ISO 19101-1:2014]

Modelování reálného světa se zabývá vždy jen určitou podmnožinou ze všech jeho jevů (fyzických i abstraktních) a informací o nich, které jsou relevantní pro zamýšlený model. Tato podmnožina tvoří obor úvah a nazývá se také univerzum diskurzu. Koncept tohoto modelu je ovlivněn aplikací nebo několika

aplikacemi, pro něž je model vytvářen. Jako příklad může posloužit jev reálného světa „ulice“. Na ulici ve městě může být pro aplikace katastru nemovitostí nahlíženo jako na parcelu, která má vlastníka a výměru a je reprezentována plochou vymezenou hranicemi parcely. Pro řešení dopravních aplikací však je tatáž ulice částí komunikační sítě spojující dvě místa a je reprezentována středovou linií, koncovými uzly a je charakterizována potřebnými dopravními a provozními parametry. Je tedy zřejmé, že každá z aplikací potřebuje vytvořit jiný konceptuální model a následně datovou strukturu, jinou vlastní geometrii a sémantiku. Pro minimalizaci nákladů spojených se sběrem informací je však vždy hledáno takové řešení – **konceptuální model** –, který bude pro obě aplikace kompromisem a bude vyhovovat těmto plánovaným nebo i dalším předvídatelným aplikacím. Z reálného světa je tedy procesem objektové generalizace, která některé vybrané informace zvýrazní a jiné potlačí, vytvořen prvotní model (základní objektový model), který je aproximací reality závisející na očekávaných aplikacích.

- **konceptuální model (conceptual model)** – model, který definuje pojmy univerza diskurzu [ISO 19101-1:2014]
- **konceptuální schéma (conceptual schema)** – formální popis konceptuálního modelu [ISO 19101-1:2014]
- **aplikační schéma (application schema)** – konceptuální schéma pro data požadovaná jednou nebo více aplikacemi [ISO 19101-1:2014]

Obrázek 5 znázorňuje vztah mezi modelovanou realitou, přesněji mezi její vybranou částí tzv. **univerzem diskurzu**, která má být popsána v modelu a o níž mají být vlastně pořízena digitální data, a požadovaným konceptuálním schématem. Tento popis se nazývá **konceptuální model** a zahrnuje nejen vybrané vzhledy jevů, ale i jejich atributy, operace a vztahy, které jsou z hlediska zamýšlené aplikace relevantní a existují mezi vzhledy jevů.



Obrázek 5: Z reality do konceptuálního schématu (převzato z ČSN EN ISO 19101-1:2017)

Konceptuální schéma je přesný popis tohoto konceptuálního modelu, které klasifikuje vzhledy jevů do typů a tříd tím, že identifikuje jejich vlastnosti (struktury a chování) a asociace mezi typy vzhledů jevů. Konceptuální schéma rovněž definuje vztahy mezi jednotlivými prvky konceptuálního modelu a omezení u těchto prvků a jejich vztahů. Tato omezení definují platné stavy modelovaného systému.

Konceptuální schéma, které definuje určitou část univerza diskurzu pro konkrétní aplikaci nebo více aplikací, se nazývá **aplikační schéma**. Konkrétní část modelovaného univerza diskurzu je reprezentována **digitálními daty**, která jsou strukturována podle specifikace v konceptuálním schématu.

Ke zpracování konceptuálních schémat jsou používány slovníky definic vzhledů jevů (Feature concept dictionaries) vedené jako registry a zpracované v souladu s ČSN EN ISO 19126:2012, pro aplikační schémata jsou využívány nebo zpracovány konkrétní katalogy vzhledů jevů – ČSN EN ISO 19110:2017. Blíže je tato problematika rozepsána v podkapitole 4.2.

Konceptuální model je zpracováván v libovolné formě, nezřídka běžným jazykem. K přesnému, úplnému a jednoznačnému vyjádření konceptuálního modelu a vytvoření konceptuálního schématu se však používá **Jazyk konceptuálního schématu** (CSL). Je to formální jazyk, který může být snadno a srozumitelně pochopen a analyzován člověkem i počítačem a obsahuje všechny jazykové konstrukty k přesnému vyjádření a manipulaci s obsahem konceptuálního schématu. CSL je založen na konceptuálním formalismu, který stanoví pravidla, omezení, mechanismy dědičnosti, události, funkce, procesy a další prvky tvořící jazyk konceptuálního schématu.

Z mnoha existujících CSL byl pro podporu norem ISO geografické informace řady 19100 vybrán **univerzální modelovací jazyk** (UML) doplněný **jazykem omezení objektů** (OCL). Používá se ve formě profilu tohoto UML popsaného normou ISO/IEC 19501:2005 uzpůsobeného potřebám specifickým pro oblast geografické informace. Pravidla a směrnice pro použití takto modifikovaného UML k vytváření normalizovaných modelů geografických informací a služeb jsou stanoveny v technické specifikaci ČSN P ISO/TS 19103:2005.

- **konceptuální formalismus (conceptual formalism)** – sada modelovacích pojmů použitých k popisu konceptuálního modelu [ČSN EN ISO 19101-1:2018]

Konceptuální formalismus poskytuje obecný základ pro formální definování veškerých znalostí relevantních pro aplikaci informační technologie. Platí to v plném rozsahu i pro aplikace geografických informací a tvorbu konceptuálních schémat pro části ISO norem geografické informace. Příkladem konceptuálního formalismu je UML meta model. Jeden konceptuální formalismus může být vyjádřen v několika jazycích konceptuálního schématu. K jednomu stejnému použitému konceptuálnímu formalismu může náležet i více než jeden, buď lexikální, nebo grafický CLS.

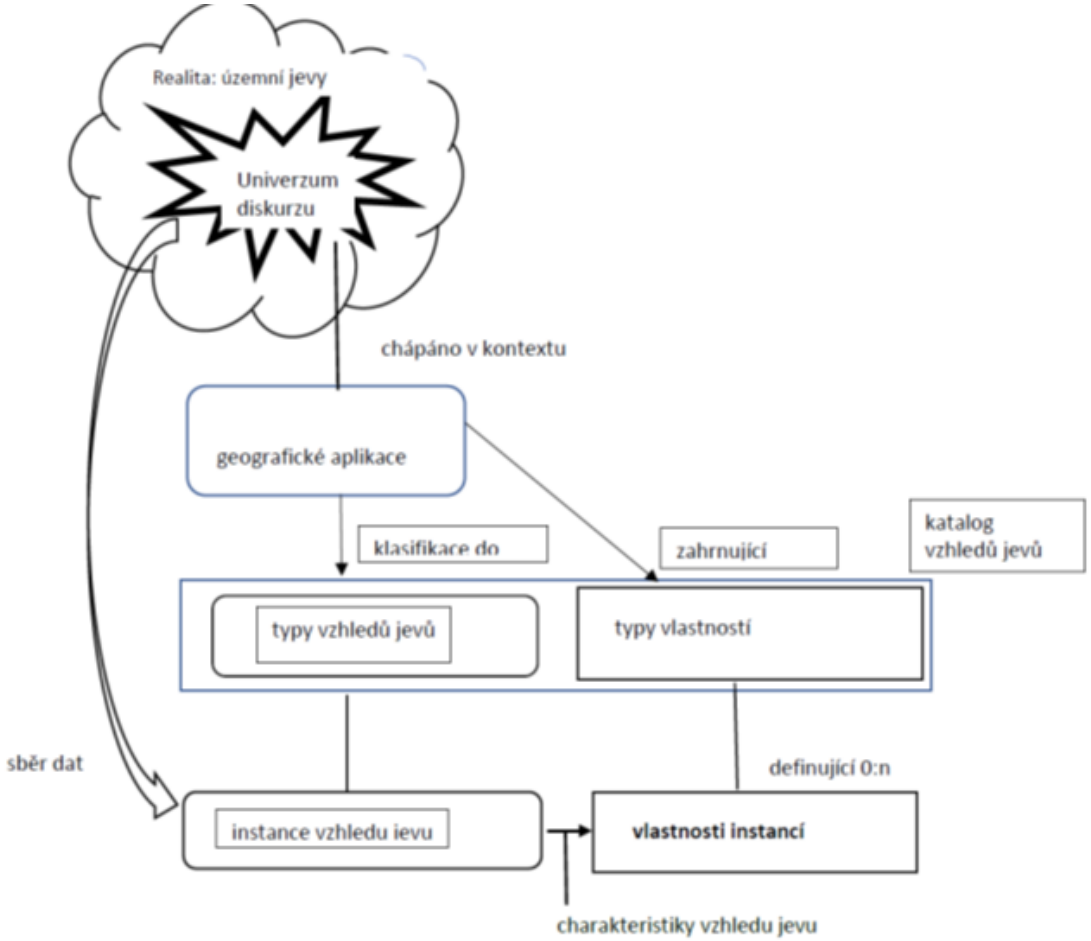
- **Grafický jazyk** – jsou použity UML diagramy tříd a balíčků s OCL konkrétně definované v ISO 19103 umožňující syntaktickou a sémantickou interoperabilitu.
- **Lexikální jazyk** – je použit OWL k podpoře interoperability napříč obory prostřednictvím Sémantického webu. ISO 19150-2 definuje pravidla pro konverzi diagramu UML aplikačních schémat do OWL.

Aplikační schémata, jako koncepční schémata pro jednu nebo několik konkrétních aplikací, jsou tedy také vytvářena podle pravidel a směrnic stanovených v technické specifikaci ČSN P ISO/TS 19103 a v souladu s normou ČSN EN ISO 19109:2020 Geografická informace – Pravidla pro tvorbu aplikačních schémat, která určuje jako závazný CSL unifikovaný modelovací jazyk (UML) doplněný jazykem omezení objektů (OCL).

Tato velmi aktuální a před nedávnou dobou jako ČSN EN ISO 19109:2020 publikovaná norma specifikuje obecná pravidla pro tvorbu formalizovaného aplikačního schématu nezávisle na konkrétním CLS a určuje základní principy pro definování vzhledů jevů. Tato norma rozlišuje čtyři aspekty pro definování tohoto pojmu:

- definice nebo popisy použité pro jejich seskupení do typů
- atributy asociované s každým typem
- vztahy mezi typy
- chování vzhledů jevů.

Jednou ze stěžejních částí těchto obecných pravidel je definování Obecného modelu vzhledu jevu. Nejabstraktnější úroveň definování a strukturování geografických dat je schematicky vyjádřena na následujícím obrázku 6.



Obrázek 6: Proces od univerza diskurzu k datům (převzato z ČSN EN ISO 19109:2020)

Klasifikace jevů reálného světa jako vzhledů jevu závisí na jejich významu pro konkrétní vybrané univerzum diskurzu a požadovanou geografickou aplikaci.

Normativní rámec pro organizování a hlášení klasifikace vzhledů jevů poskytuje
ČSN EN ISO 19 110:2017 – Geografická informace – Metodologie katalogizace vzhledů.

Vzhled jevu je jako fundamentální jednotka geografické informace podrobněji a v různých kontextech tohoto pojmu popsán v čl. 3.3.3.3 Čtyřvrstvá architektura ISO norem geografické informace.

3.3.3.1 Referenční model

K zajištění integrity geografické informace i výsledků její aplikace se musí vývoj a využití opírat o stanovený referenční model. Referenční model popisuje základní požadavky, kterými je nezbytné se řídit. Popis referenčního modelu se opírá o konceptuální formalismus.

Referenční model je abstraktním rámcem pro konkrétní doménu, sestávající z vzájemně provázané množiny zřetelně definovaných pojmů vytvořených experty nebo skupinami expertů s cílem:

- vytvořit standardy pro objekty a jejich vzájemné vztahy,
- poučit a napomáhat při řešení složitých komplexních problémů jejich rozdělením na více menších problémů, které jsou snadněji uchopitelné, lze je lépe projednat a vylepšovat řešení,
- podpořit srozumitelnou komunikaci mezi lidmi,
- stanovit jednoznačné role a definovat zodpovědnosti,
- umožnit porovnání možných alternativ řešení problémů.

Referenční modely jsou často ilustrovány jako množiny konceptů s nějakou indikací vztahů mezi nimi.

Podle OASIS (Organization for the Advancement of Structured Information Standards) je referenční model abstraktním rámcem pro pochopení důležitých vztahů mezi entitami nějakého prostředí a pro vývoj konzistentních standardů nebo specifikací toto prostředí podporujících. Referenční model je založen na nevelkém počtu sjednocujících konceptů a může být použit jako základ pro vzdělávání a vysvětlování těchto standardů neodborníkům. Referenční model není přímo svázán s žádnými standardy, technologiemi nebo jinými konkrétními detaily implementací. Usiluje však o poskytování jednotné sémantiky, která může být použita jednoznačně napříč různými implementacemi i mezi nimi navzájem.

Pojetí OASIS bylo převzato jako definice referenčního modelu v normách ISO geografické informace:

- **referenční model (reference model)** – rámec pro pochopení významných vztahů mezi entitami určitého prostředí a pro vývoj konzistentních norem nebo specifikací podporujících toto prostředí [ČSN EN ISO 19101-1:2018]

POZNÁMKA: Referenční model je založen na malém počtu sjednocujících pojmů a může být použit jako základ pro vzdělávání a vysvětlování norem laikům.

Referenční model Otevřeného distribuovaného zpracování (RM-ODP)

RM-ODP (*Reference Model of Open Distributed Processing*) je referenčním modelem v oblasti IT. Vznikl jako výsledek společného úsilí ISO, IEC a ITU-T a poskytuje koordinační rámec pro standardizaci Otevřeného distribuovaného zpracování. Je založený na přesných pojmech pocházejících ze současného vývoje distribuovaného zpracování a, pokud je to možné, i na využívání metod formálního popisu pro specifikaci architektury. RM-ODP napomáhá distribuci, interagování, nezávislosti na platformách a technologiích a vzájemné přenositelnosti.

RM-ODP charakterizují čtyři základní prvky:

- objektově modelový přístup při specifikování systémů;
- specifikace systémů z hlediska oddělených, avšak vzájemně souvisejících pohledů;
- transparentnosti distribucí pro systémové aplikace jako podmínka definování infrastruktury systémů;
- rámec pro vyhodnocení systémové shody.

RM-ODP sestává ze čtyř základních doporučení ITU-T a mezinárodních norem ISO/IEC, které definují systém navzájem souvisejících nezbytných konceptů potřebných ke specifikování systémů otevřeného distribuovaného zpracování a které poskytují dobře propracovaný rámec podnikové architektury pro strukturování specifikací rozsáhlých systémů.

Modelování pohledů a rámec RM-ODP

Specifikace velmi složitých systémů, kterým je i NaSaPO, bývají natolik rozsáhlé, že jen jeden člověk nemůže plně porozumět všem jejich aspektům a potřebám, analyzovat je nebo dokonce navrhnout a vyřešit efektivní systém. K vypracování komplexních specifikací pro otevřené distribuované zpracování těchto složitých systémů byl vyvinut rámec pohledů, který představují metodické rozčlenění specifikací celého systému do oddělených, relativně samostatných částí, z nichž každá zodpoví otázky specialistů pro určité konkrétní aspekty specifikace. S každým z těchto pohledů je asociován jeho vlastní jazyk a každý pohled používá svůj optimalizovaný slovník pojmů a svůj způsob a formy prezentace.

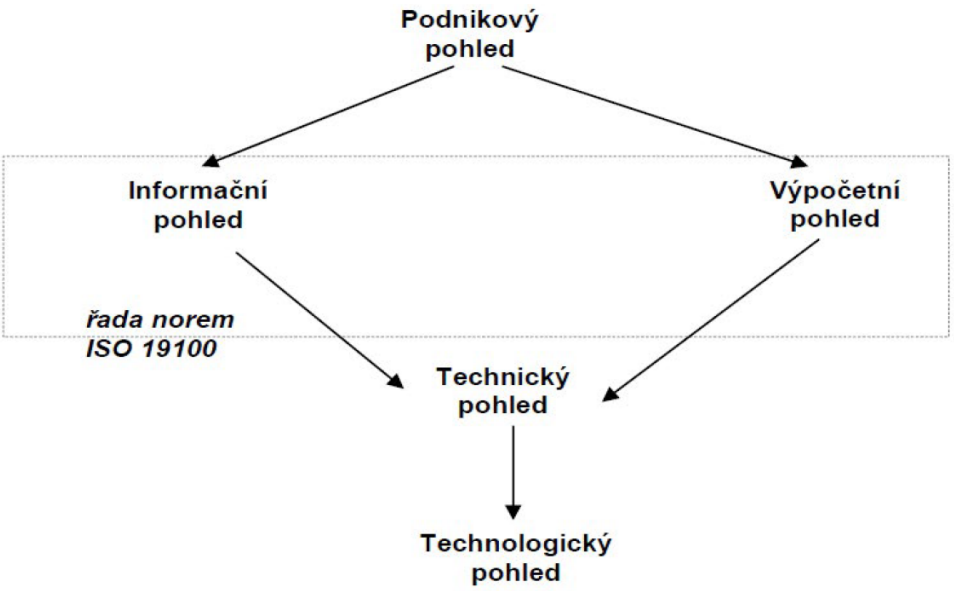
Ačkoliv jsou jednotlivé pohledy specifikovány odděleně, nejsou zcela nezávislé. Klíčové položky v každém z nich jsou identifikovány jako související s položkami v jiných pohledech a všechny pohledy používají v podstatě stejné základní pojmy. Na druhé straně však jsou pohledy dostatečně nezávislé, aby zjednodušily zdůvodňování celkové specifikace. Vzájemná konzistence mezi pohledy je zajištěna architekturou definovanou RM-ODP a navzájem všechny pohledy stmeluje dohromady používání společného objektového modelu.

RM-ODP používá na systém a jeho prostředí následujících pět obecných a vzájemně se doplňujících pohledů:

- **Podnikový pohled**, který se zabývá účelem a uplatněním, metodami, odpovědnostmi a podnikovým procesem, tj. definuje roli infrastruktury prostorové informace NIPI v jeho vlastním prostředí. Popisuje podnikové požadavky a jak je splnit.
- **Informační pohled**, který se zaměřuje na sémantiku informací a vlastní zpracování informací. Popisuje informace spravované systémem a strukturu a typ obsahu podpůrných dat.
- **Výpočetní pohled**, který souvisí s funkčním rozkladem NIPI na služby včetně rozhraní a operací. Popisuje funkcionalitu poskytovanou systémem a její funkční dekompozici. K popisu jsou používány

diagramy balíčků UML, diagramy tříd UML a omezení OCL. Tento model je pevně vztažen k modelům domény geografické informace a je implementačně a technologicky neutrální.

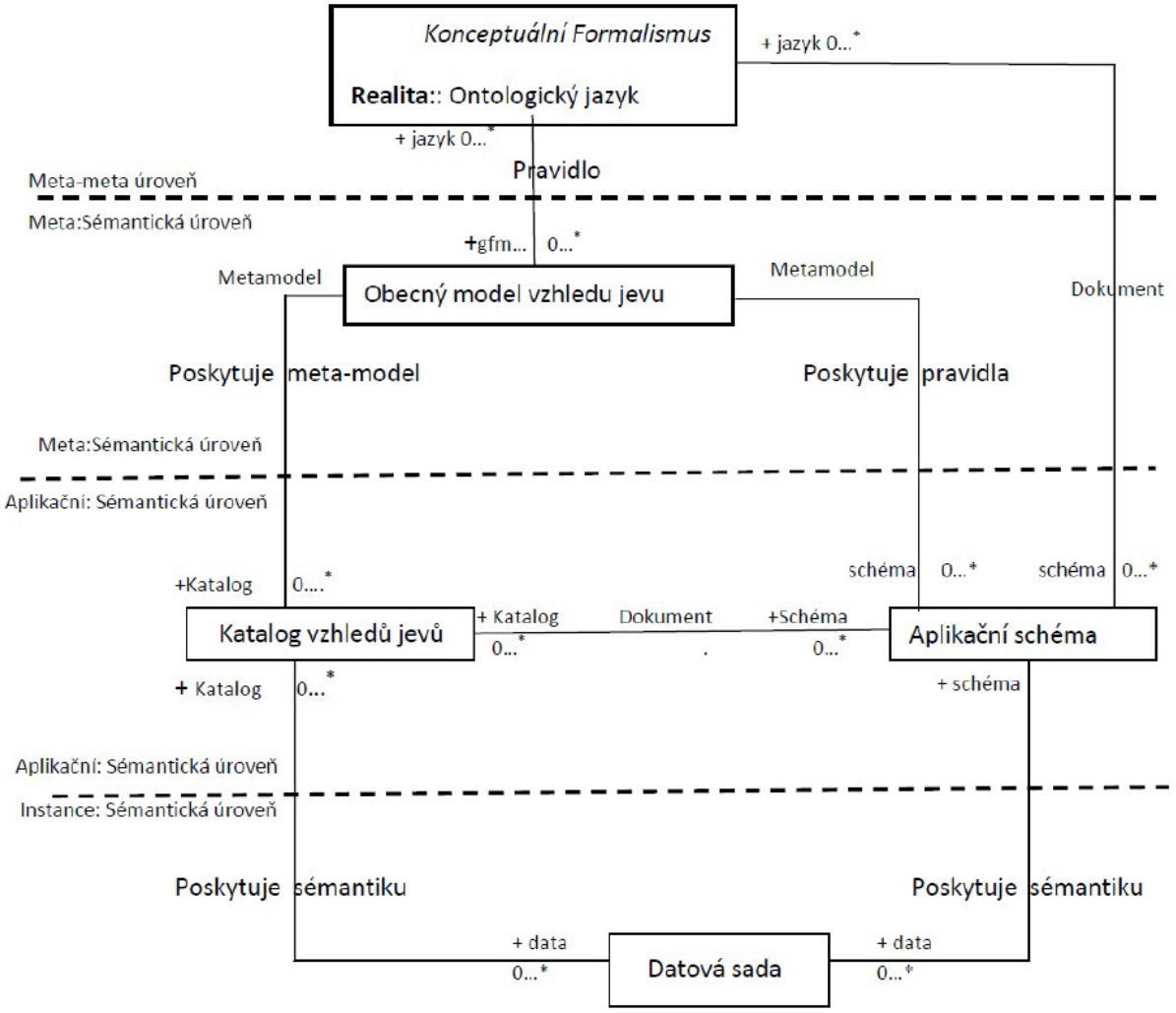
- **Technický pohled**, který se týká interakce mezi daty a službami a vzájemných propojení systémů. Zaměřuje se na mechanismy a funkce potřebné k podpoře distribuovaných interakcí mezi objekty v systému.
- **Technologický pohled**, který se zabývá výběrem technologie systému. Popisuje technologie vybrané k implementaci NIPI, tj. na poskytnutí, zpracování, funkcionalitu a prezentace informací.



Obrázek 7: Pohledy RM-ODP (převzato z ČSN EN ISO 19101-1:2018)

Normy geografické informace se týkají zejména informačního a výpočetního pohledu. Z informačního pohledu přinášejí ISO normy geografické informace pravidla, metodologie a modely od meta-meta úrovně (např. UML), meta úrovně (např. jazyk konceptuálního schématu OCL, Obecný model vzhledu jevu GFM), aplikační úrovně (např. aplikační schéma, katalog vzhledů jevů, profil metadat) po instanční úroveň (např. sady geografických dat a sady metadat pro popis vhledů jevů), viz následující obrázek 8. Podrobnější aktuální přehled norem ISO geografické informace aplikovaných v jednotlivých úrovních modelování ve shodě s referenčním modelem je uveden jako tabulka v příloze C.

Informační pohled na vzájemné vztahy v konceptuálním formalismu je v různých úrovních navrhování informačního modelu NaSaPO podrobně ilustrován na následujícím obrázku 8.



Obrázek 8: Informační pohled RM-ODP (převzato z ČSN EN ISO 19101-1:2018)

3.3.3.2 Interoperabilita

Interoperabilita je obecně definována jako schopnost systému nebo jeho komponent umožnit sdílení informace a řídit proces spolupráce mezi aplikacemi. V normách třídy „Aplikace IT ve vědě (včetně digitální geografické informace)“ je poněkud širší v tom smyslu, že se jedná o schopnost komunikovat, realizovat programy nebo přenášet data mezi různými funkčními jednotkami způsobem, který vyžaduje od uživatele malé nebo žádné znalosti o jednotlivých charakteristikách těchto jednotek. Další vlastní definice interoperability geoprostorových dat jsou uvedeny např. v Terminologickém slovníku zeměměřictví a katastru nemovitostí nebo ve Strategii rozvoje infrastruktury pro prostorové informace v ČR do roku 2020, jež byla schválena usnesením vlády č. 815 ze dne 8. 10. 2014 a v jejím Akčním plánu. Podstata problému interoperability však zůstává beze změny.

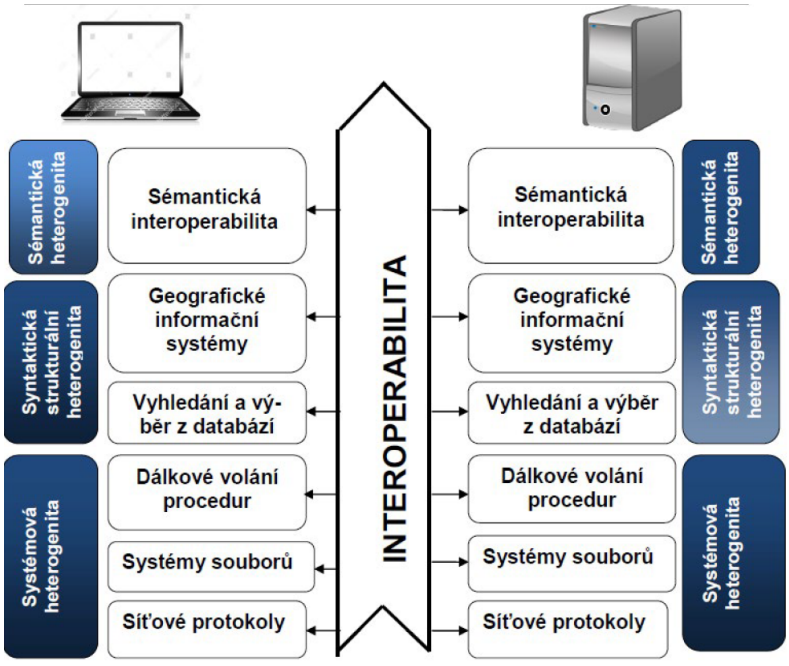
Interoperabilita je široce definována v různých kontextech. Základní definice pochází z normy ISO/IEC 2382-1:1993:

- *Způsobilost dvou nebo více systémů nebo komponent vyměňovat si informace a takto vyměněnou informaci používat.*
- *Schopnost komunikovat, provádět programy nebo přenášet data mezi různými funkčními jednotkami tak, že uživatel vystačí jen s malou znalostí konkrétních vlastností těchto jednotek, popř. je vůbec nepotřebuje znát.*

Interoperabilita je tedy proces, s jehož pomocí nezávislé systémy manipulují, vyměňují a integrují data, jež jsou automatizovaně získána z jiných systémů. Interoperabilitu geografické informace lze popsat jako proces komunikace (interakce) mezi uživatelem a poskytovatelem geografické informace.

Geografická informace je velmi významnou součástí dat e-Governmentu. Propojení geografické informace s jinými druhy informace interoperabilním způsobem má klíčový význam a je rozhodujícím předpokladem pro její výměnu a manipulaci s ní napříč distribuovanými systémy i v mezinárodním rozsahu. Pro evropský kontext je definován soubor doporučení, návodů, metod, zásad a norem pro data a služby e-Governmentu, které vytváří Evropský rámec interoperability (EIF) pro vzájemné interakce mezi organizacemi.

Protože systémy a aplikace uživatele a poskytovatele mohou být odlišné, musí být pro dosažení interoperability překonána heterogenita. Jedním z nejčastěji používaných způsobů popisu interoperability je její rozvrstvení do šesti úrovní. V této dekompozici, která je znázorněna na následujícím obrázku 9, je každá vrstva závislá na stavu vrstvy předchozí. To tedy znamená, že pokud existuje problém s interoperabilitou v některé z nich, má to negativní dopad i na všechny vrstvy nad ní.



Obrázek 9: Dekompozice interoperability geografické informace (převzato z ČSN EN ISO 19101-1:2018)

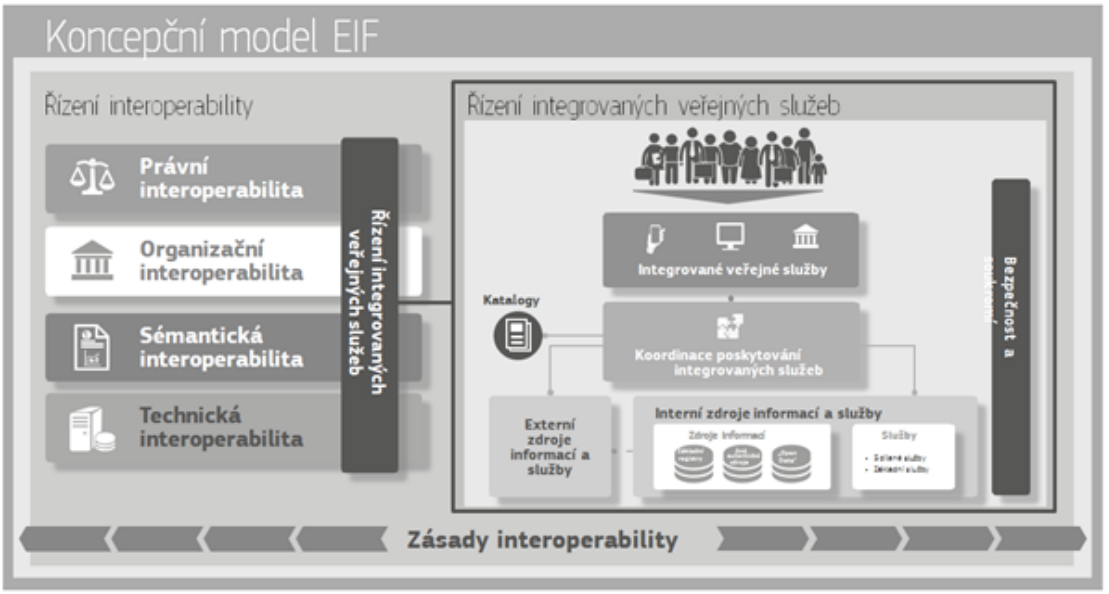
Heterogenita geografické informace

Heterogenitu v geografické informaci rozděluje norma ČSN EN ISO 19101-1 do čtyř typů:

- **Systémová heterogenita** – rozdíly v technickém vybavení, mezi operačními systémy a komunikačními systémy (sítě a protokoly). Nastavení propojení systémů je nezbytným krokem projektu NaSaPO. Mezinárodní normy ISO geografické informace se tímto ustanovením nezabývají.
- **Syntaktická heterogenita** – rozdíly ve fyzické reprezentaci dat, vztahující se k vnější podobě zprávy odesílatele k příjemci, nikoliv k obsahu zprávy. Syntax definuje symboly a gramatiku pro přenos dat ve zprávě. Syntaktická heterogenita se projevuje při komunikaci v odlišném kódovacím formátu (odesílání Shapefile/příjem DXF) nebo v odlišných strukturách dat aplikací (vektor/rastr).
- **Strukturální heterogenita** (schématická heterogenita) – rozdíly různých konceptuálních modelů geografických vzhledů jevů, které jsou popsány rozdílným pojetím tříd, atributů a vztahů.
- **Sémantická heterogenita** – rozdíly ve významu mezi pojmy a daty reprezentujícími realitu. Vzniká z důvodu různých kontextů (hledisek), které byly použity pro abstrakce jevů reálného světa. Například most mohou vnímat různí jedinci odlišně (jako silniční infrastrukturu, železniční infrastrukturu nebo jako překážku pro lodní dopravu).

Řízení interoperability se týká rozhodování o rámcích interoperability, institucionálních opatřeních, organizačních strukturách, rolích a odpovědnosti, politikách, dohodách a jiných aspektech zajišťování a sledování interoperability na vnitrostátní úrovni a na úrovni EU.

Významnou součástí řízení interoperability na úrovni EU je Evropský rámec interoperability, akční plán interoperability a architektura evropské interoperability (EIRA). Pro veřejné služby v EU jsou základní zásady, úrovně a koncepční model interoperabilit publikovány v unijním dokumentu EIF Strategie provádění [21].



Obrázek 10: Koncepční model EIF a řízení integrovaných veřejných služeb (převzato z EIF Strategie provádění)

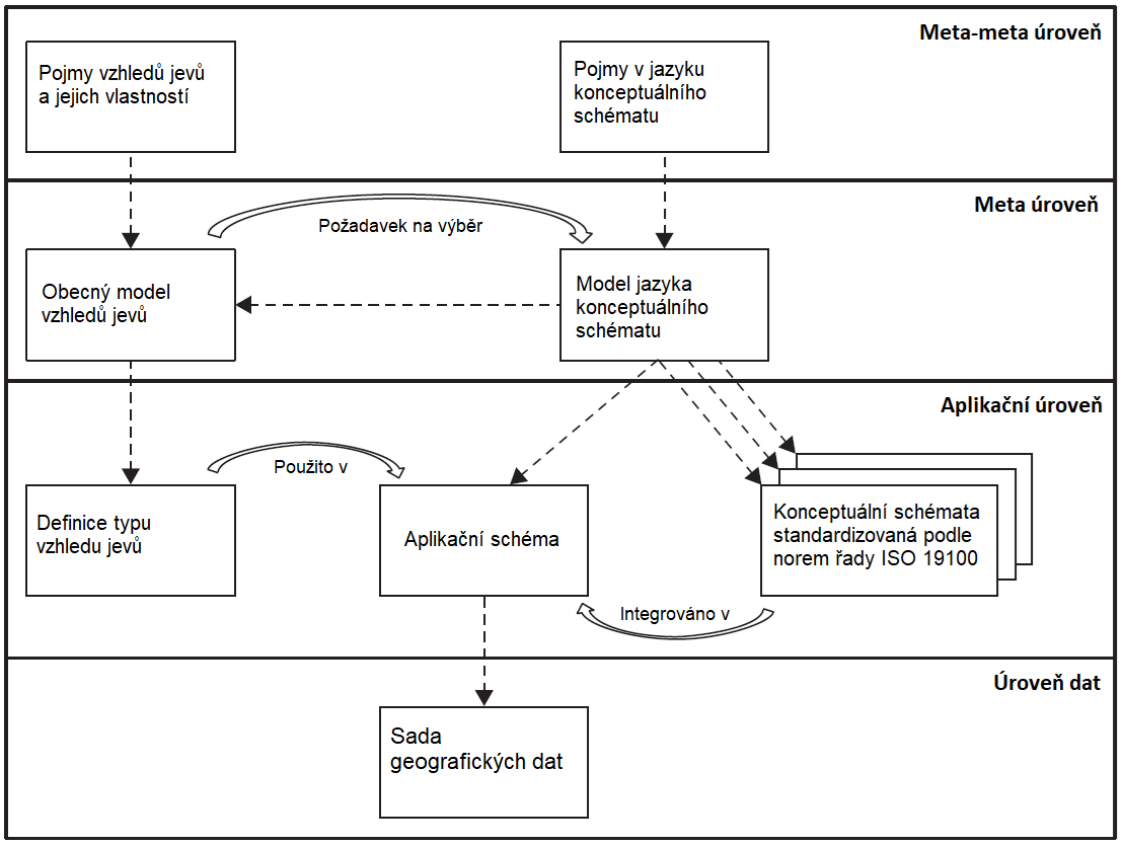
Příkladem rámce interoperability včetně právní interoperability, koordinačních struktur a technických podmínek interoperability pro specifickou oblast životního prostředí, ale nejen pro ni, je v evropském kontextu směrnice INSPIRE.

3.3.3.3 Čtyřvrstvá architektura ISO norem geografické informace

V současné době je při projektování složitých ICT systémů nejčastěji aplikován přístup založený na modelem řízené architektuře. Model vždy odpovídá nějakému příznačnému meta modelu. Jedním z těchto typických meta modelů navržených OMG je UML. Rovněž organizace ISO publikovala standardní meta model v normě ISO/IEC 24744 Software Engineering – Metamodel for Development Methodologies.

Také architektura modelování geografické informace podle norem ISO je založena na modelem řízeném přístupu, který definovala OMG, a je rozvrstvena do čtyř hierarchických úrovní. Pro specifikaci modelu je používán meta model, jehož základem je UML model. Přednostmi meta modelového přístupu je jeho intuitivnost a snadné praktické používání při implementaci.

Formální přístup k modelování geografické informace musí být ve shodě s čtyřvrstvou architekturou. Následující obrázek 11 popisuje vztahy mezi obecným modelem vzhledu jevu a aplikačním schématem a vztahy ve čtyřvrstvé architektuře.



Obrázek 11: Čtyřvrstvá architektura (převzato z ČSN EN ISO 19109:2021)

Meta-meta model

Úroveň meta-meta modelu je nejvyšší vrstvou architektury schématu a tvoří základy meta modelové hierarchie. Jejím hlavním účelem je specifikovat pojmy, termíny a formalismus, a tím definovat jazyk pro specifikaci meta modelu. Na této úrovni se nacházejí pojmy definování vzhledů jevů a také pojmy definování jazyků konceptuálního schématu. Tato úroveň obsahuje definující schéma, které specifikuje pojmy, terminologii, operace a předpoklady potřebné pro specifikaci základních konstruktů na meta-meta úrovni. Tato úroveň je obvykle vyjádřena v přirozeném jazyce a samotná není předmětem standardizace.

Meta-meta model však umožňuje definici konceptů a aplikačních schémat ve standardním schématu a podporuje vyhodnocení navrhovaných změn vůči výchozímu standardnímu schématu a tvorbu konstruktů.

- **meta-meta model (meta-metamodel)** – rámce, pravidla, omezení, modely a teorie aplikovatelné a užitečné pro modelování předem vymezené třídy problémů [OMG]
  POZNÁMKA: Vztahy mezi jednotlivými vrstvami architektury schématu (modely) lze popsat následující zkratkou: Model je instancí meta modelu a meta model je instancí meta-meta modelu.

Meta model

Meta modely (meta-modelová úroveň) obsahují definice pojmů, terminologii, operace a předpoklady potřebné ke konstrukci složitějších aplikačních schémat. Tato složitější aplikační schémata nazýváme **konstrukty**. Pracují se syntaxí a sémantikou různých jazyků pro modelování nebo reprezentaci včetně jazyků konceptuálního schématu (UML, EXPRESS), nebo jen schémat nebo vzorů použitých pro elementární modelování. Konstrukty standardních schémat zahrnuté do této úrovně jsou popsány v jazyce založeném na základních konceptech v **referenčním modelu (definujícím schématu)**.

- **meta model (metamodel)** – model, který definuje jazyk pro vyjádření jiných modelů

[UML 2.4.1]

Aplikační model

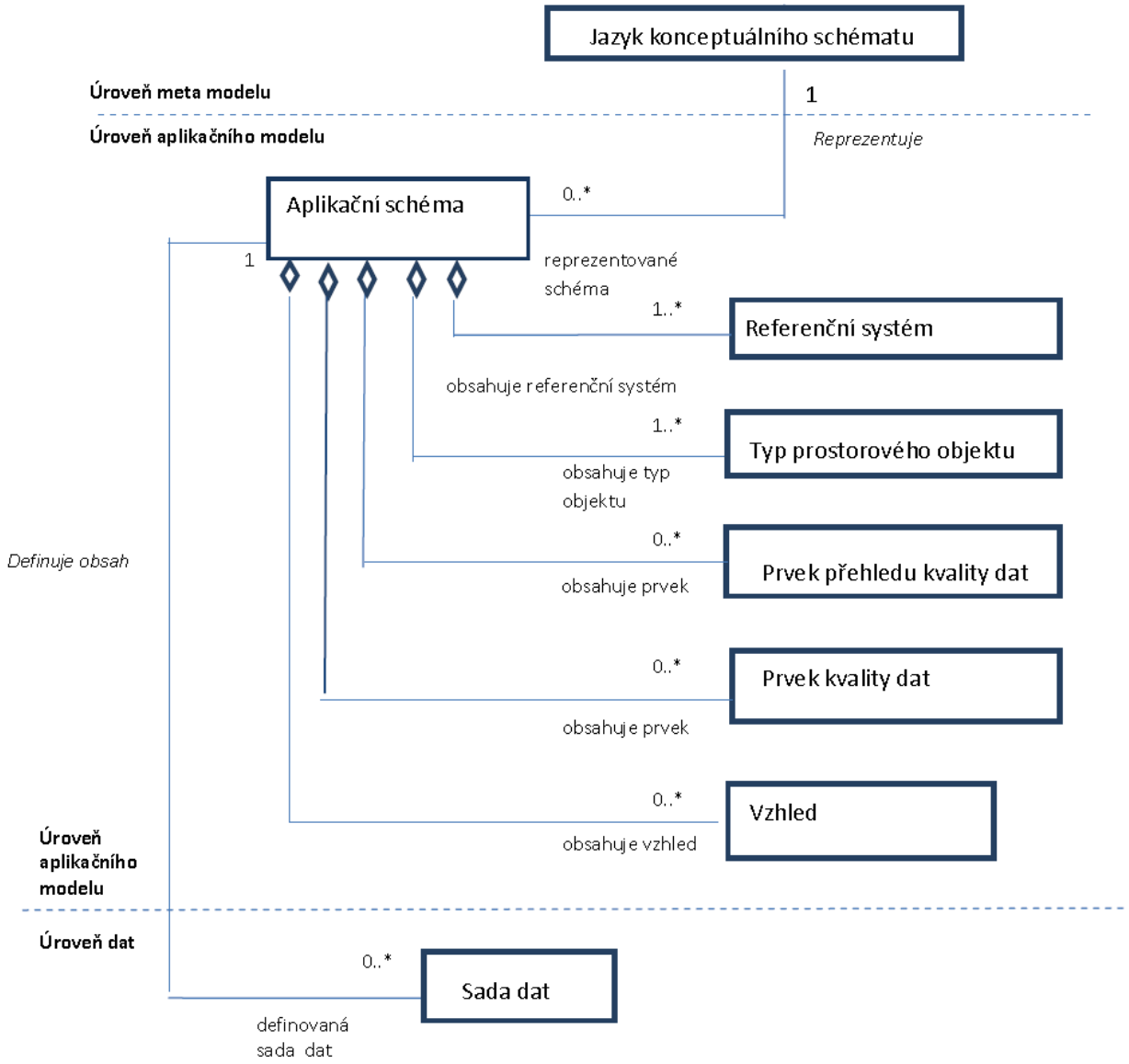
Aplikační schémata (úroveň aplikačního modelu) definují typy objektů, data a zpracování, která jsou doložena příklady v informační bázi. Používají konstrukty, které jsou definovány v meta-modelech. Aplikační schéma může být složeno z několika částí podle jednoho nebo více jazyků pro modelování, pro geografickou informaci obecně UML.

- **model (model)** – abstrakce některých aspektů reality

[ČSN EN ISO 19109:2021]

Úroveň dat

Aplikační schéma je konceptuální schéma, které definuje, jak je obor úvah popsán daty. Aplikační schémata obsahují sémantiky pro interpretaci dat stejně jako datové struktury pro jejich generování, např. XML schémata. Aplikační schéma poskytuje popis sémantické struktury datové sady a rovněž identifikuje typy prostorových vzhledů jevů, typy atributů vzhledů jevů, typy vztahů mezi vzhledy jevů a typy operací nad vzhledy jevů. Tyto budou přehledně zpracovány ve formě „katalogu typů vzhledů jevů“, blíže viz podkapitola 4.2. Aplikační schéma definuje také referenční systém a prvky kvality požadované pro poskytnutí úplného popisu geografické informace v datové sadě.



Obrázek 12: Prvky aplikačního schématu definující sadu dat
(Převzato z TNI CEN/TR 15449-3)

Výčet norem týkajících se geografické informace s vyznačením úrovní modelu geografické informace, v nichž je daná norma při modelování používána, je obsahem Přílohy C – Přehled ISO norem geografické informace.

POZNÁMKA: Předmětem zájmu těchto norem ISO geografické informace jsou i negeografické vzhledy jevů, které mohou být začleněny do aplikačního schématu neobsahujícího žádné prostorové charakteristiky.

Termín vzhled jevu ve čtyřvrstvé architektuře

Zavádění specifického a do jisté míry neobvyklého termínu „vzhled jevu“ pro jeden ze základních anglických pojmů „feature“ vyvolalo v odborné komunitě poměrně velmi živou diskusi a trvalo určitou dobu, než byl tento termín všeobecně přijat. K jeho zavedení do českých verzí norem ISO geografické informace došlo především na základě výsledků detailního zkoumání a odborného doporučení prof. RNDr. PhDr. Zdeňka Neubauera, CSc. z PřF UK a v současné době již v komunitě odborníků pro oblast geografické informace používání pojmu vzhled, jako ekvivalentu anglického pojmu feature, nemá alternativu. Širší význam informatického pojmu „vzhled“ je velmi kvalifikovaně a přitom srozumitelně popsán a jeho používání pregnantně zdůvodněno v [22].

Používání termínu vzhled jevů a jeho významy ve čtyřvrstvé architektuře geografické informace je přehledně uvedeno v následující tabulce 2.

Úroveň architektury	**Použití termínu „vzhled jevu“**
Meta-meta úroveň	„vzhled jevu“ jako obecný pojem, není specifikován typ ani instance
Meta úroveň	třída CLASS v UML vyjadřující Obecný model vzhledů jevů se jménem třídy „FeatureType“
Aplikační úroveň	specifický typ vzhledu jevu reprezentující třídu jevu reálného světa, např. „Road“ instance třídy CLASS „FeatureType“ Obecného modelu vzhledů jevů, vyjádřená v aplikačním schématu. Výsledkem v aplikačním schématu v UML je třída CLASS nazvaná „Road“.
Úroveň dat	Instance vzhledu jevu reprezentující sadu dat pro nějakou instanci typu vzhledu jevu, např. silnice „Route 66“

Tabulka 2: Termín vzhled jevu a jeho významy v čtyřvrstvé architektuře

POZNÁMKA: Zpravidla lze z kontextu pochopit, ze které úrovně je termín „vzhled jevu“ použit. Termín „typ“ nebo „instance“ je použit jako dodatek k termínu „vzhled jevu“, když je myšlen pouze jeden z nich.

Vztah meta modelů a ontologií

Meta modely úzce souvisí s ontologiemi. Obojí jsou často používány k popisu a analýze vztahů mezi pojmy.

- **Ontologie** vyjadřují něco významného ve specifikovaném oboru úvah nebo v uvažované doméně s uplatněním gramatiky pro používání slovní zásoby. Gramatika specifikuje správná a úplná prohlášení, tvrzení, dotazy a pod. (formální omezení) o tom, jak mohou být spolu použity pojmy v řízeném slovníku ontologie.
- **Meta modelování** může být považováno za explicitní popis (konstrukty a pravidly) toho, jak je vytvořen model specifický pro doménu. Zahrnuje to zejména formalizovanou specifikaci notací specifických pro danou doménu. Meta modely jsou obvykle množinou přísných pravidel a měly by se vždy dodržovat. Validní metamodel je ontologií, avšak ne všechny ontologie jsou explicitně modelované jako metamodely.

Problematice ontologie se v tomto dokumentu podrobně věnuje II. část Metodika pro budování otevřené a rozšířené ontologické struktury. K jejímu snazšímu porozumění mohou přispět, vedle četné odborné literatury uvedené v závěru druhé části, i následující definice základních pojmů, informace a poznámky:

- **ontologie** – formální reprezentace jevů oboru úvah včetně podkladového slovníku zahrnujícího definice a axiomy, které činí zamýšlený význam explicitním a popisují jevy a jejich vzájemné vztahy [ISO 19101-1:2014]

Ontologie v počítačové vědě a informatice zahrnuje reprezentaci, formální pojmenování a definice kategorií, vlastností a vztahů mezi pojmy, daty a entitami, které dokládají jednu, mnoho nebo všechny domény diskurzu. Zjednodušeně, ontologie je způsob, jak ukázat vlastnosti odborné oblasti a jak tyto vlastnosti navzájem souvisí, a to definováním množiny pojmů a kategorií, které tuto odbornost reprezentují.

Dalšími souvisejícími termíny jsou:

Taxonomie je teorie a praxe klasifikace věcí nebo pojmů, včetně principů, na nichž je tato klasifikace založena. Postupným vývojem se termín taxonomie stal synonymem pro termín **klasifikace**.

Tezaurus je referenční dílo k vyhledání synonym nebo antonym k danému slovu. Seskupuje termíny v hierarchické taxonomii pojmů nebo v abecedním pořadí, případně jiným způsobem. Tezaury většinou neobsahují definice pojmů, ale popisují rozdíly mezi podobnými slovy s poznámkou o jejich konotacích a významových odstínech. V informatice je tezaurus pokládán za druh **řízené slovní zásoby**. Tezaury jsou používány při automatizovaném zpracování lidské řeči pro zajištění jednoznačného rozpoznání smyslu slova a ke zjednodušení textů při jejich strojovém překladu do cizích jazyků.

Tezaurus může tvořit součást ontologie a být vyjádřen ve SKOS (the Simple Knowledge Organization System).

3.4 Podkladové terminologické slovníky

3.4.1 Slovník pojmů GeoInfoStrategie

Dokumentace GeoInfoStrategie [1] obsahuje rovněž Slovník pojmů definující 16 základních termínů. Využitelnost tohoto slovníku je však značně omezená, protože uvedené termíny a jejich definice neuvádí obor, zdroj ani odkazy, nejsou navzájem provázané, jsou zacyklené a jejich definice obsahují řadu dalších termínů, které je nutné také definovat. Konsolidace vybraných termínů pro účely projektu NaSaPO je uvedena v příloze A.

3.4.2 Terminologický slovník zeměměřictví a katastru nemovitostí

Tento poměrně rozsáhlý ve velké míře využívaný oficiální terminologický slovník resortu zeměměřictví a katastru nemovitostí je zpracováván Terminologickou komisí Českého úřadu zeměměřického a katastrálního. Do tohoto digitálního díla byl převzat značný počet termínů ze seznamů odborných termínů

v ISO normách geografické informace. Při jejich výběru byla rozhodujícím kritériem jejich užitečnost v jazykové doméně ČÚZK.

Tento resortní terminologický slovník je dostupný na : http://www.vugtk.cz/slovnik/index.php.

3.4.3 Terminologický výkladový slovník pro potřeby realizace Akčního plánu Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020

Tento terminologický slovník byl původně vypracován s ohledem na potřeby jednotného chápání a výkladu pojmů a jejich přesných definic k vzájemnému porozumění jak mezi tvůrci Akčního plánu GeoInfoStrategie vzájemně, tak i mezi budoucími uživateli geografických informací. Tomuto účelu podléhal i výběr termínů, který byl později rozšířen o některé vybrané termíny z oblasti územního plánování, dopravy, bezpečnosti, krizového řízení a životního prostředí. Slovník obsahuje i anglické ekvivalenty všech termínů a rozsáhlý seznam nejpoužívanějších zkratek. Terminologický slovník byl sestaven ve Výzkumném ústavu geodetickém, topografickém a kartografickém, v. v. i., skupinou předních expertů pod vedením doc. Ing. Jiřího Šímy, CSc. V roce 2016 byl jako jeden z výsledků projektu Technologické agentury ČR TB0500MV004 „Vypracování certifikované metodiky pro vyhodnocování stavu infrastruktury pro prostorové informace v České republice“ schválen jako výchozí výkladový slovník pro implementaci GeoInfostrategie. Jeho správcem a garantem dalšího rozvoje je MV ČR.

V současné době je uživatelům k dispozici jeho 4. verze, kterou je možné ve formátu PDF volně stáhnout na: https://www.mvcr.cz/soubor/terminologicky-vykladovy-slovnik-verze-4-pdf.aspx.

3.4.4 Terminologie ISO norem geografické informace v gesci ISO/TC211

Terminologii mezinárodních norem ISO pro geografickou informaci vydávaných ISO/TC 211 spravuje její pracovní skupina Terminology Maintenance Group (TMG), která průběžně udržuje terminologické repository a poskytuje projektovým týmům a pracovním skupinám podporu při harmonizaci terminologie. Současně koordinuje a publikuje „ISO/TC 211 Multi-Lingual Glossary of Terms“. Více informací o TMG lze nalézt na https://github.com/ISO-TC211/TMG.

Tento mnohojazyčný slovník termínů byl sestaven ze všech mezinárodních norem vydaných ISO/TC 211, obsahuje všechny anglické termíny v nich obsažené a tisíce jejich překladů dodaných členskými zeměmi. Jeho účelem je podporovat konzistenci při užití a interpretaci geoprostorových termínů. Multi-Lingual Glossary of Terms (MLGT) je volně dostupný ve dvou formátech:

- Původní verze v Microsoft Excel: 2020602 Edition ISO/TC211 Multi-Lingual Glossary of Terms (MLGT), 5. edice ze dne 02. 06. 2020
- Nová verze vyvinutého online vydání dostupná na Geolexica.

3.4.5 Další terminologické slovníky

Terminologický slovník OGC

K dosažení shody používaných odborných termínů, k jednotnému a správnému chápání jejich definic, a tím k zajištění sémantické interoperability, také OGC Technical Committee vytváří a poskytuje svůj oborový slovník „normativních" termínů, definic a akronymů z oboru geoprostorových technologií. Tento otevřený dokument je kompilován z mnoha zdrojů, je doplňován průběžně a je volně přístupný na https://www.ogc.org/ogc/glossary/.

Terminologický slovník W3C

Obdobně i W3C shromáždilo slovníky ze specifikací vypracovaných tímto konsorciem a různých dalších zdrojů a v roce 2003 vyvinulo W3C Glossary and Dictionary. Tento nástroj je množinou 1725 anglických termínů a jejich definic, které jsou propojeny se zdrojovým dokumentem, z něhož byl termín převzat. Termíny jsou rozčleněny do skupin podle odborných témat nebo tvoří úplný nečleněný seznam. Tento slovník již není udržován, ale obsah jeho archivované verze z roku 2010 je dostupný veřejnosti na: https://www.w3.org/2003/glossary/subglossary/All/.

Terminologický slovník INSPIRE

Společný mnohojazyčný terminologický slovník INSPIRE Glossary je jedním z registrů projektu INSPIRE. Jak již bylo uvedeno v článku 3.3.2.2, je vlastněn Evropskou komisí a udržován Společným výzkumným střediskem (JRC). Jednotné termíny a jejich definice ve 23 jazycích jsou určeny k používání v dokumentech Směrnice INSPIRE a Prováděcí pravidla INSPIRE. Slovník obsahuje odborné pojmy a definice převzaté jak z mezinárodních ISO norem geografické informace, tak i z dalších zdrojů, nebo vytvořené týmy v JRC. Tímto způsobem vzniklá terminologie je nekoherentní a je příčinou některých nedorozumění a ztráty sémantické interoperability, jak již bylo zmíněno v článku 3.3.2.2. Terminologický slovník je veřejně dostupný na: http://inspire.ec.europa.eu/glossary.

Příklady dalších terminologických slovníků

Z poměrně velkého množství existujících terminologických slovníků pro oblast geografické informace jsou dále uvedeny odkazy jen na několik významnějších, které jsou co do počtu hesel i rozsáhlejší:

https://www.fgdc.gov/policyandplanning/a-16/lexicon-of-geospatial-terminology – Lexicon of Geospatial Terminology — Federal Geographic Data Committee,

http://www.geoinformatik.uni-rostock.de/lexikon.asp – Geoinformatik Service GI-LEXIKON,

https://www.dgiwg.org/dgiwg-standards/106 – Technical Specification for the DGIWG Terminology Register, v. 1.0.1, 2008,

https://www.gov.uk/government/publications/geospatial-glossary/geospatial-glossary – Geospatial Glossary,

https://www.geo.ed.ac.uk/agidict/welcome.html – GIS Dictionary AGI UK,

https://www.wiley.com/college/chrisman/glossary.htm – John Wiley & Sons.

Zejména od roku 2000 jsou ve světě i v České republice věnovány značné odborné kapacity tvorbě a vytváření podmínek pro osvojování a používání správných termínů a využívání disponibilních technických norem pro oblast geografické informace. Výsledkem provedené rešerše poměrně velkého počtu relevantních terminologických slovníků je jednoznačné potvrzení zásadního významu jednotné závazné terminologie a její interpretace při modelování geografické informace, které je nutným předpokladem pro sémantickou homogenitu, srozumitelnost a mezioborovou a mezinárodní komunikaci. Byla zjištěna všeobecná a široká shoda v účelnosti používání mezinárodních, evropských a harmonizovaných národních norem geografické informace a byla tak potvrzena opodstatněnost ustanovení Aktualizovaného Akčního plánu GeoInfoStrategie i požadavků zadání tohoto projektu NaSaPO ohledně vytvoření a používání závazné terminologie a zcela jednoznačně a bez výjimek hovoří pro důsledné využívání konzistentní terminologie vycházející z ISO norem řady 19100 geografické informace.

Metodika a technologie pro tvorbu odborných tezaurů a slovníků pro potřeby rozvoje národní infrastruktury pro prostorové informace (projekt MV702)

Jednou z významných aktivit v rámci plnění Akčního plánu GeoInfoStrategie souvisejících s projektem NaSaPO a řešením problematiky terminologie byl v roce 2020 ukončený vývoj „Metodiky a technologie pro tvorbu odborných tezaurů a slovníků pro potřeby rozvoje národní infrastruktury pro prostorové informace“. Tento úkol byl řešen Výzkumným ústavem geodetickým, topografickým a kartografickým, v. v. i., v rámci projektu TITIMV702, jehož poskytovatelem byla Technologická agentura ČR. Cílem bylo optimalizovat a zefektivnit metody tvorby terminologických slovníků a tezaurů, zejména pro potřeby Českého úřadu zeměměřického a katastrálního, Ministerstva vnitra ČR a dalších uživatelů z oblasti veřejné správy a zájmových organizací a spolků. Metodika stanovuje formát, který umožňuje vzájemné propojení vytvářených terminologických slovníků a tezaurů, a základní principy tvorby obsahu terminologických slovníků a tezaurů.

Metodika byla certifikována a technologie je v současné době ověřována v rámci pilotního projektu jejím provozovatelem Ministerstvem vnitra ČR.

3.5 Principy tvorby mezioborových slovníků

Vývoj informačních produktů vyžaduje často získávání a kombinaci více datových sad od různých poskytovatelů. Způsob, jakým jsou data zkombinovaná, závisí na povaze zvažované potřeby a může kolísat od jednoduchého sestavení tematické reprezentace přes sofistikovanou integraci, analýzu a vyjádření. V každém případě dodavatelé dat i jejich zpracovatelé potřebují sdílet obecně přijímané pochopení datových charakteristik, aby byla zajištěna jejich správná interpretace a použití.

Při kombinování nesourodých datových sad vznikají problémy vyplývající z odlišných terminologických konvencí přijatých zúčastněnými dodavateli. Velmi často množina dat bude pocházet z komunity odborníků, kteří poskytují podporu konkrétnímu průmyslovému odvětví, např. silniční dopravě. Terminologie použitá k popisu obsahu, vztahů a chování dat odráží spojení pojmů z průmyslového odvětví s kulturou, konvencemi a praxí jeho specialistů. Konkrétní pojem může být tudíž identifikován různými termíny v závislosti na kontextu v odvětví, v němž je použit.

Schopnost zkombinovat data pocházející z různých odborných komunit závisí na všeobecném pochopení termínů a pojmů použitých k popisu významu dat. K tomu slouží křížové mapování, které sladí sémantické rozdíly mezi slovními zásobami zúčastněných komunit.

Metodika pro cross-mapping mezi různými slovními zásobami je stanovena mezinárodní normou ČSN EN ISO 19146:2018 Geografická informace – Mezidoménové slovníky, která byla primárně zamýšlena pro komunity z oblasti geoprostorových informací, ale možnosti jejího aplikování jsou širší. Specifikuje také implementaci mezinárodní normy ISO 19135-1:2015 při registraci položek mezidoménových (mezioborových) slovníků jako výsledků křížového mapování. Tato norma používá následující definici křížového mapování:

- **křížové mapování (cross-mapping)** – porovnání terminologických položek z různých domén k určení jejich sémantického vztahu

[ČSN EN ISO 19146:2018]

Účelem této normy není definovat ontologii nebo taxonomii, ale stanovit pravidla pro zajištění konzistence při implementaci procesů křížového mapování. Tato pravidla byla vytvořena s ohledem na taxonomické a ontologické pojmy a se záměrem umožnit budoucí sémantickou interoperabilitu. Lze tedy očekávat, že aplikování křížového mapování na slovní zásoby může být příspěvkem k budoucím iniciativám v oblasti taxonomie a ontologie.

3.5.1 Principy křížového mapování mezi doménovými (oborovými) slovníky

Křížové mapování je prostředek, kterým dvě a více komunit navzájem komunikují a formalizují své slovníky (slovní zásoby). Termíny a definice používané v komunitě identifikují pojmy charakterizující její filosofii, technologie a činnosti. Struktura slovní zásoby a podpůrné informace identifikují vztahy mezi pojmy a kontext, v němž musí být používány.

Ke vzájemné komunikaci těchto komunit je nutný prostředek, který tyto odlišné terminologie integruje, harmonizuje a logicky zdůvodní. Tuto funkci bude plnit všemi zúčastněnými stranami aktivně vypracovaný a odsouhlasený mezioborový slovník, který vyřeší sémantické rozdíly a umožní společné chápání pojmů.

Ve výkladových slovnících lze nalézt následující obecnou definici domény, která je relevantní pro kontext křížového mapování mezi různými slovními zásobami:

- **doména (domain)** – zřetelně vymezená oblast lidských znalostí, k nimž je přiřazena terminologická položka

[ČSN EN ISO 19146:2018]

Křížové mapování je integrální součástí šířeji pojatého procesu standardizace. Při vývoji slovníku pro mezioborovou (mezidoménovou) komunikaci i při jakékoliv další iniciativě křížového mapování by tedy měly být dodržovány základní principy standardizační praxe, jako jsou:

- společné angažmá všech dotčených komunit,
- nastolení otevřeného a transparentního prostředí,
- uznání autorit oprávněných jednat za zúčastněnou komunitu,
- práce se stabilními referenčními slovními zásobami,

- dodržování stanovených a společně odsouhlasených postupů,
- vytváření konsolidovaného a racionálního slovníku bez duplicit, zaměřeného na praktické využívání,
- uznání výsledného slovníku a jeho publikace v registru,
- vedení mezidoménového křížového mapování k rozšířené bázi znalostí,
- pravidelné revidování výsledného slovníku a zahrnování jakýchkoliv změn do něho,
- a další.

Mezidoménové křížové mapování by mělo souviset se slovními zásobami, které jsou udržovány v různých typech báze znalostí, jako jsou tezaury, taxonomie a ontologie. Spojením jejich obsahu se zavedenými doménovými pojmy je vytvořena rozšířená báze znalostí.

3.5.2 Metodika mezidoménového křížového mapování

Metodika křížového mapování mezi slovními zásobami různých domén specifikovaná v ČSN EN ISO 19146:2018 sestává z následujících čtyř fází:

a) **řídící rámec** – společná formalizace pracovních vztahů mezi zúčastněnými komunitami s cílem zajistit hladký průběh prací a legitimizaci výstupů;

b) **referenční slovní zásoba** – volba referenční slovní zásoby a ostatních slovních zásob v každé zúčastněné komunitě, které budou předmětem mapování nebo budou při něm využity;

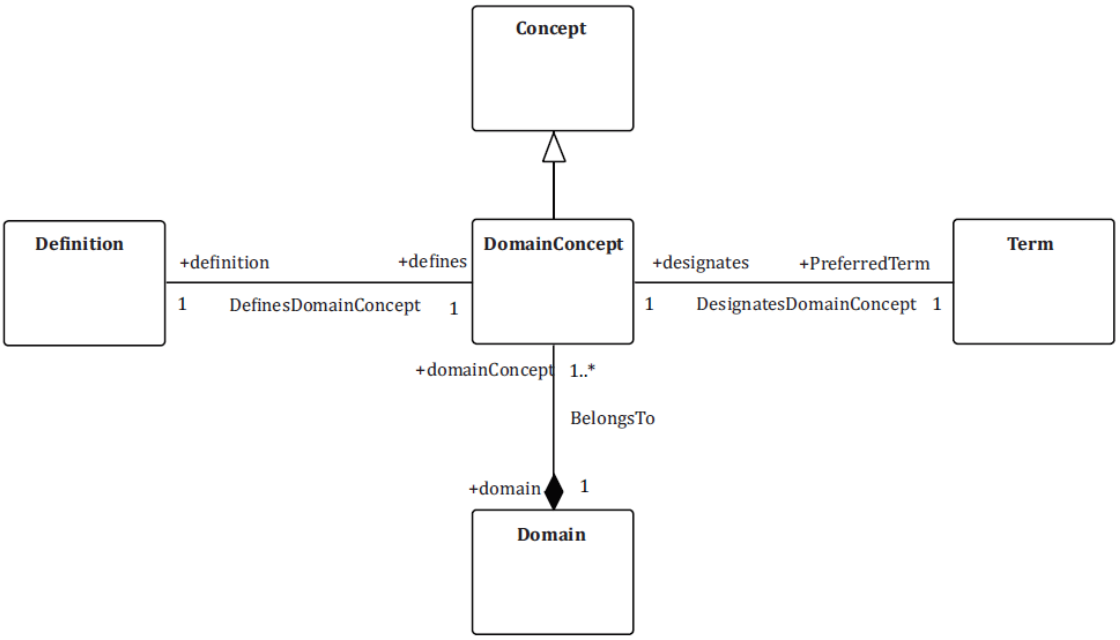
c) **vlastní křížové mapování** – tento proces sestává z postupných kroků:
   - ***identifikace kandidátních položek na křížové mapování*** – položky slovní zásoby jsou současně všemi zúčastněnými komunitami vyhodnoceny za účelem identifikování terminologických párů nebo shluků (clusterů), které se stanou kandidáty na křížové mapování. Terminologické páry nebo shluky by měly zahrnovat všechny výskyty, kde slovní zásoby obsahují:
     - identické termíny s identickými definicemi,
     - identické termíny s odlišnými definicemi,
     - podobné definice pro různé termíny.
   - ***porovnání termínů*** – analyzuje vztahy mezi pojmy a termíny v jednotlivých slovních zásobách.
     Termín je označení, pod nímž je nějaký pojem známý. Vztah mezi pojmem a označením není nutně jedinečným. Pojem může být označen dvěma nebo více termíny. Podobně jeden termín může označovat dva nebo více pojmů. U kandidátů na mezidoménové křížové mapování mohou být vzájemné vztahy mezi termíny v různých slovních zásobách klasifikovány jako:
     - *identické* – stejné termíny označující stejný pojem
     - *synonymické* – různé termíny označující stejný pojem, vzájemně zaměnitelné v jakémkoliv kontextu
     - *polysemické* – stejné termíny označující různé pojmy sdílející určité charakteristiky

- *homonymické* – stejné termíny označující různé pojmy, které navzájem nijak nesouvisí
- *žádné*

- ***porovnání pojmů*** – charakteristiky pojmu, které jej činí jedinečným, jsou popsány jeho definicí. Porovnání pojmů identifikuje rozsah, v němž jsou tytéž charakteristiky reprezentovány v záznamech slovní zásoby, která je porovnávaná. Vztahům mezi pojmy referenční a další slovní zásoby musí být přidělena jedna z těchto klasifikací:
  - identické
  - generalizace
  - specializace
  - rovnocenné
  - součást
  - agregace
  - uživatelem definovaná klasifikace (nutno konkrétně specifikovat)

d) **dokumentování a publikování** – spočívá v prezentování výsledků křížového mapování zúčastněným komunitám a v případě potřeby i veřejnosti, včetně zpřístupnění výsledků ve formě registru. Výsledkem jsou vzájemné vztahy mezi termíny a pojmy identifikátory domén, do nichž tyto termíny přísluší.

Požadovaný vztah mezi doménami, pojmy, definicemi a termíny je schematicky ilustrován na následujícím obrázku 13.



Obrázek 13: Termíny, definice, pojmy a domény (převzato z ČSN EN ISO 19146:2018)

3.5.3 Registr položek mezidoménového křížového mapování

Výsledky mezidoménového křížového mapování jsou zpravidla publikovány ve formě registru, přičemž jeden registr může být použit pro více iniciativ křížového mapování. V takovém případě musí registr pro každou iniciativu obsahovat samostatný subregistr pro referenční slovní zásobu, subregistr pro další pracovní slovní zásoby a samostatný subregistr pro položky křížového mapování.

Racionálním požadavkem je vytvářet registr výsledků mezidoménového křížového mapování v souladu s normou ČSN EN ISO 19135-1:2016 Geografická informace – Postupy registrace položek – Část 1: Základy s upřesňujícími požadavky specifikovanými v ČSN EN ISO 19146:2018.

4 Metadatové struktury a výměnné formáty pro třídy objektů NaSaPO

Tato kapitola se věnuje konceptuálnímu modelu standardizovaného popisu informačních zdrojů metadaty podle ISO norem geografické informace. Třídy objektů prostorových dat jsou popsány v rámci metadat, resp. katalogu vzhledu jevů. Výměnným formátem pro třídy objektů je tak formát pro přenos metadat. Standardně zvoleným formátem dokumentů metadat je formát XML.

Mnoho subjektů napříč VS vytváří důležité datové sady prostorových dat velmi vysoké hodnoty, které jsou pořizovány přímým měřením různými geodetickými, fotogrammetrickými a jinými metodami sběru geografické informace, nebo jsou z takto pořízených základních prostorových dat odvozeny (modelovou generalizací). Rozvoj Národní infrastruktury pro prostorové informace bude do budoucna provázen zvýšenou produkcí, obohacováním i případnou významnější změnou digitální prostorové informace. Tím bude narůstat množství, složitost a různorodost zdrojů prostorových informací. Rozvoj distribuovaných informačních systémů, který souvisí především s využíváním internetu, tak klade stále vyšší požadavky na standardizaci popisu těchto informačních zdrojů, který zprostředkovává porozumění všem jejich aspektům.

Zejména v souvislosti s řešením NaSaPO bude velmi důležité nastavit a používat správný rozsah metadat. Narozdíl od projektu INSPIRE, který je příkladem užití pouze malé části základních metadat podle skupiny norem ISO 19115, je nezbytné v rámci NaSaPO a obecně i v rámci celé Národní infrastruktury pro prostorové informace doplnit i další metadatové popisy podle ISO norem geografické informace týkající se především sběru, kvality a obsahu a struktury prostorových dat. Tento požadavek je o to naléhavější, protože v současnosti jsou metadata VS zaměřena pouze na popis datových sad a některých služeb, tedy spíše na distribuci prostorových dat, než na prostorová data samotná.

Ačkoliv primárním účelem metadat podle ISO 19115-1:2014 je popis digitálních geografických informací, mohou být schémata podle této normy použita i k popisu informačních zdrojů, které nemají geografický charakter nebo nejsou v digitální formě. Například k popisu tištěných textových dokumentů, softwarů, specifikací produktů a úložišť dat, map apod. Na druhé straně některé jiné odborné domény zpracovávají metadata podle jiných schémat, např. v knihovnictví jsou používána standardně schémata iniciativy Dublin Core. Běžnou praxí je, že si v případě potřeby komunita vytvoří pro svá metadata vlastní profil schématu z norem geografické informace skupiny ISO 19115 nebo z jí používaného specifického metadatového standardu. Je nutné poznamenat, že normám ISO řešícím problematiku metadat je ze strany TC 211 věnována značná pozornost, která odpovídá jejich významu a tyto normy jsou na základě odezvy z praxe průběžně doplňovány a rozvíjeny.

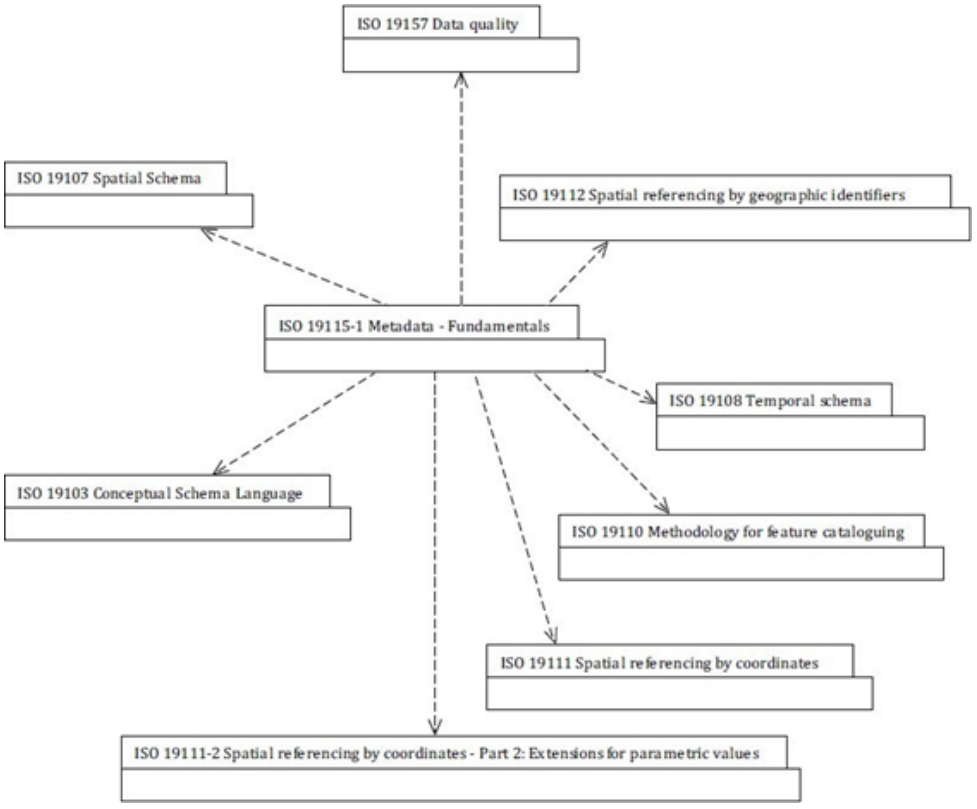
V České republice je pro metadata o prostorových informacích zaveden a již běžně používán Národní metadatový profil ČR, který byl navržen a vypracován v souladu s Nařízením Komise č. 1205/2008 a požadavky specifikovanými v Nařízení Komise č. 1089/2010. Tento profil odpovídá mezinárodním normám ISO geografické informace ČSN EN ISO 19115, ČSN EN ISO 19119 a ČSN EN ISO 19139. Národní profil je průběžně aktualizován a v současné době je k dispozici uživatelům již jeho verze 4.2. Je včetně editoru šablon pro validaci prostorových dat bezplatně dostupný např. na https://geoportal.cuzk.cz. Z hlediska potřeb NaSaPO nebo NIPI je však tento profil nevyhovující, protože neobsahuje řadu kriticky významných položek metadat (o kvalitě, o struktuře – aplikačním schématu, katalogu vzhledů jevů a dalších).

V současnosti je vývoj informačních systémů zaměřen na služby, především na webové služby, které jsou významnými zdroji informací. Mnohé z těchto služeb poskytují na místě založené informace nebo funkcionality. Popsat tyto služby pro jejich vyhledávání a využívání je rovněž jedním ze stěžejních úkolů metadat.

Vytváření metadat o zdrojích NaSaPO bude založeno na využívání všech tří částí normy ČSN EN ISO 19115. Schéma metadat požadované pro úplný popis všeobecně používané geografické informace a služeb je specifikováno v její 1. části. Poskytuje informace o jejich identifikaci, rozsahu, kvalitě, prostorových a časových aspektech, obsahu, prostorových referencích, zobrazení, distribuci a dalších vlastnostech. Tuto část normy ISO 19115-1 je možné aplikovat i na katalogizaci všech typů zdrojů a aktivity všech účastníků procesů tvorby NaSaPO.

Z tohoto výčtu je zřejmá těsná závislost této normy na dalších specifických normách ISO, přičemž využívá koncepty definované v jednom nebo ve více balíčcích těchto norem ISO, především norem řady 19100 pro geografické informace.

Obrázek 14 zobrazuje tyto balíčky vypracované ISO/TC 211, na nichž jsou základy normy ISO 19115 závislé.



Obrázek 14: Balíček základů metadat a závislosti (převzato z ISO 19115-1:2014)

4.1 Základní metadata

Za základní metadata jsou normou ISO 19115-1:2014 Geographic information – Metadata Part 1: Fundamentals považovány metadatové skupiny, entity a prvky pro popis zdrojů prostorové informace.

- **zdroj (resource)** – identifikovatelné aktivum nebo způsob, který splní nějaký požadavek [ČSN EN ISO 19115-1:2020]

  PŘÍKLAD: Datová sada, série datových sad, služba, dokument, iniciativa, software, osoba nebo organizace.

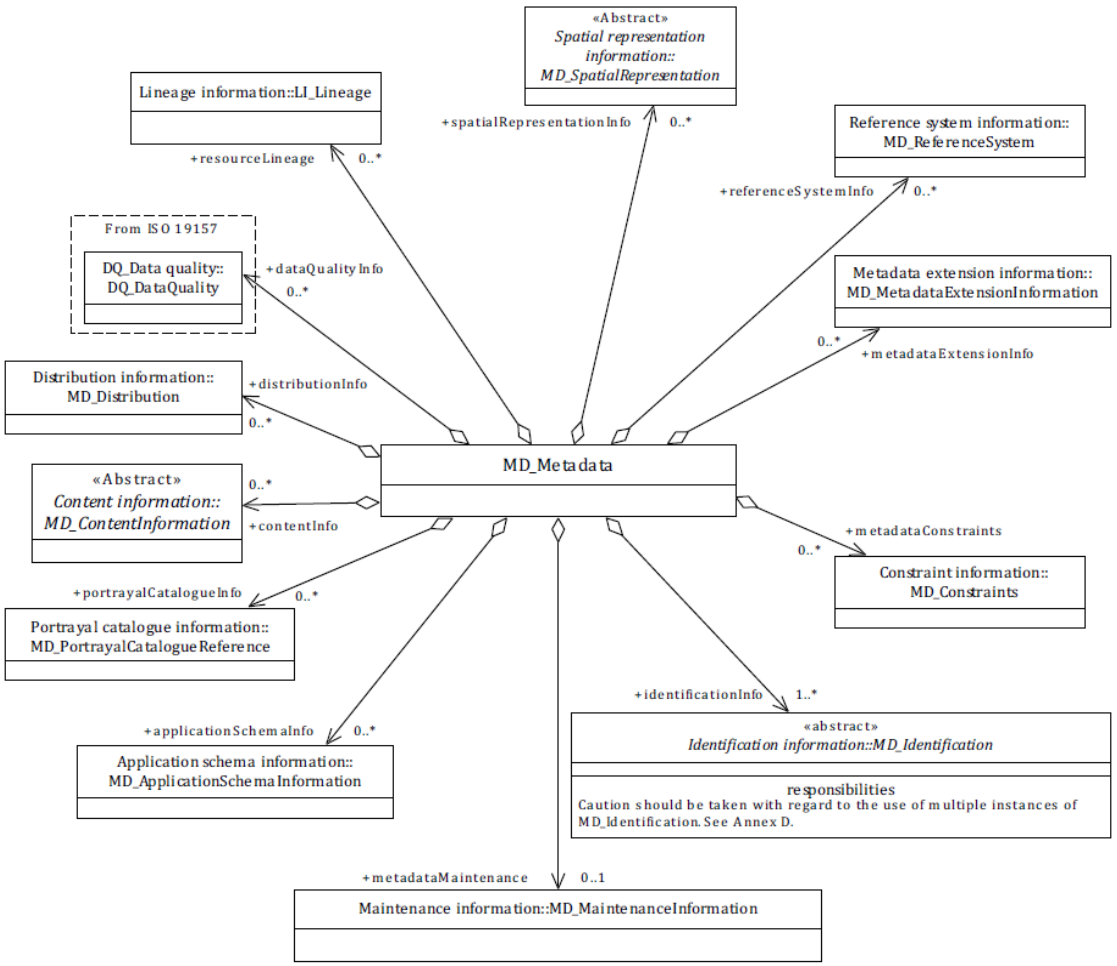
Norma podrobně definuje datové typy a vztahy mezi entitami pomocí modelu vyjádřeného v UML. Příloha B k ISO 19115-1:2014 definuje více než 400 metadatových prvků, jejichž cílem je zvýšit efektivnost vyhledání, hodnocení a zpřístupnění zdrojů prostorové informace tak, aby bylo umožněno jejich dlouhodobé opakované používání i k dalším účelům a jejich uchování.

Metadata jsou popsána pomocí jednoho nebo více balíčků, kdy každý balíček poskytuje oddělenou komponentu metadatové informace. ISO 19115-1:2014 definuje 13 balíčků, které jsou použity k popisu metadat (viz obrázek 15) a jsou udržovány v jednom integrovaném UML modelu. Jsou to balíčky: Metadata information, Identification information, Constraint information, Lineage information, Content information, Distribution information, Reference system information, Spatial representation information, Portrayal catalogue information, Metadata application information, Application Schema information, Metadata extension information a Service metadata information.

Jak je zřejmé z obrázku 15, úplná metadata jsou poskytnuta balíčkem MD_Metadata a agregátem 12 dalších tříd. Třída DQ_DataQuality je definována v ISO 19157:2013.

V následujícím výčtu tříd metadat podle ISO 19115-1:2014 jsou uvedeny stručné popisy informace, kterou poskytují:

- **Metadata o metadatech** (MD_Metadata) – o metadatech
- **Informace o identifikaci** (MD_Identification) – k jednoznačné identifikaci zdroje
- **Informace o omezeních** (MD_Constraints) – týkající se právních a bezpečnostních omezení vztažených na zdroje a metadata o zdrojích
- **Informace o údržbě** (MD_MaintenanceInformation) – vztahující se na rozsah a frekvenci údržby pro zdroje nebo týkající se zdrojů v jedné třídě
- **Informace o distribuci** (MD_Distribution) – týkající se distributorů a možných variant získání zdroje
- **Informace o prostorové reprezentaci** (MD_SpatialRepresentation) – identifikující prostorová primitiva použitá zdrojem a mechanismus použitý k modelování jevu reálného světa v digitálním informačním systému
- **Informace o referenčních systémech** (MD_ReferenceSystem) – identifikující prostorové, časové a parametrické referenční systémy použité zdrojem v jedné třídě



Obrázek 15: Schéma metadat (převzato z ISO 19115-1:2014)

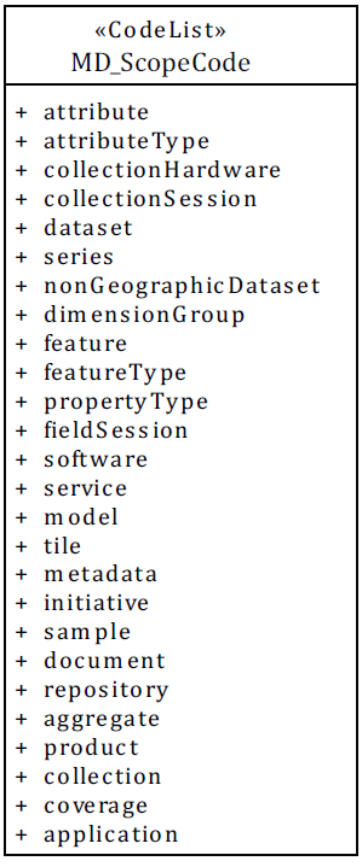
- **Informace o původu** (LI_Lineage) – týkající se výchozích zdrojů a produkčních procesů použitých při produkci zdrojů
- **Informace o obsahu** (MD_ContentInformation) – identifikující obsah zdroje citováním katalogů vzhledů jevů použitých při definování obsahu (viz ISO 19110:2016)
- **Informace o kvalitě** (DQ_DataQuality) – o kvalitě zdroje (viz ISO 19157:2013)
- **Informace o katalogu zobrazení** (MD_PortrayalCatalogueReference) – identifikující použitý katalog zobrazení
- **Informace o rozšíření metadat** (MD_MetadataExtensionInformation) – týkající se specifikovaných rozšíření metadat
- **Informace o aplikačním schématu** (MD_ApplicationSchemaInformation) – popisující aplikační schéma použité pro definování a odhalení struktury zdroje
- **Informace o metadatech služby** – týkající se služeb. Struktura metadat zahrnuje tři hlavní třídy:
    - základní metadata služeb (SV_ServiceIdentification) – obecný popis funkčností poskytovaných službou,
    - popis operací dostupných z konkrétní služby (SV_OperationMetadata),
    - popis dat dostupných z konkrétní služby (MD_DataIdentification).

Výše uvedené balíčky metadat jsou dnes již poměrně dobře známy díky aplikaci směrnice INSPIRE v ČR. Jak už bylo v tomto dokumentu uvedeno, metadata podle směrnice INSPIRE jsou aplikací (UseCase) norem ISO geografické informace. Pro Národní infrastrukturu pro prostorové informace je však zcela nezbytné umožnit plné využití potenciálu metadat stanovených v normách ISO geografické informace. Směrnice INSPIRE nepožaduje poskytovat mnohá metadata, která jsou normami ISO považována za významná a naopak. V rámci NIPI je třeba se podrobněji zaměřit především na informace o kvalitě, obsahu a původu zdrojů prostorové informace.

Metadata zpracovaná podle ISO 19115-1:2014 se mohou vztahovat k široké škále zdrojů informací. Možné rozsahy platnosti metadat zdrojů informací jsou definovány v rámci balíčku MD_Metadata v seznamu kódů MD_ScopeCode (viz obr. 16).

Další 4 balíčky Citation information, Responsible party information, Language-characterset localisation information a Extent information jsou použity jinými balíčky. Jednotlivé balíčky mohou být použity samostatně, aby pro splnění konkrétního případu užití poskytly vyčleněné komponenty metadat. Pokud je však poskytována kompletní sada metadat, musí být použity minimálně balíčky Metadata a Identification information.

V případě potřeby poskytnout doplňková metadata, mohou být přidány dodatečné balíčky.



Obrázek 16: MD_ScopeCode (převzato z ISO 19115-1:2014)

Vedle vlastního obsahu metadat vytvořeného podle výše uvedených zásad jsou neméně významné i metody a problémy implementace a správy metadat pro účely vyhledávání a prohlížení i výměny a prezentace. Tyto otázky a možné přístupy jsou velmi podrobně, včetně schémat a praktických příkladů, zpracovány zejména v následujících v přílohách k základní mezinárodní normě ISO 19115-1:2014.

- Příloha C – návody na tvorbu profilů k rozšíření metadat v případě specifických požadavků uživatele
- Příloha D – vzorový příklad tvorby metadat pro sadu geoprostorových dat
- Příloha E – přehled metod implementace a správy
- Příloha F – seznam metadatových položek, které usnadní zjišťování zdrojů geoprostorových dat

K racionálnímu a efektivnímu zpracování a využívání metadat významně přispívá možnost sdílení vytvořených agregovaných záznamů metadat. Při něm jsou použity objekty MD_Scope k indikování rozsahu a/nebo typu zdroje, který metadatový záznam popisuje. Zahrnuje MD_ScopeCode, je užitečný jako rychlý indikátor pro vyhledávání, a MD_ScopeDescription sloužící k doplnění podrobností. Praktické použití jednotlivých kódů oblastí je popsáno v Příloze E.5 k ISO 19115-1:2014.

4.1.1 Rozšíření pro získávání a zpracování

Nově jsou v metadatech specifikovány požadavky na standardizované údaje o sběru a zpracování prostorových dat, jejichž existence a dostupnost je jedním z významných předpokladů, aby data mohla být považována za garantovaná. Schéma potřebné ke zvýraznění významu popisu sběru a zpracování geografických informací, včetně informací z leteckého snímkování a dálkového průzkumu Země, je definováno v mezinárodní normě ISO 19115-2:2019 Geographic information – Metadata – Part 2: Extensions for acquisition and processing. Tato Část 2 zahrnuje požadavky na sběr a zpracování, vlastnosti měřících systémů, numerických metod a výpočetních postupů používaných k odvození geografické informace z dat pořízených těmito systémy. Rozšiřuje základní Část 1 přidáním dalších modelů: metadat (MD_Metadata), kvality dat (DQ_DataQuality v mezinárodní normě ISO 19157:2013), prostorové reprezentace (MD_SpatialRepresentation) a informací o obsahu (MD_ContentInformation).

4.1.2 Výměnný formát metadat

Části 1 a 2 normy ISO 19115:2014 společně definují konceptuální modely pro obsah metadat, které jsou nezávislé na nějakém konkrétním schématu kódování. (Schématům pro kódování se věnuje technická specifikace ISO/TS 19139-1:2019 Geographic information – XML schema implementation – Part 1: Encoding rules, která definuje schémata XML (eXtensible Markup Language). Kódování obsahu integrovaných konceptů metadat vytvořených v souladu s Částmi 1 a 2 normy ISO 19115:2015 řeší další 3. část této normy, jíž je technická specifikace ISO/TS 19115-3:2016 Geografická informace – Metadata – Part 3: XML schema implementation for fundamental concepts, která fakticky nahrazuje zrušené ISO/TS 19139:2007 a ISO/TS 19139-2:2012 a umožňuje automatizovanou validaci standardními programovými prostředky. Tato technická specifikace popisuje postup používaný k vytvoření XML schématu z konceptuálních modelů geografické informace dle norem ISO vztahujících se k metadatům.

Postup (je obsahem především její kap. 8) spočívá v tom, že je nejprve aplikováním pravidel definovaných v Příloze A k ISO 19118:2011 z konceptuálních modelů podle ISO 19115-1:2014 a ISO 19115-2:2019 odvozeno integrované schéma a poté je aplikována technická specifikace ISO/TS 19139-1:2019 na přijatou a k XML implementaci připravenou verzi konceptuálních UML modelů.

Výsledkem aplikování této technické specifikace ISO/TS 19115-3:2016 jsou:

a) sada XML schémat požadovaných k validaci dokumentů instancí metadat vyhovujících prvkům konceptuálního modelu definovaným v ISO 19115-1:2014, ISO 19115-2:2019 a ISO/TS 19139-1:2019;
b) sada pravidel stanovených normou ISO/IEC19757-3 (Schematron), která implementují validační omezení v UML modelech podle ISO 19115-1:2014 a ISO 19115-2:2019 která nejsou validována XML schématem;
c) Extensible Stylesheet Language Transformation (XSLT) pro transformování metadat podle ISO 19115-1:2014 kódovaných s použitím XML schématu podle ISO/TS 19139-1:2019 a metadat podle ISO 19115-2:2019 kódovaných s použitím XML schématu podle ISO 19139-1:2019 do ekvivalentního dokumentu, který je validní podle XML schématu definovaného v této technické specifikaci ISO/TS 19115-3:2016.

Přijetím tohoto XML schématu geografických dat v komunitě sdílející informace budou získány výhody standardizace při zjišťování, přístupu, použití informačních zdrojů a jejich chápání.

4.1.3 Specifikace datového produktu

Specifikace datového produktu (SDP) je jedním z prvních vytvořených a také nejdůležitějších dokumentů. Vytváří základ pro produkci nebo pořizování dat. Požadavky jsou definovány v úvodní fázi vývoje datového produktu a jsou široce využívány ve všech etapách vývoje a využívání produktu prakticky všemi účastníky. SDP shrnuje základní požadavky zadavatele produktu, řídí se jimi jeho vývojáři a výrobci. Podle informací v tomto dokumentu a podle metadat si potenciální uživatelé produktu posuzují jeho vhodnost k použití pro svůj konkrétní účel. Určení a obsah SDP nutně zakládá pevný vztah s metadaty tohoto produktu, i když obsahují do jisté míry odlišné údaje. Rozdíl je dán tím, že SDP popisuje, jaký produkt má být, a metadata, jaký produkt je.

Přijatý způsob formalizovaného vyjádření požadavků na datový produkt s použitím pojmů z ostatních mezinárodních norem ISO, jeho shoda a popisy testů jsou obsahem mezinárodní normy ČSN ISO 19131:2010 Geografická informace – Specifikace datového produktu. Tato norma poskytuje pomoc při vytvoření snadno srozumitelné specifikace produktu, který bude odpovídat zamýšlenému účelu. Norma také konkretizuje kódování SDP v XML.

SDP zpracovaná podle mezinárodní normy ČSN ISO 19131:2010 je tvořena především těmito povinně zpracovanými oddíly, které pokrývají hlavní stránky datového produktu:

- ***Přehled*** – je určen k poskytnutí krátkého slovního úvodu k pochopení SDP. Obsahuje informace o vytvoření SDP, název a akronymy produktu, použité zkratky, informativní popis datového produktu (obsah, prostorový a časový rozsah, konkrétní účel, zdroje a procesy produkce, údržba apod.)

- ***Oblasti specifikace*** – popisuje oblasti SDP, které vymezují prostorový a časový rozsah, typy vzhledů jevů a zahrnuté vlastnosti, prostorovou reprezentaci nebo postavení v hierarchii produktů. Může specifikovat rozdělení datového obsahu produktu na základě jednoho nebo více kritérií. Takové rozdělení může být pro různé části SDP odlišné, přičemž každá takto vzniklá část datového obsahu musí být popsána oblastí specifikace (tzv. scope).

- ***Identifikace datového produktu*** – musí obsahovat titul, stručné narativní shrnutí obsahu datového produktu, hlavní témata, rozsah geografického území pokrytého produktem. Volitelně také např. alternativní titul, účel produktu, typ prostorové reprezentace, prostorové rozlišení a další popisné informace.
- ***Obsah a struktura dat*** – závisí na tom, zda jsou data založená na vzhledech jevů nebo na pokrytí a zobrazení:
    a) obsah datového produktu založeného na vzhledech jevů je popsán jako aplikační schéma a katalog vzhledů jevů. SDP musí obsahovat popis každého vzhledu jevu v tomto produktu, proto do ní musí být zahrnut buď odkaz na katalog vzhledů jevů a atributů, nebo přímo jeho popis. Bližší informace k tvorbě aplikačních schémat podle ČSN EN ISO 19109:2021 a katalogů vzhledů jevů podle ČSN EN ISO 19110:2017 Geografická informace – Metodologie katalogizace vzhledů jsou uvedeny v podkapitole 3.3 a dále v čl. 4.2.1.
    b) SDP založený na pokrytí a zobrazení musí identifikovat každé pokrytí a každý obraz v tomto produktu a musí poskytnout pro každé z nich narativní popis. Informace o obsahu musí být popsána podle ČSN EN ISO 19123:2007 Geografická informace – Schéma pro geometrii a funkce pokrytí.
- ***Referenční systémy*** – musí zahrnovat prostorové a časové referenční systémy použité v datovém produktu. Použitý prostorový referenční systém může být souřadnicový referenční systém definovaný v ČSN EN ISO 19111:2020 nebo prostorový referenční systém používající geografické identifikátory definované v ČSN EN ISO 19112:2019. Časový referenční systém musí být definován podle ČSN EN ISO 19108:2004.
- ***Kvalita dat*** – požadavky na kvalitu dat v produktu musí být vyspecifikovány včetně určení přijatelné shody úrovní kvality a odpovídajících měr kvality pro všechny prvky a podprvky kvality dat v souladu s ČSN EN ISO 19157:2015. Pokud některý z prvků a podprvků kvality dat nemůže být použit, je nutné tuto skutečnost ve SDP jmenovitě uvést. Výsledky hodnocení kvality jsou publikovány v metadatech způsobem stanoveným v ČSN EN ISO 19115-1:2020.
- **Dodávka datového produktu** – SDP musí identifikovat jakékoliv požadavky na jeho dodávku. Tyto musí zahrnovat informace o formátu (název a verze, struktura, jazyky, norma kódování) a případně i o nosiči dodávky (popis jednotek, velikost, název nosiče a další).
- ***Metadata*** – obsahuje hlavní prvky metadat definované v ČSN EN ISO 19115-1:2020, která musí být zahrnuta do datového produktu. Jakákoliv požadovaná další dodatečná metadata musí být stanovena, včetně jejich formátu a kódování v SDP.

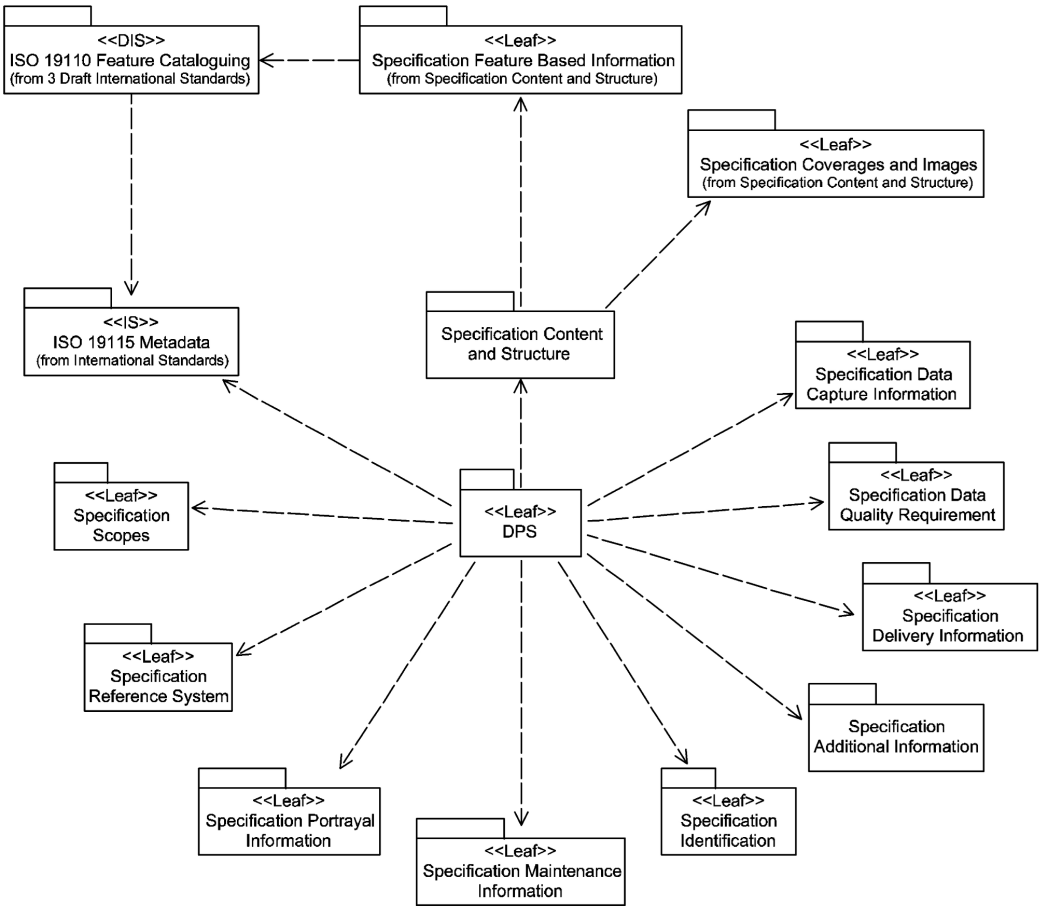
Dále je doporučováno zpracovat v SDP také problematiky:

- ***Sběr dat*** – informace o tom, jak jsou data sbírána. Je-li tento oddíl zařazen, musí být popsány zdroje a procesy, jež musí být použity. Může být ponechána volnost ve výběru procesů nebo být vyspecifikován jen jeden určitý proces sběru dat.
- ***Údržba dat*** – je-li tento oddíl zařazen, musí popsat zásady a kritéria aplikované v údržbě dat. Musí zahrnovat frekvenci údržby a aktualizace, a frekvenci, s jakou se změny a doplňky promítají do produktu.

- ***Znázornění dat*** – představuje popis, jak budou data v produktu prezentována, jako grafický výstup, diagram nebo obraz. Pokud je zařazen tento oddíl, musí mít formu odkazů na množinu pravidel a znázornění a množinu specifikací znázornění. Katalog znázornění musí být zpracován podle ČSN EN ISO 19117:2014.
- ***Další informace*** – do tohoto oddílu mohou být zahrnuty jakékoliv další informace SDP, které nejsou obsaženy ve výše uvedených oddílech nebo jinde v ČSN EN ISO 19131:2010.

S výjimkou oddílu Přehled, který je zpracováván ve formě běžně čitelného volného textu, je každý oddíl zpracován jako balíček UML. Model pro specifikaci datového produktu je popsán množinou balíčků, jejichž uspořádání je ilustrováno na obrázku 17. Datový produkt je plně specifikován v diagramech modelu UML a částečně v datovém slovníku pro každý balíček, které jsou uvedeny v přílohách k této mezinárodní normě ISO 19131:2010. Minimální popis datového produktu je v každém oddíle tvořen stanovenými povinnými prvky.

Pokud jsou v rámci některé sekce uplatněny konkrétní požadavky a popisy jen pro části datového produktu, tyto rozčlení sekci na ohraničené oblasti, přičemž pro každou z nich bude platit unikátní sada charakteristik. Každá taková sada charakteristik je s touto konkrétní oblastí specifikace asociována. Tato možnost platí pro všechny sekce s výjimkou sekce identifikace produktu, která vždy určuje celý datový produkt.



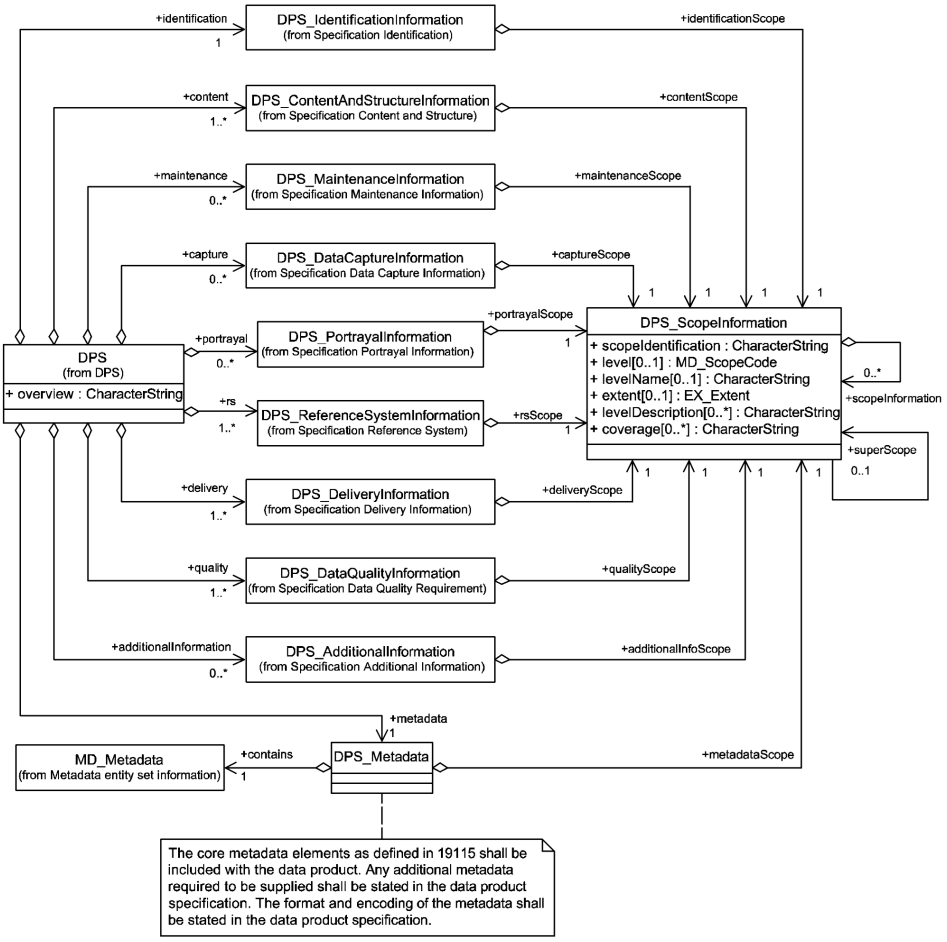
Obrázek 17: Balíčky specifikace datového produktu (převzato z ČSN ISO 19131:2010)

V praxi je běžné, že jsou vzneseny odlišné požadavky na různé části obsahu dat. Nejčastěji jsou jako základ sloužící pro členění obsahu dat do částí používána následující kritéria:

- prostorový nebo časový rozsah
- typ vzhledu jevu
- typ vlastnosti
- hodnota vlastnosti
- prostorová reprezentace a další.

V tomto případě musí být v SDP přesně popsáno toto rozdělení obsahu datového produktu do částí na základě jednoho nebo více kritérií. Každá tato část obsahu dat je považována za oblast specifikace (specification scope), přičemž může dědit charakteristiky celkové oblasti specifikace, případně charakteristiky jiné jí nadřazené parciální oblasti, nebo na ni nedbat. Obdobně může být tato oblast naopak agregátem jiných parciálních oblastí, které jsou jejími částmi.

Každá jednotlivá oblast je popsána položkami, jako jsou: identifikace oblasti, indikace hierarchické úrovně dat (vyjádřeno hodnotou MD_ScopeCode) a její jméno, prostorový, časový a výškový rozsah, podrobný popis úrovně dat, pokrytí na něž se informace vztahuje, oblasti které jsou částmi této oblastmi a oblast nadřazená této oblasti.



Obrázek 18: UML model pro specifikaci datového produktu (převzato z ČSN ISO 19131:2010)

4.2 Katalog vzhledů jevů a pojmové slovníky vzhledů jevů

Tato kapitola se podrobně věnuje katalogu vzhledů jevů a pojmovým slovníkům vzhledů jevů, které jsou vedeny jako registry. Katalog vzhledů jevů je často odvozen z pojmových slovníků vzhledů jevů. Oba typy registrů mohou výrazně pomoci vyhledání, zpřístupnění a využití prostorových informací.

Definované schéma registrů je ve shodě s ISO 19135-1:2015 Geographic information – Procedures for item registration – Part 1: Fundamentals, která specifikuje postupy registrace položek geografické informace a dále specifikuje strukturu hierarchického registru, tedy strukturu, kdy hlavní registr obsahuje množinu položek určených k popisu dílčích registrů. Pro Národní infrastrukturu pro prostorové informace se jeví jako nejvhodnější budování hierarchického registru, který obsahuje dílčí registry pojmových slovníků vzhledů jevů a registry katalogů vzhledů jevů. Tento „národní“ hierarchický registr katalogů a pojmových slovníků poslouží především harmonizaci prostorových informací a zajištění interoperability mezi různými subjekty veřejné správy. Z uvedených důvodů by měl být pevnou součástí centrálního informačního systému veřejné správy v oblasti prostorových informací IS NIPI.

IS NIPI, resp. metadatový IS, registry katalogů vzhledů jevů a registry pojmových slovníků vzhledů jevů slouží jako zdroj odkazů (referencí) pro křížové odkazování mezi registry založené na všech úrovních VS. Křížové odkazování může být však složité, pokud struktura registrů by byla rozdílná. Proto je nezbytné jednotlivé registry založit na normovaných strukturách pro zajištění slučitelnosti jednotlivých registrů.

katalog vzhledů jevů (feature catalogue) – katalog obsahující definice a popisy typů vzhledů jevů, atributy vzhledu jevů a vztahů mezi vzhledy jevů vyskytujícími se v jedné nebo více množinách geografických dat, a to zároveň s jakýmikoliv operacemi vzhledů jevů, které mohou být uplatněny

[ČSN EN ISO 19115-1:2020]

pojem vzhledu jevu (feature concept) – pojem, který může být podrobně specifikován jako jeden nebo více typů vzhledů jevů

[ČSN EN ISO 19126:2012]

pojmový slovník vzhledů jevů (feature concept dictionary) – slovník, jenž obsahuje definice a související popisnou informaci o pojmech, které mohou být podrobně specifikovány v katalogu vzhledů jevů

[ČSN EN ISO 19126:2012]

hierarchický registr (hierarchical register) – strukturovaná množina registrů pro doménu položek registrů, sestávající z hlavního registru a množiny dílčích registrů

[ISO 19135-1:2015]

4.2.1 Katalog vzhledů jevů

Katalogy vzhledů jevů (feature catalogues) definují typy vzhledů jevů, jejich operace, atributy a asociace, které jsou nezbytné pro přeměnu dat na použitelné informace. Dodavatelé a uživatelé musí mít jednotné chápání druhů jevů reálného světa reprezentovaných v prostorových datech tak, aby uživatel mohl jasně vyhodnotit, zda prostorová data jsou vhodná pro zamýšlený účel. Katalog vzhledů jevů tak podporuje šíření, sdílení a použití prostorových dat prostřednictvím poskytování informací o obsahu a významu dat.

Katalog vzhledů se vytváří pro konkrétní kontext, jako je například aplikační oblast nebo specifikace produktu. Je přímo sdružený s aplikačním schématem a s jednou nebo více datovými sadami vytvořenými v souladu s tímto aplikačním schématem. Obsahuje úplnou specifikaci každého typu vzhledu jevu, jeho vlastností a vztahů, včetně seznamu jeho povinných nebo volitelných vlastností. Každá z těchto vlastností je rovněž explicitně specifikována.

Klíčovými mezinárodními normami ISO geografické informace, které řeší vymezení pojmu vzhled jevu a stanoví metodiku tvorby katalogu vzhledů jevů, jsou:

ČSN EN ISO 19109:2021 Geografická informace – Pravidla pro aplikační schéma definuje pravidla pro tvorbu aplikačních schémat, pro které využívá Obecný model vzhledu jevu (GFM). Tento stanovený koncept popisu typů vzhledů jevů a jejich vlastností je dále rozvinut v rámci ISO 19110:2016.

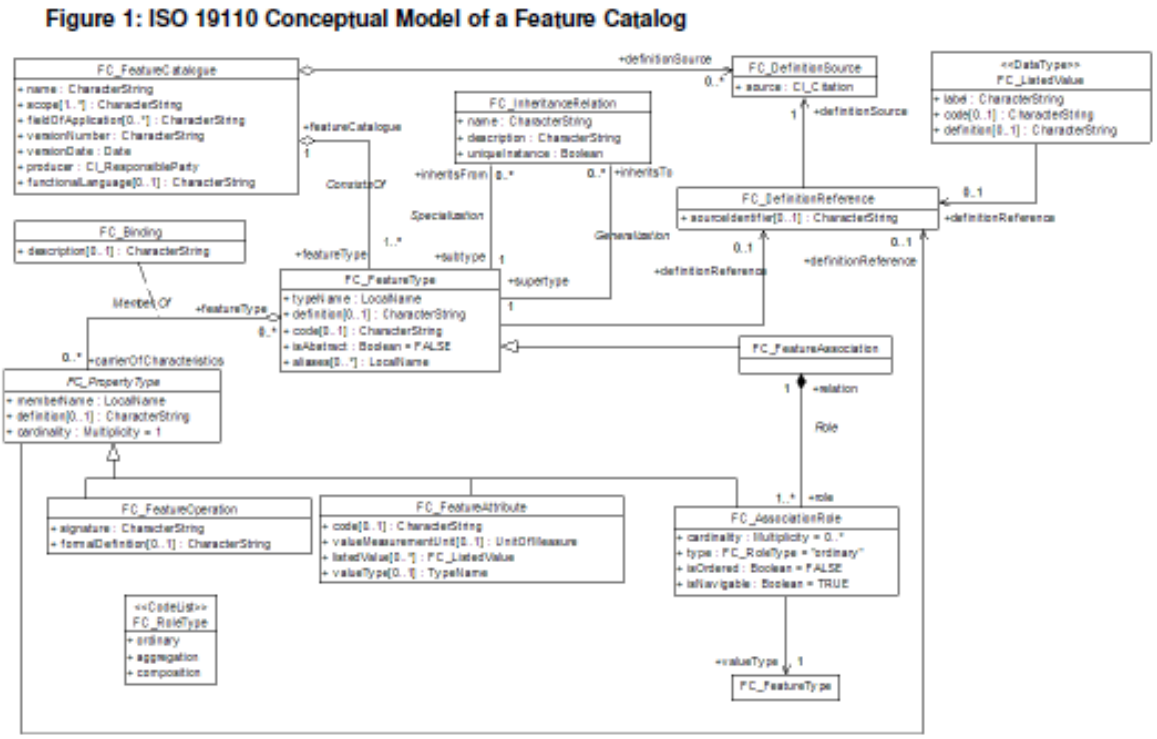
ISO 19110:2016 Geographic information – Methodology for feature cataloguing, která je do soustavy ČSN převzata vyhlášením (viz ČSN EN ISO 19110:2017 Geografická informace – Metodologie katalogizace vzhledů) a poskytuje normovaný rámec pro organizaci a zápis klasifikace abstrakcí jevů reálného světa v množinách prostorových dat, který umožňuje uživatelům snadnou orientaci v nich a jejich správné pochopení.

Principy definování vzhledů jevů jsou popsány v kapitole 7 mezinárodní normy ISO 19109:2015, která specifikuje Obecný model vzhledu jevu (GFM). Obě výše uvedené normy se zabývají popisnými (tematickými) stránkami vzhledů jevů. K vyjádření prostorových charakteristik vzhledů jevů jsou používány prostorové (geometrické a topologické) atributy. Jejich struktury jsou definovány v prostorovém schématu, který je předmětem především mezinárodních norem:

ISO 19107:2019 Geographic information – Spatial schema, která specifikuje konceptuální schémata pro popisování prostorových charakteristik a množinu prostorových operací konzistentních s těmito schématy. Ošetřuje vektorovou geometrii a topologii. Pro prostorové (geometrické a topologické) objekty definuje standardy prostorových operací pro používání při přístupu, dotazování, správě, zpracování a výměně dat.

ISO 19136-1:2020 Geographic information – Geography Markup Language (GML) – Part 1: Fundamentals – Geografický značkovací jazyk (GML) je XML kódováním v souladu s mezinárodní normou ISO 19118:2011 používaným při přenosu a ukládání geografických informací modelovaných ve shodě s konceptuálním modelovacím rámcem daným mezinárodními normami ISO řady 19100 a obsahuje prostorové i neprostorové vlastnosti geografických vzhledů jevů. Část 1 – Základy této normy definuje syntaktická pravidla XML Schématu, mechanismy a konvence.

POZNÁMKA: Jestliže je za základ pro ukládání a přenos geografických informací použito aplikační schéma zpracované ve shodě s ISO 19109 a popsané v UML, tato norma ISO 19136-1:2020 poskytne normativní pravidla pro mapování tohoto aplikačního schématu do aplikačního schématu GML v XML Schématu a jeho převod kódováním XML do dat v logické struktuře, která je v souladu s tímto aplikačním schématem zpracovaným podle ISO 19109:2015.



Obrázek 19: Konceptuální model katalogu vzhledů jevů (převzato z ISO 19110:2016)

Katalog vzhledů jevů má být odvozen z pojmových slovníků vzhledů jevů a má být pevnou součástí metadat, resp. je spojen s konkrétní datovou sadou nebo specifikací datového produktu.

Různé komunity uživatelů mohou své požadavky na soubory informací zformulovat s využitím katalogů vzhledů jevů:

- všeobecných – buď samostatných nebo sestavených z referenčních pojmů z externích pojmových slovníků vzhledů jevů
- specializovaných – zavádějících obsah vzhledů jevů a strukturu specifikací datových produktů a/nebo aplikačních schémat (buď samostatných nebo z referenčních typů vzhledů jevů a pojmů z externích univerzálních katalogů vzhledů jevů a z externích pojmových slovníků vzhledů jevů)
- představujících mapování a/nebo korespondenci mezi různými specializovanými a/nebo univerzálními katalogy vzhledů jevů.

4.2.2 Pojmové slovníky

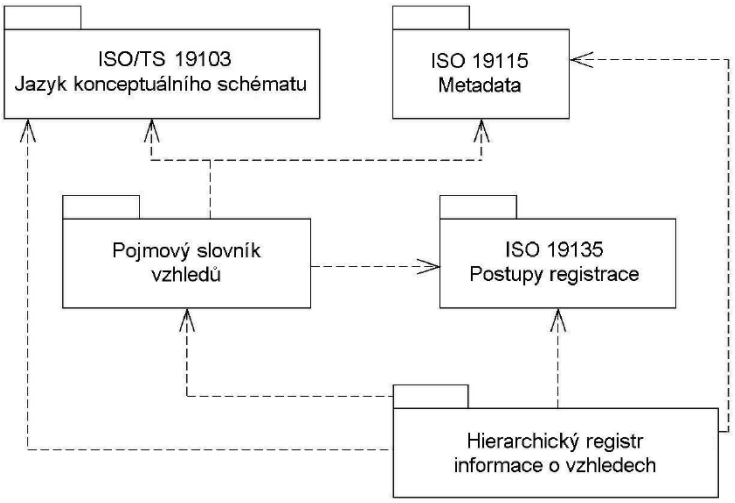
Pojmové slovníky jsou významnými dokumenty, které zásadním způsobem pomáhají vzájemnému porozumění mezi tvůrci a uživateli geografické informace z různých domén. Pojmové slovníky vzhledů jevů se na rozdíl od katalogů vzhledů jevů vytváří v širším kontextu. Je tvořen pro překlenutí mnoha aplikačních oblastí nebo k podpoře rozmanitých specifikací produktů. Není přímo spojen ani s aplikačním schématem ani s žádnou specifikací datového produktu.

Obsahuje definice, které identifikují podstatné charakteristiky typu vzhledů jevů, nespecifikuje však, které typy vlastností musí být v typu vzhledů jevů obsaženy. Může zahrnovat definice typů vlastností vzhledů jevů a případně i identifikovat typy vzhledů jevů, v nichž by mohly být takové typy vlastností obsaženy, ale nestanoví pro to žádné požadavky.

Pojmové slovníky vzhledů jevů obsahují pojmy, které se využívají k popisu vzhledů jevů a jsou sdíleny a využívány více aplikačními oblastmi, resp. jsou opakovaně používané jedním nebo více katalogy vzhledů jevů. Pomocí pojmů z pojmových slovníků vzhledů jevů lze konkrétně charakterizovat typy vzhledů jevů a typy jejich vlastností pro jejich dokumentaci v katalogu vzhledů jevů.

Pro založení a správu pojmových slovníků vzhledů jevů spravovaných jako registry byla přijata norma ČSN EN ISO ISO 19126:2012 Geografická informace – Pojmové slovníky vzhledů a registry, která specifikuje schéma těchto FCD.

Jak už bylo uvedeno v úvodu této podkapitoly 4.2, pojmové slovníky lze vést v rámci hierarchického registru, stejně jako katalogy vzhledů jevů. Pojmové slovníky jsou však vytvořeny bez spojení s konkrétním aplikačním schématem, a tak mohou výrazně napomoci k porozumění mezi různými producenty prostorové informace v rámci veřejné správy, ale i soukromoprávních subjektů. Základní náplní pojmových slovníků NIPI by měl být především stanovený konceptuální formalismus NIPI, který by měl vycházet a být provázán s konceptuálním formalismem definovaným v ISO 19103:2015.



Obrázek 20: Závislosti balíčků specifikovaných v ČSN EN ISO 19126:2012

Obdobně jako u katalogů vzhledů jevů patří k požadavkům různých komunit na tvorbu souborů geografické informace následující druhy pojmových slovníků:

- všeobecné – zakládající známou referenční množinu pojmů
- specializované – zavádějící v komunitě jedinečné pojmy, které nejsou v rozporu s pojmy v existujících referenčních pojmových slovnících vzhledů jevů
- mapování a/nebo korespondence mezi pojmy v různých pojmových slovnících vzhledů jevů, především těch, na nichž závisí významná datová úložiště geografické informace.

Množina pojmových slovníků vzhledů jevů a/nebo katalogů vzhledů může být spravovaná jako hierarchický registr (podle normy ČSN EN ISO 19135-1: 2016). Katalogy vzhledů jevů, které jsou odvozeny z pojmových slovníků vzhledů jevů, by měly být spravovány jako oddíl hierarchického registru, který obsahuje zdrojový pojmový slovník vzhledů jevů.

Pojmové slovníky vzhledů jevů mohou zahrnovat definice následujících pojmů:

1. vzhledů jevů
2. atributů vzhledů jevů
3. asociací vzhledů jevů
4. operací vzhledů jevů
5. jmenovitých hodnot i(jiné než číselné hodnoty) – jméno objektu, typu nebo kategorie.

Jako příklad pojmového slovníku vzhledů jevů lze uvést slovník The Defence Geospatial Feature Concept Dictionary (DGFCD). [26] Byl vypracovaný v roce 2017 pracovní skupinou DGIWG zabývající se komplexně geografickou informací v NATO. Tento slovník je strukturovanou sbírkou informací o vzhledech jevů (pojmy vzhledů jevů, pojmy atributů a přidružená data), jejíž schéma je v souladu s konceptuálním modelem specifikovaným v ISO 19126:2009. Blíže viz https://www.dgiwg.org/dgiwg-standards/206.

Mezi normami ISO/TC 211, které se zabývají vzhledy jevů, existuje řada závislostí – viz obrázek 21. Zmínit je třeba především:

- Specifikace datového produktu (ISO 19131:2007) obsahuje aplikační schéma (ISO 19109:2015) a buď obsahuje nebo se odkazuje na katalog vzhledů jevů (ISO 19110:2016). Podrobněji o specifikacích datových produktů viz čl. 4.1.3.

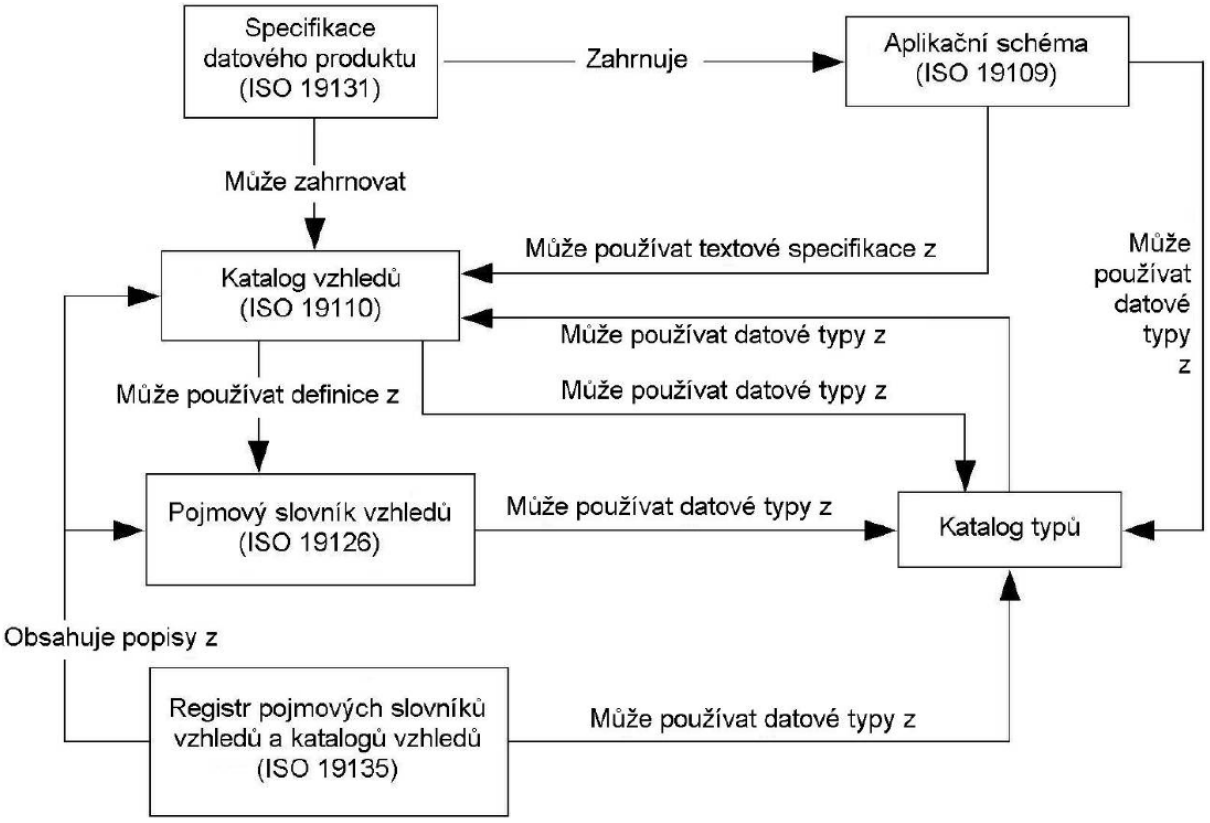
- Aplikační schéma může odkazovat na katalog vzhledů jevů jako na zdroj svých definic typů vzhledů jevů a typů jejich vlastností.

- Katalog vzhledů jevů může čerpat své definice pro typy vzhledů jevů a typy jejich vlastností z pojmového slovníku vzhledů jevů (ISO 19126:2009).

- Pojmové slovníky vzhledů jevů jsou uchovávány jako registry (ISO 19135-1:2015); katalogy vzhledů jevů mohou být také uchovávány jako registry, ale mohou být zahrnuty do dokumentů, jako jsou specifikace datových produktů.
- Katalogy a pojmové slovníky vzhledů jevů mohou být začleněny do systémů slovníků, které se odkazují jeden na druhý.
- Datové typy, které mohou být přiřazeny k typům atributů vzhledů jevů nebo zahrnuty do příznaků operací vzhledů jevů, jsou specifikovány v ISO 19103:2015.

Podrobně jsou požadavky na realizaci vzájemných vztahů a začleňování katalogu vzhledů jevů do aplikačního schématu specifikovány v ISO 19131:2007/Amd 1:2011 Requirements relating to the inclusion of an application schema and feature catalogue and the treatment of coverages in an application schema.

Podrobné informace a návrh vlastního Katalogu tříd objektů NaSaPO lze nalézt v dokumentaci výsledku V03 „Rámcové vymezení objektů NaSaPO včetně referenčních atributů“.



Obrázek 21: Závislosti mezi normami ISO řady 19100 (převzato z ČSN EN ISO 19126:2012)

4.3 Kvalita

Kvalitu v nejširším významu tohoto slova je možno považovat za rozhodující aspekt garantovaných geoprostorových dat. Z mnoha jejích definic byla vybrána následující, která nejlépe vystihuje místo a funkce (role) kvality prostorových dat ve všech fázích jejich životního cyklu, to je ve všech procesech jejich navrhování, vývoje, produkce, distribuce i archivace a rušení (náhrady).

- **kvalita (quality)** – stupeň, po který množina podstatných charakteristik splňuje požadavky

[ISO 9000:2015]

Jednoduchou interpretací této obecné definice shledáváme, že kvalitou prostorových dat je dána míra uspokojení požadavků jejich uživatele.

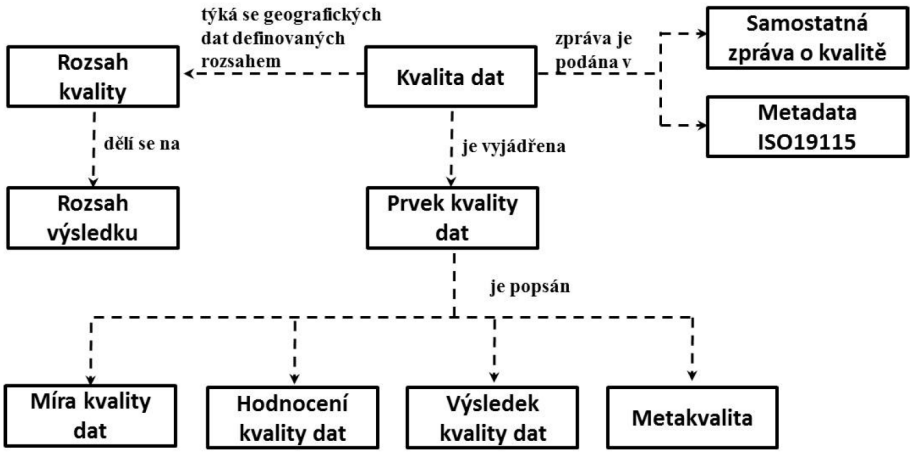
Při výběru množiny dat je pro uživatele nezbytné porovnat kvalitu dostupných množin dat pro stanovení, která z nich splňuje jeho požadavky nejlépe. Účelem kvality je usnadnit porovnání a volbu množin dat pro potřeby uživatele, a tím podpořit sdílení a využití množin dat. Zavedení nástrojů a procesů definování, zajištění, vyhodnocení a prezentace kvality dat v rámci NIPI je nezbytné pro dosažení tzv. garantovaných dat.

Problematiku kvality dat řeší ISO 19157:2013 Geographic information – Data quality, která je do soustavy ČSN převzata překladem (viz ČSN EN ISO 19157:2015 Geografická informace – Kvalita dat). Tato norma podrobně definuje komponenty kvality dat, prvky kvality dat, pojmy, postupy hodnocení a vykazování kvality, řízení kvality atd.

Předmětem této mezinárodní normy je poskytnout principy pro popis kvality geografických dat a pojmy pro nakládání s informací o kvalitě geografických dat, a dále konzistentní a normovaný způsob pro určení a zápis informace o kvalitě množiny dat. Cílem je poskytnout návod hodnotících postupů kvantitativní informace o kvalitě geografických dat.

[ČSN EN ISO 19157:2015]

Následující obrázek 22 představuje konceptuální model kvality geografických dat odpovídající normě ISO 19157:2013.



Obrázek 22: Konceptuální model kvality geografických dat (převzato z ČSN EN ISO 19157:2015)

Kvalita dat kvantitativně vyjadřuje, jak jsou splněna kritéria specifikace datového produktu (viz ISO 19131:2007) nebo požadavky uživatelů, resp. vyjadřuje, do jaké míry množina dat odráží daný obor úvah. Pokud specifikace datového produktu neexistuje nebo neobsahuje požadované míry kvality, může být hodnocení provedeno subjektivně, nekvantitativně jako popis výsledného produktu. V případě vytvoření a změny specifikace datového produktu nebo požadavků uživatelů musí být hodnocení kvality provedeno opakovaně.

4.3.1 Požadavky na kvalitu dat

Ke splnění velmi významné role kvality prostorových dat je nutné vytvořit podmínky, aby byly standardizovanými postupy

a) stanoveny jednoznačné a komplexní požadavky na kvalitu dat a tyto definovány podle ISO 19131:2007 Geografická informace – Specifikace datového produktu,
b) stanovené úrovně kvality zajištěny ve všech fázích a etapách procesu výroby nebo akvizice dat a jejich distribuce uživatelům (zavedením systému řízení jakosti podle norem ISO řady 9000 u producenta dat, případně i podle norem AQAP řady 2100, aplikací normy ISO 19158:2012, vypracováním plánů kvality apod.),
c) všechny prvky kvality objektivně hodnoceny v rámci vytvořeného systému hodnocení dosažené kvality datového produktu podle normy ISO/TS 19157:2013 (vytvořením plánů testování a jejich exaktním vyhodnocením po jednotlivých prvcích kvality, plněním plánů přejímky a certifikace dat, atd.),
d) zpřístupněny uživatelům informace o kvalitě publikováním výsledků hodnocení kvality v metadatech dle ISO 19115-1:2014,
e) pravidelně revidována a činěna aktivní opatření k zachování a případně zvýšení kvality dat po celou dobu jejich životnosti.

Primární roli v otázkách kvality dat hrají specifikace datového produktu. Tyto specifikace musí obsahovat především ucelený soubor požadavků, jejichž splnění je předpokladem souladu produktu s ISO 19157:2013. SDP musí stanovit přijatelné úrovně shody a odpovídající míry kvality dat specifikované v této normě, přičemž konkrétní úrovně musí být stanoveny pro všechny prvky a podprvky kvality dat z ISO 19157:2013. Pokud některý prvek kvality není aplikovatelný, je nutné tuto skutečnost písemně zaznamenat. K prvkům kvality uvedeným v této normě lze libovolně přidávat další a stávající prvky je možné rozvinout a upřesnit. Seznamy prvků včetně jejich rozšíření je třeba vytvářet pro každou oblast dat (scope) a popisovat podle ISO 19115:2014.

Při stanovování úrovní shody kvality v SDP by mělo být bráno v úvahu:

- že na různé části datové sady (oblasti kvality dat) mohou být aplikovány odlišné metody hodnocení kvality,
- že pro tentýž prvek kvality dat může být s různými intervaly spolehlivosti a různými mírami hodnocení kvality dosaženo různých výsledků hodnocení,

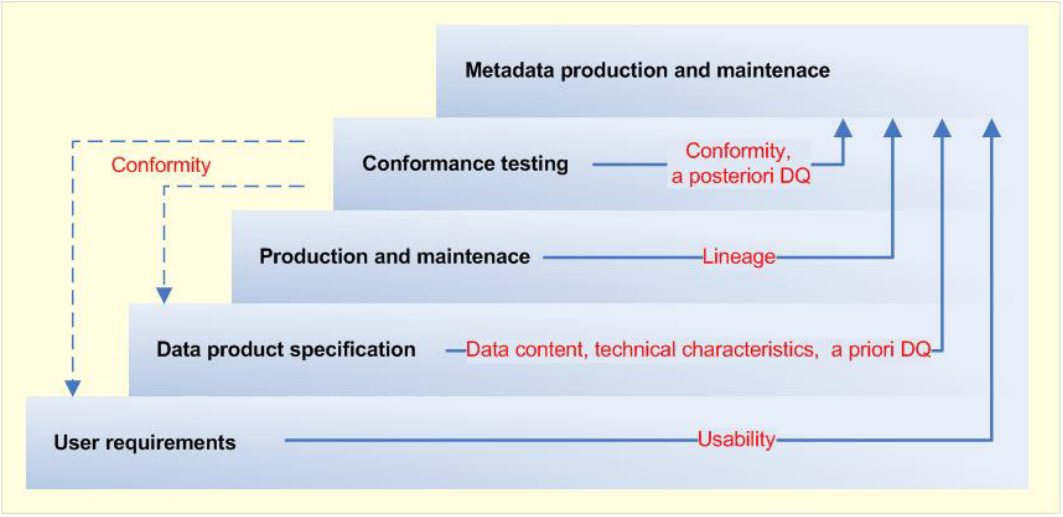
- že úrovně shody kvality mohou být v datové sadě různé pro různé vzhledy jevů, např. polohová přesnost vzhledů jevů s neurčitými hranicemi je obvykle mnohem nižší, než polohová přesnost liniových a dobře vymezených vzhledů jevů.

Klasický proces výroby prostorových dat je poměrně rigidní a sestává z následujících fází:

1. Studium požadavků uživatelů
2. Vývoj specifikace datového produktu
3. Výroba
4. Testování shody
5. Publikování metadat

K zajištění požadované kvality výsledného datového produktu, stanovené ve specifikaci datového produktu, musí být každá tato fáze doprovázena procedurami a opatřeními zajišťujícími dosažení požadované úrovně shody kvality přičemž každý krok musí být zdokumentován v metadatech. Z toho vyplývá, že data o kvalitě Data Quality (DQ) v metadatech budou zahrnovat informace o použitelnosti dat, technických charakteristikách dat, aspektech hodnocení kvality, shodě a distribuci dat. V ideálním případě jsou metadata tvořena akumulováním jejich přírůstků z každého postupného kroku, přičemž výstupy předchozího kroku jsou vstupem do následujícího (viz obr. 23).

Specifikování požadavků a zajištění potřebné kvality prostorových dat je mimořádně významná součást NIPI, je jedním ze základů systému garance za prostorová data a je potřebné alokovat pro tuto komplexní problematiku odpovídající kapacity.

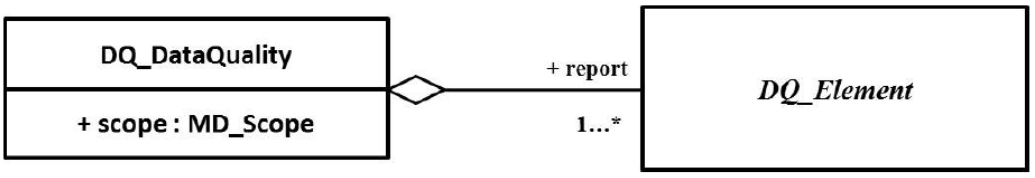


Obrázek 23: Akumulace přírůstků metadat o kvalitě v procesu výroby informací (převzato z [25])

4.3.2 Prvky kvality dat

Norma ISO 19157:2013 konstatuje, že kvantitativní prvky kvality dat mohou mít přidruženou meta kvalitu, která popisuje kvalitu výsledku hodnocení kvality dat pomocí normou stanovených charakteristik.

Kvalita dat může být vztažena na vybranou množinu dat, která sdílí společné charakteristiky. Pro popis kvality dat jsou použity různé jednotky kvality dat, které jsou kombinací rozsahu kvality (MD_Scope) a prvků kvality (DQ_Element).



Obrázek 24: Jednotka kvality dat (převzato z ISO 19157:2013)

Prvky kvality a jejich parametry jsou nejčastěji stanovovány a tedy také hodnoceny pro následující oblasti kvality:

a) řada množin dat;
b) množina dat;
c) podmnožina dat definovaná jednou nebo více následujícími charakteristikami:
    1) typy položek (množiny typů vzhledů jevů, atributy vzhledů jevů, operace na vzhledech jevů a vztahy mezi vzhledy jevů);
    2) specifické položky (množiny instancí vzhledů jevů, hodnoty atributů nebo instance vztahů mezi vzhledy jevů);
    3) geografická rozloha;
    4) časový rozměr (časový rámec odkazu a přesnost časového rámce).

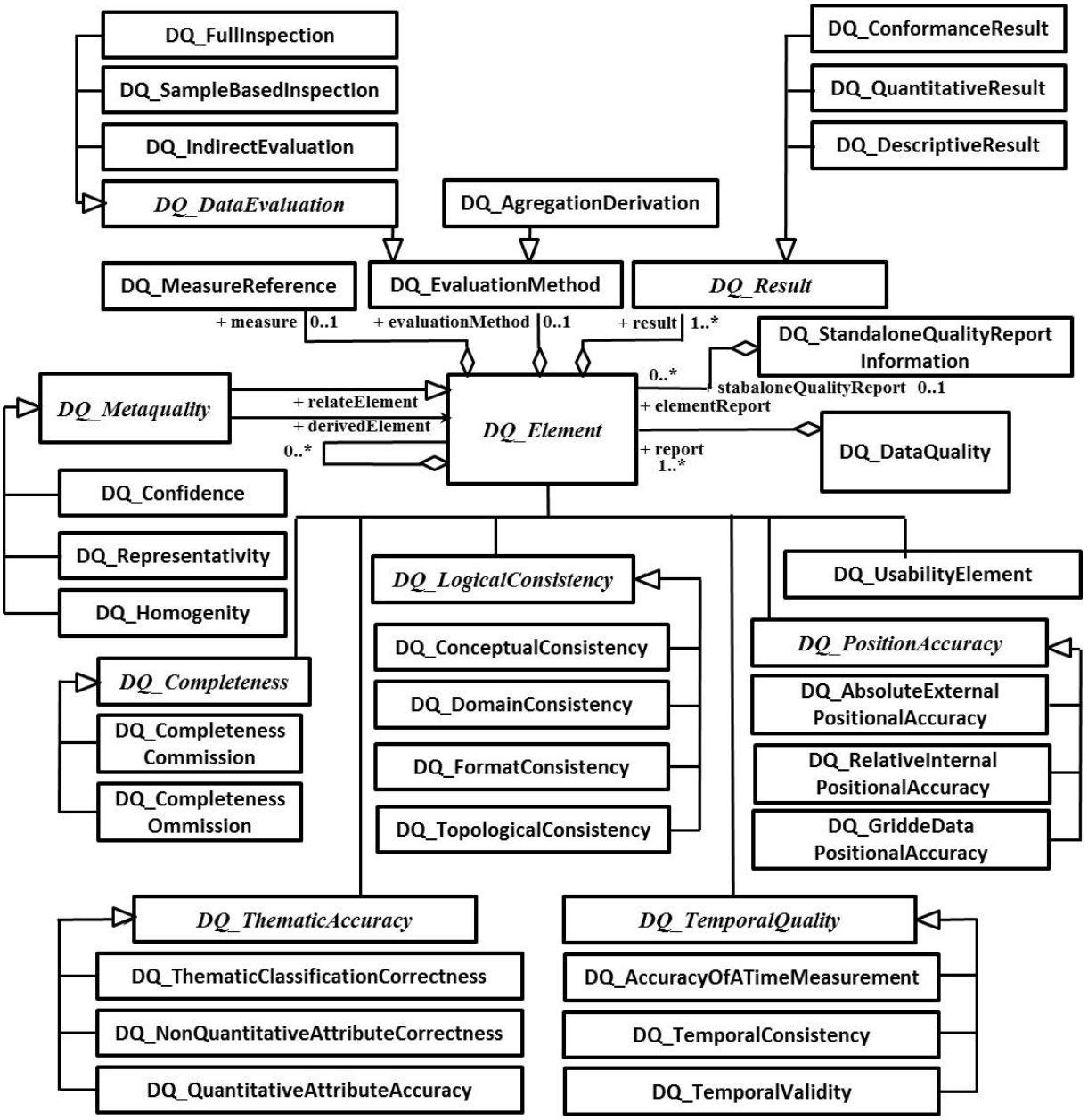
[ISO 19157:2013]

Jejich úplný seznam je však podstatně delší a je specifikován v MD_Scope v ISO 19115-1:2014. Pro oblasti kvality a oblasti metadat je používán společný seznam kódů MD_ScopeCode (viz obr. 16). Pro specifikování a pro vyhodnocování jednotlivých prvků kvality lze efektivně uplatňovat tatáž pravidla a postupy agregace jako pro metadata, jejichž jsou údaje o kvalitě velmi významnou součástí.

Prvky kvality dat (DQ_Element) jsou organizovány do skupin:

- úplnost (DQ_Completeness)
- logická bezespornost (DQ_LogicalConsistency)
- polohová přesnost (DQ_PossitionAccuracy)
- tematická přesnost (DQ_ThematicAccuracy)
- časová kvalita (DQ_TemporalQuality)
- prvek použitelnosti (DQ_UsabilityElement)

Ucelený přehled všech komponent kvality a jejich vzájemných vztahů představuje UML schéma na obr. 24.



Obrázek 24: Přehled komponent kvality dat (převzato z ISO 19157:2013)

4.3.3 Hodnocení kvality

Základní principy a procedury řízení kvality prostorových dat a služeb nad těmito daty jsou určeny platnými a obecně široce využívanými normami patřícími především do řady norem ISO 9000 a normami pro vytváření a provozování informačních systémů řady ISO 27001. Proces hodnocení kvality sestává z několika na sebe navazujících technologických kroků v různých fázích životního cyklu datového produktu, přičemž v každé fázi má jiné cíle. Těmito fázemi jsou specifikace, produkce, dodání, použití a aktualizace.

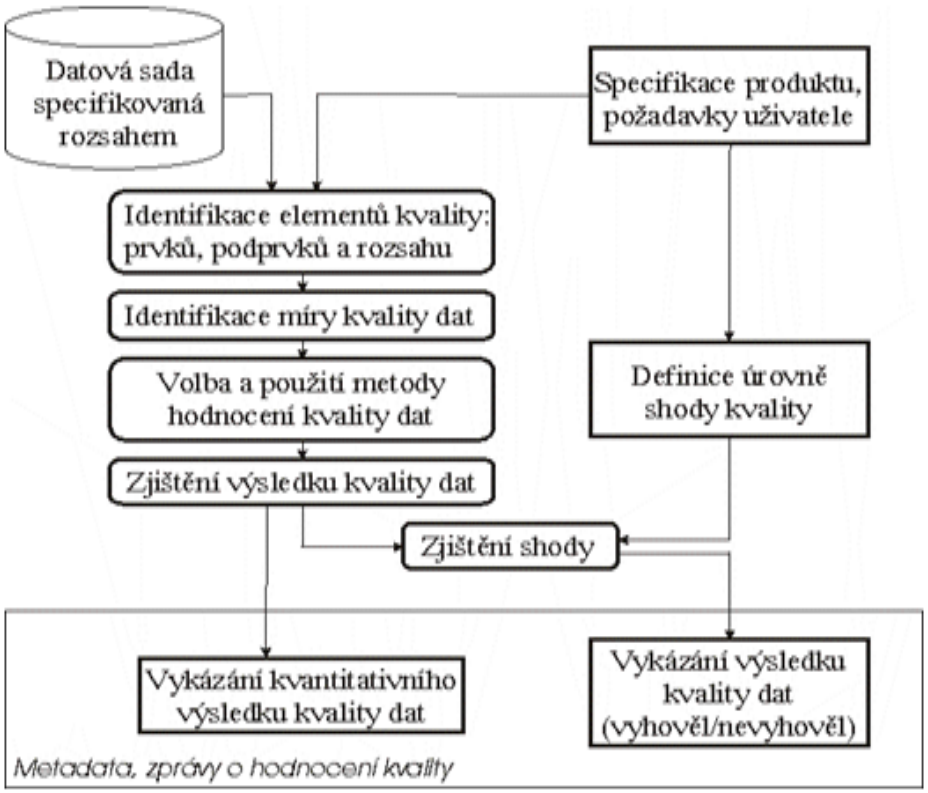
Specifické otázky kvality vztahující se ke geoprostorové informaci jsou konkretizovány ve skupině norem vyvinutých ISO TC 211. Aby bylo dosaženo komparativních hodnocení zpráv o kvalitě různých datových produktů jsou v Příloze D normy ISO 19157:2013 definovány normované míry kvality dat, které je touto normou důrazně doporučeno při hodnocení kvality prostorových dat používat.

4.3.3.1 Hodnocení kvality dat

Vyhodnocení kvality dat je základním nezbytným popisem množiny dat nebo podmnožiny dat v metadatech, který kvantitativně vyjadřuje, do jaké míry jsou splněna kritéria ze specifikace datového produktu nebo z požadavků uživatelů. Z logických důvodů je výhodné, aby struktura požadavků na kvalitu prostorových dat ve specifikaci datového produktu byla stejná jako struktury údajů jejich standardizovaného vyhodnocení.

Posloupnost i obsahy kroků při hodnocení, které vedou k výsledku hodnocení dat, jsou zřejmé z obecného schématu na obr. 25. Obrázek byl převzat z výzkumné zprávy VÚGTK, v. v. i., [23]. Specialisté ústavu se již dlouhodobě věnují vývoji metodik pro stanovení, hodnocení a prezentaci výsledků hodnocení kvality geoprostorových dat podle standardů a norem ISO.

Pokud hodnocená geografická data jsou heterogenní s různou kvalitou pro různé oblasti (scopes), měly by být zkoušky aplikovány pro vhodné části dat.



Obrázek 25: Obecné schéma technologických kroků procesu hodnocení kvality (převzato z [23])

Hodnocení prvků kvality dat je popsáno pomocí:

- měr – aspektů hodnocení každého prvku kvality dat, které jsou vyjádřeny pomocí generických základních měr
- metod – postupů hodnocení kvality dat (přímých – úplná kontrola, kontroly vzorkováním; nepřímých – podle externích referenčních dat)
- výsledků – výstupy hodnocení kvality dat
- metakvality – informace popisující kvalitu kvality dat.

4.3.3.2 Hodnocení kvality služeb nad prostorovými daty

Pro hodnocení kvality služeb nad prostorovými daty jsou používána nejčastěji následující kritéria:

- dostupnost – pravděpodobnost že je webová služba připravená k okamžitému použití
- dosažitelnost – stupeň, do něhož je služba schopna obsloužit webový požadavek
- integrita – úroveň v jaké služba udržuje správnost prováděných transakcí
- výkonnost – průchodnost daná počty webových požadavků obsloužených službou za určitou časovou jednotku a odezva jako časový interval mezi vysláním požadavku a přijetím odpovědi;
- spolehlivost – úroveň udržování a kvality služeb. Mírou spolehlivosti je počet výpadků a závad za měsíc nebo rok. Spolehlivost může být dána také poměrem odeslaných a doručených zpráv mezi žadatelem a poskytovatelem služby;
- zajištění právních norem – míra shody služby s pravidly, zákony a platnými verzemi norem a standardů;
- bezpečnost – úrovně škálovatelných opatření k zachování mlčenlivosti, ochrany a šifrování dat, prokázání věrohodnosti účastníků apod.

Problematiku služeb nad prostorovými daty, včetně aspektů jejich kvality, řeší podrobně norma ČSN EN ISO 19119:2019 Geografická informace – Služby.

4.3.4 Formy výsledného hodnocení

Vyhodnocením každého prvku a podprvku kvality podle stanovených měr kvality jsou získány soubory kvantitativních a/nebo kvalitativních výsledků, které jsou dále zpracovány statistickými metodami s uplatněním kritérií stanovených v SDP (zpravidla prahových hodnot) a je posouzeno, zda bylo dosaženo shody úrovně kvality.

- **shoda úrovně kvality (conformance quality level)** – prahová hodnota nebo sada prahových hodnot pro výsledky kvality dat použitá pro určení, jak dobře bude množina dat odpovídat množině kritérií ze specifikace datového produktu nebo požadavkům uživatele.

[ČSN EN ISO 19157:2015]

Ke komplexnímu vyhodnocení kvality datového produktu je zpravidla uskutečněna agregace představující kombinaci existujících výsledků hodnocení kvality dat založené na různých prvcích kvality dat nebo na různých rozsazích kvality dat.

Jednotlivým dílčím výsledkům lze v SDP, nebo zájemcem o data, přisoudit různou váhu podle významu hodnoceného prvku kvality pro zamýšlené použití datového produktu. Soubory těchto dílčích kvantitativních výsledků, případně jejich vážený průměr, jsou porovnány s požadovanou úrovní shody a výsledek tohoto porovnání je spolu s výsledky jednotlivých testů podkladem k rozhodování garanta o vydávání osvědčení o shodě tohoto datového produktu.

Stejný postup hodnocení datového produktu s vlastními kritérii, limitními hodnotami a požadovanou úrovní shody kvality může uplatňovat i zájemce o využití datového souboru před jeho akvizicí.

4.3.5 Publikování výsledků hodnocení kvality

Hodnocení kvality datového produktu spočívá v postupném provádění testů stanovených v plánech kvality. Tyto plány jsou vyhotoveny před zahájením výroby a jejich obsah závisí na zvolené technologii a metodě sběru dat a konkrétním typu dat. Pro každý výstup z jednotlivého technologického kroku jsou stanoveny míry kvality a úrovně shody, jejichž dosažení je podmínkou zahájení následujícího technologického kroku. Finální datový produkt je podroben výstupní kontrole (přejímce) přičemž je posuzováno dosažení shody s požadavky kvality stanovenými ve formě úrovně shody ve specifikaci datového produktu.

Postup produkce dat a provedené kontroly s dosaženým kvantitativními i kvalitativními výsledky kontrol musí být velmi podrobně zdokumentovány a publikovány, což je jedním z podstatných aspektů garantovaných dat.

Výsledky hodnocení kvality musí být prezentovány formou metadat ve shodě s kapitolou 7, kapitolou 10 a přílohou C, ISO 19115-1:2014 a ISO 19115-2:2019. Pro poskytnutí více detailních informací o kvalitě dat, než která jsou součástí metadat, je doporučováno vytvořit navíc samostatnou zprávu o kvalitě. Její struktura je volná. Samostatná zpráva o kvalitě nesmí nahradit metadata, je jen jejich doplněním. Metadata by však měla poskytnout odkazy na tuto samostatnou zprávu o kvalitě, pokud je tato zpracována.

- **samostatná zpráva o kvalitě (standalone quality report)** – volný text, poskytující plnou podrobnou informaci o hodnocení kvality dat, výsledcích a použitých mírách

[ČSN EN ISO 19157:2015]

Specifickou a praktickou formou popisování výsledků hodnocení kvality prostorových dat je použití pokrytí (coverage). Poskytuje velmi názorný přehled o hodnotách kvality dat pro konkrétní prvky kvality v prostorově vymezených částech území a/nebo časových úsecích. Pravidla a postupy využívání této formy prezentace výsledků jsou uvedeny v 1. doplňku normy ISO 19157:2013/Amd 1:2018 Describing data quality using coverages.

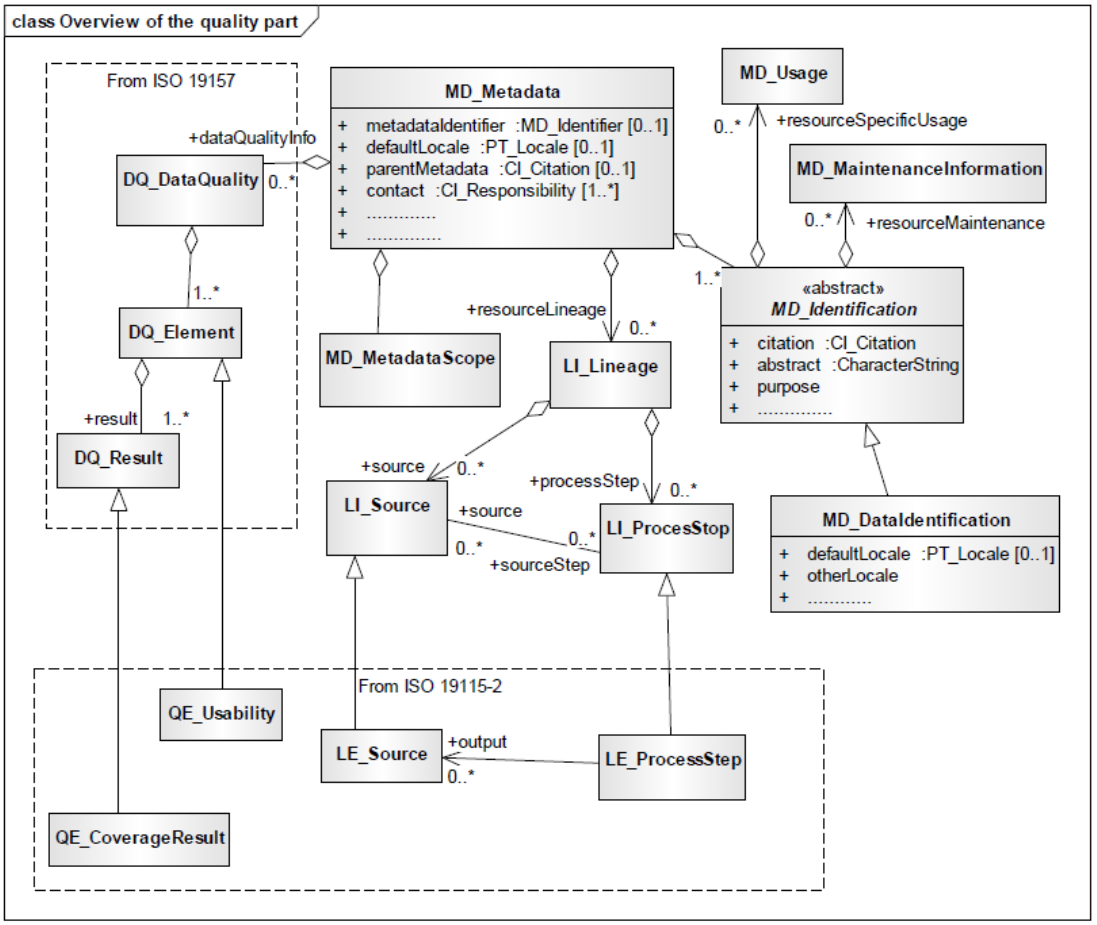
Problematika kvality prostorových dat je velmi komplexně a detailně zpracována v normě ČSN EN ISO 19157:2015, která detailně popisuje veškeré hlavní komponenty a míry kvality prostorových dat, postupy a metody hodnocení kvality a publikování výsledků tohoto hodnocení. Norma obsahuje množství praktických příkladů a je rozhodně jednou ze stěžejních a nejčastěji aplikovaných norem ISO geografické informace. Otázkám kvality geografické informace je v rámci ISO/TC 211 věnována mimořádná pozornost. V současné době je norma ISO 19157:2013 revidována z její nové struktury byla v roce 2016 již vydána technická specifikace ISO/TS 19157-2 Geografická informace – Data quality – Part 2: XML Schema implementation a před vydáním je nová část normy ISO 19157-1 Geographic information – Data quality – Part 1: General requirements. Tuto klíčovou normu i technickou specifikaci ISO 19158:2012 Geographic information – Quality assurance of data supply je nutné důsledně využívat řídícími i odbornými orgány, stejně jako využívat výsledky společných projektů řešených v JRC Ispra, v rámci EuroGeographics – Q KEN (Quality Knowledge Exchange Network) a dalších odborných institucí, iniciativ a fór, a to ve všech fázích životního cyklu prostorové informace.

4.4 Implementace metadat

Soubor norem ISO 19115 se komplexně zabývá metadaty, včetně řešení jejich implementace a způsobů publikování. Neřeší však otázky vytváření a správy jejich katalogů, registrů a informačních systémů, které umožňují splnit jejich primární účel – rychlé vyhledání a zhodnocení využitelnosti informačních zdrojů pro uživatelův záměr.

Implementace metadat představující konverzi UML schémat do kódování XML je popsána v čl. 4.1.2 Výměnný formát metadat a detailní postupy jsou uvedeny v normách ISO/TS 19139-1:2019, ISO 19115-2:2019 a ISO 19118:2011. Několik názorných praktických příkladů implementace metadat o datových sadách, včetně jejich vícejazyčného provedení, je uvedeno v Příloze D „Implementation examples“ a v Příloze E „Metadata implementation“ k normě ISO 19115-1:2014.

Implementace metadat je dále ilustrována na skupině metadat o kvalitě. Jejich na kódování nezávislý tvar je vyjádřen v UML schématu uvedeném na obr. 26. Aby nepoužíval každý uživatel při implementaci normy jinou verzi, byla vypracována a je používána technická specifikace poskytující univerzální implementaci normy ISO 19115-1:2014 pro zpracování UML schémat a pro XML schéma, které odpovídá pravidlům popsaným v ISO 19118:2018 Geografická informace – Kódování.



Obrázek 26: UML schéma balíčku *Kvalita* v metadatech (převzato z [24])

V nich jsou balíčky s prvky kvality pojmenovány *Lineage information* a *Identification information*, které společně vytváří UML třídu *MD_Metadata*.

Pro každý balíček uvedený v ISO 19115-1:2014 je zpracován soubor XML schématu téhož jména jako má balíček, např. soubor dataQuality.xsd obsahuje implementaci informací balíčku *Data quality*.

Na stránkách https://schemas.isotc211.org/schemas/19115/ je možné nalézt XML schémata, která jsou sdružená s příslušnými jmennými prostory a použít je pro kódování UML modelu definovaného v této normě.

4.4.1 Rozšíření metadat do komunitního profilu

ISO 19115-1:2014 v příloze B poskytuje normovaná metadata pro různorodé zdroje prostorové informace. Pokud se po podrobném prozkoumání těchto metadat prokáže, že existují specifické potřeby metadat, které není možné pokrýt těmito generickými metadaty, je možné definovat profil rozšíření metadat. Detailní návod rozšíření metadat je uveden v ISO 19115-1:2014 příloha C a dále v ISO 19106:2004 definující tvorbu profilů. Obdobně můžeme postupovat i u katalogu vzhledu jevů (ISO 19110:2016), respektive Katalogu tříd objektů NaSaPO a kvality prostorových dat (ISO 19157:2013).

Povolené typy rozšíření metadat:

1) přidání nového metadatového balíčku;
2) vytvoření nového metadatového seznamu kódů pro nahrazení domény existujícího metadatového prvku, který má zapsán „volný text“ jako svou doménovou hodnotu;
3) vytvoření nového prvku metadatového seznamu kódů (rozšiřující se seznam kódů);
4) přidání nového metadatového prvku;
5) přidání nové metadatové třídy;
6) zavedení přísnějších povinností k existujícímu metadatovému prvku;
7) zavedení mnohem vyhrazenější domény pro existující metadatový prvek.

[ISO 19115-1:2014, příloha C]

Společenství uživatelů tvořící komunitu, v našem případě se jedná o účastníky NIPI ČR, mohou například definovat některé prvky metadat jako povinné, nebo doplnit prvky metadat jako „cena“ množiny dat, „agenda“ na jejímž základě je zdroj pořízen, „životní situace“ podporované zdrojem, „stav“ zdroje v rámci systému, identifikátor „IČO“, atd. Pro všechna rozšíření musí být definováno konkrétní jméno, definice, datový typ, doména hodnot, povinnost, maximální počet výskytů a vztahy pro určení struktury (schématu) vyjádřené v UML.

Pro zajištění interoperability musí být přidaná rozšíření (profil) publikována tak, aby byla známá společenství uživatelů i mimo něj. V rámci NIPI by tak mělo být zajištěno publikování „národního profilu metadat zdrojů prostorových informací“ ve spolupráci s ÚNMZ, například formou ČSN (viz čl. 3.2.3).

4.5 Garance za datový produkt

Při řešení problematiky garantovaných dat a garantovaných služeb budou prakticky využity poznatky ze všech předchozích podkapitol. Zavedení definic těchto frekventovaných pojmů a jejich vysvětlení je nejen pro účely projektu NaSaPO, ale i pro praxi v celé NIPI velmi významné a má mimořádné legislativní, organizační a administrativní dopady. Základními používanými pojmy jsou:

- **garant** – osoba nebo organizace, která ručí za kvalitu
- **garantovaná data**[3] – data splňující stanovené požadavky kvality a tato skutečnost je formálně potvrzena úředně stanovenou osobou nebo organizací
- **validace** – proces posuzování kvality datových produktů odvozených z výstupů systému nezávislými prostředky

[ISO 19101-2:2018]

- **certifikace** – podle § 10 zákona č. 22/1997 Sb., o technických požadavcích na výrobky, je to činnost
  a) autorizované osoby prováděná v rozsahu vymezeném technickým předpisem, nebo
  b) k tomu akreditované osoby prováděná na žádost výrobce, dovozce nebo jiné osoby, při níž se vydáním certifikátu osvědčí, že výrobek nebo činnosti související s jeho výrobou, popřípadě s jeho opakovaným použitím jsou v souladu s technickými požadavky v certifikátu uvedenými.

[Zákon č. 22/1997 Sb.]

Jedním z nejdůležitějších předpokladů pro začlenění prostorových dat do NaSaPO je, aby byla tato prostorová data garantovaná. Musí tedy být přezkoumána (validována) stanovenou autorizovanou nebo k tomu akreditovanou právnickou osobou (garantem), zda splňují všechny všeobecné požadavky na data NaSaPO a především požadavky na kvalitu specifikované pro každý datový produkt v jeho SDP. Pouze po splnění této podmínky může stát tato data převzít a poskytovat za ně garanci. Tento proces přezkoumání a potvrzení kladných výsledků vydáním certifikátu o shodě je obdobou procesu státního ověřování jakosti výrobků, které ve smyslu platných zákonů vzbuzují oprávněný zájem společnosti (tj. výrobky, které by mohly ve zvýšené míře ohrozit zdraví nebo bezpečnost osob, majetek nebo životní prostředí, popřípadě jiný veřejný zájem, dále jen „oprávněný zájem“ – viz zákon č. 22/1997 Sb.). Lze tedy předpokládat, že autorizovaní pracovníci ÚřOSK SOJ budou často tímto certifikačním orgánem, a to zejména v případech přezkoumávání prostorových dat NaSaPO využívaných pro fungování strategické nebo kritické infrastruktury. V běžných případech budou prostorová data NaSaPO z určité domény garantována kompetentní akreditovanou certifikační organizací nebo institucí, kterou k tomu určí gestor příslušné odborné oblasti. Jako potvrzení úspěšného certifikačního procesu (přezkoumání dosažené shody kvality a splnění všech obecných požadavků na data NaSaPO) bude garantem vystaveno časově vymezené osvědčení o shodě datového produktu (certifikační protokol) platné pro vyjmenované oblasti a po stanovenou dobu. Údaje o uděleném osvědčení budou spolu s příslušnými dalšími údaji o účastnících, o certifikačním orgánu, rozsahu a době platnosti zapsány do metadat datového produktu a publikovány. Dosažené úrovně kvality garantovaných dat umožňují, aby byla data využívána jako referenční při vyhodnocování kvality jiných zdrojů prostorových dat externími metodami hodnocení.

[3] Pozor na rozdílnost od definice garantovaných prostorových dat (*prostorová data s příslušným certifikátem kvality, nad kterými veřejná správa provádí transparentní rozhodovací procesy s následnou právní odpovědností*), která je zveřejněna v rámci [2].

4.5.1 Všeobecné požadavky na garantovaná prostorová data

V souhrnu jsou na garantovaná prostorová data NaSaPO kladeny následující požadavky:

- struktura a syntaxe prostorových dat a metadat je validována, resp. je ověřen jejich formát a úplnost,
- prostorová data jsou logicky bezesporná,
    - data mají vyplněná metadata včetně údajů o katalogu vzhledu jevů (Katalogu tříd objektů NaSaPO) a údajů o kvalitě.
    - pro popis dat (metadat) je použit stanovený konceptuální formalismus zajišťující jednoznačné chápání významu dat,
- prostorová data mají stanoveného gestora a garanta (povinná položka metadat),
- prostorová data jsou uchovávána (archivována) způsobem, který umožní jejich zpětnou rekonstrukci do stavu jaký byl ve stanoveném okamžiku (historizace dat),
- prostorová data jsou pořizována stanovenými metodami sběru dat,
- prostorová data splňují požadavky kvality stanovené ve specifikaci datového produktu pro konkrétní typ dat, resp. pro konkrétní datové série, datové sady nebo typy (třídy) vzhledů jevů. Mezi základní požadavky kvality prostorových dat patří kromě povinných prvků úplnost a logická bezespornost i polohová přesnost, tematická přesnost a časová kvalita.

Na sběr prostorových dat NaSaPO je nutné nahlížet jako na vybrané činnosti, jejichž důsledkem je kvalita rozhodovacích procesů veřejné správy. Proto jejich ověřování může provádět pouze fyzická osoba, která získala oprávnění k jejich výkonu podle zvláštního právního předpisu[4]. Podle zákona č. 200/1994 Sb., o zeměměřictví, v platném znění a jeho prováděcí vyhlášky č. 31/1995 Sb. ověření údajů o prostorové poloze staveb, majetkoprávních vztahů v území, průběhu podzemní technické infrastruktury nebo vedení digitální technické mapy je vyžadováno úředně oprávněným zeměměřickým inženýrem (ÚOZI). Obdobně je nutné nahlížet na budoucí přicházející výzvy s prostorovými daty jako je datové modelování při odvozování objektů NaSaPO nižší úrovně podrobnosti, informační modelování staveb a vystavěné prostředí, kde jsou garantovaná data základním předpokladem.

Obdobně je nutné nahlížet na postavení a funkce referenčních dat dle ustanoveními zákona č. 111/2009 Sb., o základních registrech.

Vymezení pojmu garantovaná data je věnována pozornost také ve schváleném záměru zákona o NIPI, kde lze reálně očekávat, že tato problematika bude součástí komplexně pojatého zákona o NIPI.

[4] Zákon č. 360/1992 Sb., o výkonu povolání autorizovaných architektů a o výkonu povolání autorizovaných inženýrů a techniků činných ve výstavbě, ve znění pozdějších předpisů, nebo zákon č. 200/1994 Sb., o zeměměřictví a o změně a doplnění některých zákonů souvisejících s jeho zavedením, ve znění pozdějších předpisů.
Zákon č. 200/1994 Sb., o zeměměřictví a o změně a doplnění některých zákonů souvisejících s jeho zavedením, ve znění pozdějších předpisů

4.5.2 Certifikace prostorových dat

Zásady, povinnosti a postupy poskytování garance za prostorová data se principiálně nijak neliší od záruk poskytnutých podle zákona č. 22/1997 Sb., o technických požadavcích na výrobky, které se vztahují na libovolný produkt vázaný oprávněným zájmem a jsou platné i pro dokumenty a data bez rozlišování, zda jsou v analogové nebo digitální formě. Procesy řízení jejich kvality jsou stanoveny v ISO normách řady 9000 a v českých obranných standardech (AQAP řady 2100), které řeší zajištění kvality produktů v rámci vládních zakázek, včetně digitálních prostorových dat).

Působnost akreditačních a autorizačních orgánů a proces certifikace zastřešuje výše zmíněný zákon č. 22/1997. Sb. Práva a povinnosti zúčastněných stran, principy a závazné postupy při provádění auditů pro certifikaci systému řízení kvality (QMS) nebo konkrétního produktu jsou uvedeny v ISO 19011:2018 Guidelines for auditing management systems a ve směrnicích certifikačních institucí, např. ČSJ, ÚřOSK SOJ, LRQA a mnoha dalších.

4.5.3 Abstraktní sestava zkoušek kvality dat

Vlastním obsahem procesu certifikace je v principu realizace Abstraktní sestavy zkoušek v rozsahu, který podrobně specifikuje norma ČSN EN ISO 19157:2015 – Příloha A (normativní). Cílem je detailně posoudit soulad hodnocení kvality s uvedenou normou a správnost prezentované kvality prostorových dat provedených jejich producentem. Certifikační autorita ji uskuteční na základě porovnání SDP, dokumentace použité technologie a výsledného datového produktu, včetně metadat, s detailní dokumentací QMS, provedeného hodnocení a dosažených výsledků.

Abstraktní sestavu zkoušek identifikuje pět skupin zkoušek, všechny jsou svým charakterem základními zkouškami.

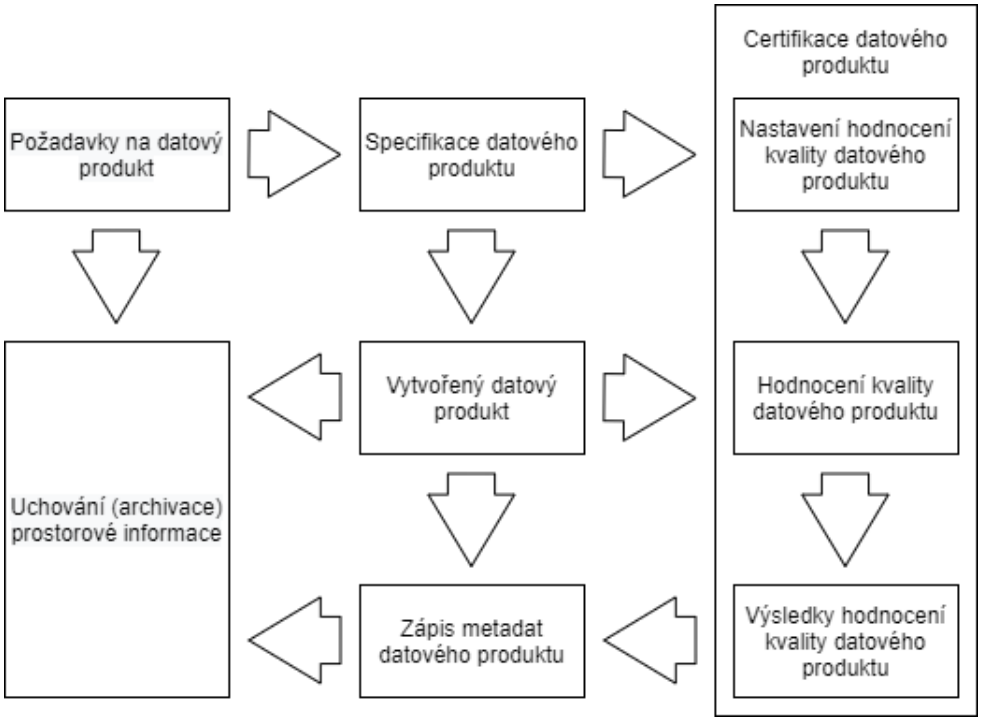
A.1 Proces hodnocení kvality

Účelem zkoušky je validovat proces hodnocení kvality dat, tzn. ověřit, zda proces hodnocení kvality zahrnuje všechny kroky specifikované v ČSN EN ISO 19157:2015 v čl. 9.1.3. To zahrnuje:

1) Identifikují se výroky specifikace datového produktu nebo požadavky uživatele relevantní ke kvalitě dat a použijí se pro identifikaci použitelných prvků kvality dat a jejich vhodného rozsahu. Poté se porovnají použitelné prvky kvality dat s prvky kvality dat hodnocenými tak, aby bylo zaručeno, že byly identifikovány všechny použitelné prvky kvality dat a ve vhodném rozsahu hodnoceny.
2) Ověří se, zda míra kvality dat, aplikovaná pro každé hodnocení kvality dat, je vhodná s ohledem na výrok specifikace datového produktu nebo na požadavky uživatelů.
3) Ověří se, zda postup hodnocení kvality dat, aplikovaný pro každé hodnocení kvality dat, je vhodný s ohledem na výrok specifikace datového produktu nebo požadavky uživatelů.

A.2 Metadata kvality dat

Přezkouší se, zda jsou metadata kvality dat modelována podle modelů UML a datového slovníku. Metadata musí obsahovat vhodné komponenty kvality dat a pro každou z nich se musí řídit existujícími pravidly.



Obrázek 27: Proces certifikace a archivace garantovaného datového produktu

A.3 Shoda metadat

Přezkouší se, zda jsou metadata o kvalitě dat vykazována ve shodě s ISO 19115-1:2014 a ISO 19115-2:2009. K ověření jsou realizovány abstraktní zkoušky stanovené v Příloze A k ISO 19115-1:2014.

A.4 Samostatná zpráva o kvalitě

Přezkouší se, zda samostatná zpráva o kvalitě obsahuje členění na všechny relevantní aspekty kvality a zda se popis všech komponent kvality dat řídí pravidly definovanými v ČSN EN ISO 19157:2015.

A.5 Míry kvality dat

Přezkouší se, zda je míra kvality dat strukturovaně a sémanticky dobře definovaná a zda jsou použité míry kvality dat popsané a modelované podle modelu UML a datového slovníku.

4.5.4 Akreditační a certifikační instituce pro oblasti ICT

Pro účely nestranného, objektivního a nezávislého posouzení způsobilosti státních i soukromých subjektů udělovat akreditaci pro všechny oblasti akreditace byl vládou ČR založen Český institut pro akreditaci, o. p. s., (ČIA), který ve vztahu k EU a zahraničí plní roli národního akreditačního orgánu ve smyslu nařízení Evropského parlamentu a Rady (ES) č. 765/2008, kterým se stanoví požadavky na akreditaci a dozor nad trhem týkající se uvádění výrobků na trh a kterým se zrušuje nařízení (EHS) č. 339/93. Pro oblast systémů managementu kvality QMS a produktů ICT tento institut akreditoval mj. také jednu z nejlépe etablovaných a dlouhodobě mezinárodně uznávaných společností v ČR, kterou je Česká společnost pro jakost, z. s., (ČSJ). Z širokého spektra odborných domén, pro které má ČSJ uděleno oprávnění certifikovat jsou „Systémy managementu IT služeb podle ISO/IEC 20000-1:2018 Information technology – Service management –

Part 1: Service management system requirements“. Lze očekávat, že držení tohoto certifikátu bude jedním z významných požadavků na firmy v rámci výběrových řízení pro poskytování garantované kvality dodávek IT služeb pro NaSaPO k doložení jejich způsobilosti. Dalšími relevantními oblastmi certifikace producentů dat pro NaSaPO jsou „Systém managementu bezpečnosti informací podle ISO/IEC 27001:2013 Information technology – Security techniques – Information security management systems – Requirements“ a další. Bližší informace lze nalézt např. na webových stránkách ČSJ https://www.csq.cz/informacni-a-komunikacni-technologie/.

4.6 Archivace prostorových dat

Aby mohl být datový produkt patřící do NaSaPO používán k rozhodovacím procesům ve státní a veřejné správě nebo pro potřeby veřejnosti je zcela nezbytné zajistit archivaci (uchování) datového produktu a jeho metadat tak, aby prostorová informace byla po stanovenou dobu plně obnovitelná. V případě požadavků, např. sporů, musí být vytvořeny podmínky pro zpětnou rekonstrukci dat pro zadaný okamžik, a to včetně informací o kvalitě a všech ostatních metadat. Uchováním prostorových dat se zabývá norma ISO 19165-1:2018 Geographic information – Preservation of digital data and metadata – Part 1: Fundamentals a ISO 19165-2:2020 Geographic information – Preservation of digital data and metadata – Part 2: Content specifications for Earth observation data and derived digital products.

5 Klasifikace datových modelů

Analýza klasifikace datových modelů se zaměřila na procesy klasifikace datových modelů podle procesů výběru témat. Teoretickým i praktickým otázkám generalizace geografické informace se již dlouhodobě věnuje řada odborných institucí především z akademické sféry a odborných komisí ICA. Pro analýzy a klasifikaci konceptuálních modelů generalizace a vícenásobné reprezentace byly využity především významné teoretické práce prof. Grünreicha, uskutečněné v rámci vývoje systému ATKIS [14] a souhrnný a hodnotící příspěvek L. T. Sarjakoski do publikace [17]. Z domácích odborných prací znamenala významný přínos pro tuto analýzu a klasifikaci datových modelů především disertační práce Ing. Mildorfa v rámci jeho doktorandského studia na ZČU Plzeň. [13].

Generalizace – *úplnost procesu odehrávajícího se při získávání nebo kartografické reprezentaci jevu, při němž – podmíněno měřítkem nebo tématem – nevýznamné detaily jsou opomenuty v geometrickém, konceptuálním a časovém aspektu, a významné zůstávají zachovány nebo převedeny do obecnějších konceptů* [15]

Generalizace je pak dále dělena na objektovou, modelovou a kartografickou:

1. **Objektová generalizace:** *Již při sběru původní topografické informace topografem nebo fotogrammetrem probíhají procesy výběru, konceptuálního a geometrického propojení, jakož i zjednodušení, což vyúsťuje do redukce primární informace. Tento proces, jež se nazývá „objektová generalizace“, vytváří primární model reálného světa (území), který svým obsahem odpovídá topografické základní mapě a označuje se jako „digitální model území (DMÚ).* [14]

   POZNÁMKA: Digitální model území (DMÚ) = Digital Landscape Model (DLM). Výsledkem objektové generalizace je primární digitální model území.

2. **Modelová generalizace:** proces generující vizualizace nebo prostorové báze dat nižší úrovně podrobnosti než je původní (zdrojová) báze dat, které odpovídají danému účelu nebo tématu, přičemž jsou původní charakteristiky prostorových informací zachovány. [28]

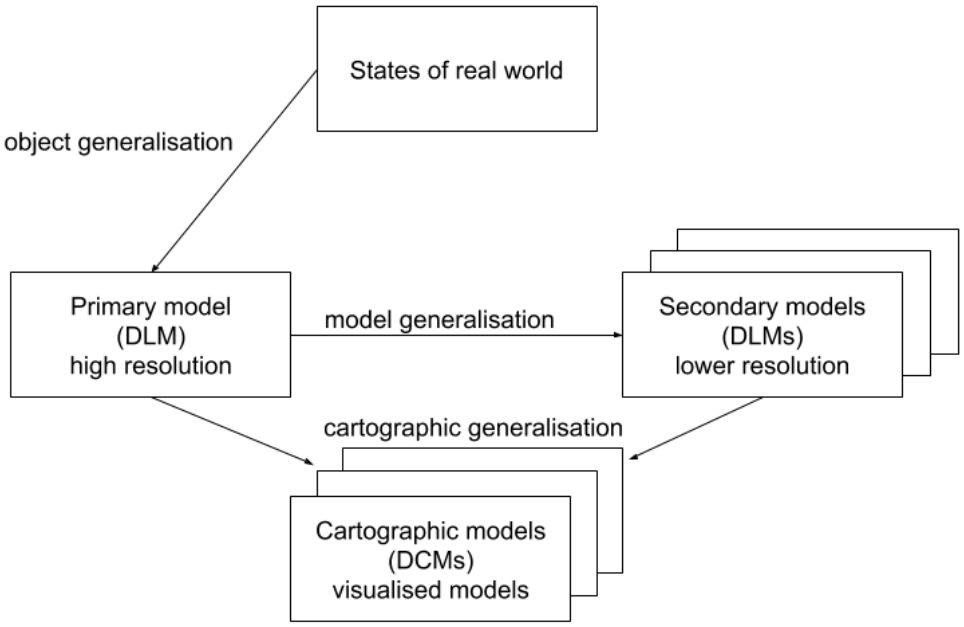
   POZNÁMKA: Výsledkem modelové generalizace je sekundární digitální model území (DMÚ).

3. **Kartografická generalizace:** výběr a zjednodušené reprezentování podrobnosti úměrné měřítku a/nebo účelu mapy. [29]

   POZNÁMKA: Výsledkem kartografické generalizace je terciální digitální kartografický model (DCM – Digital Cartographic model).

Z výše uvedených definic vyplývá **klasifikace digitálních modelů území** (viz obr. 28):

- **primární model** – Digitální model území (DMÚ)
- **sekundární model** – Digitální model území menší úrovně podrobnosti (DMÚ)
- **(terciální) kartografický model** – Digitální kartografický model (DKM)



Obrázek 28: Konceptuální model generalizace a vícenásobné reprezentace (převzato z [17])

POZNÁMKA: V odborných kruzích je obecně zažitý termín „základní model území“ pro označení primárního modelu a termín „odvozený model území“ pro označení sekundárního modelů.

Z uvedených definic vyplývá, že oborem zájmu projektu NaSaPO je objektová generalizace vztahující se ke sběru dat v terénu tvořící základní model území a modelová generalizace vztahující se k odvození datových modelů nižší úrovně podrobnosti. Kartografická generalizace je zaměřena na grafickou a vizuální stránku prostorové informace, řeší problematiku mapových značek, umístění textových polí a zvýraznění určitých aspektů podle účelu výsledného kartografického výstupu. Kartografická generalizace není předmětem projektu NaSaPO. Je řešena v rámci současného projektu čís. TIRDCUZK702 „Výzkum a vývoj programového aparátu pro generalizaci státního mapového díla“, který navazuje na předcházející projekt č. TB04CUZK001 „Výzkum a vývoj metod pro kartografickou generalizaci státního mapového díla středních měřítek“ z roku 2016, oba řešené s finanční podporou TA ČR“.

Pro případné hlubšího pochopení rozdílů ve významu pojmu „generalizace“ a pojmu „abstrakce“ popsaného v článku 3.3.1 lze doporučit např. články autorů Buttenfieldová a McMaster [27] s Sesterová [28].

5.1 Klasifikace digitálních modelů v rámci NIPI

Pro usnadnění orientace v klasifikaci digitálních modelů je potřebné nejdříve stručně představit Národní infrastrukturu pro prostorové informace (NIPI). Její níže uvedené definice byly představeny i v rámci činnosti Realizačního výkonného týmu pro legislativu v oblasti prostorových informací (RVTL) zřízeného jako jeden z orgánů organizační struktury pro realizaci Akčního plánu GeoInfoStrategie.

infrastruktura pro prostorové informace (spatial data infrastructure) – soustava politik, zásad, znalostí, institucionálních opatření, technologií, mechanismů, procesů, prostorových dat, datových sad, služeb nad prostorovými daty, metadat a lidských zdrojů zřízených pro konkrétní účely veřejné správy nebo soukromoprávních subjektů;

národní infrastrukturou pro prostorové informace – soustava politik, zásad, znalostí, technických specifikací, institucionálních opatření, technologií, mechanismů, procesů, dat a lidských zdrojů pro zajištění centrální koordinace jednotlivých infrastruktur pro prostorové informace, jejichž prostorová data, datové sady a služby nad prostorovými daty jsou využívané pro účely veřejné správy;

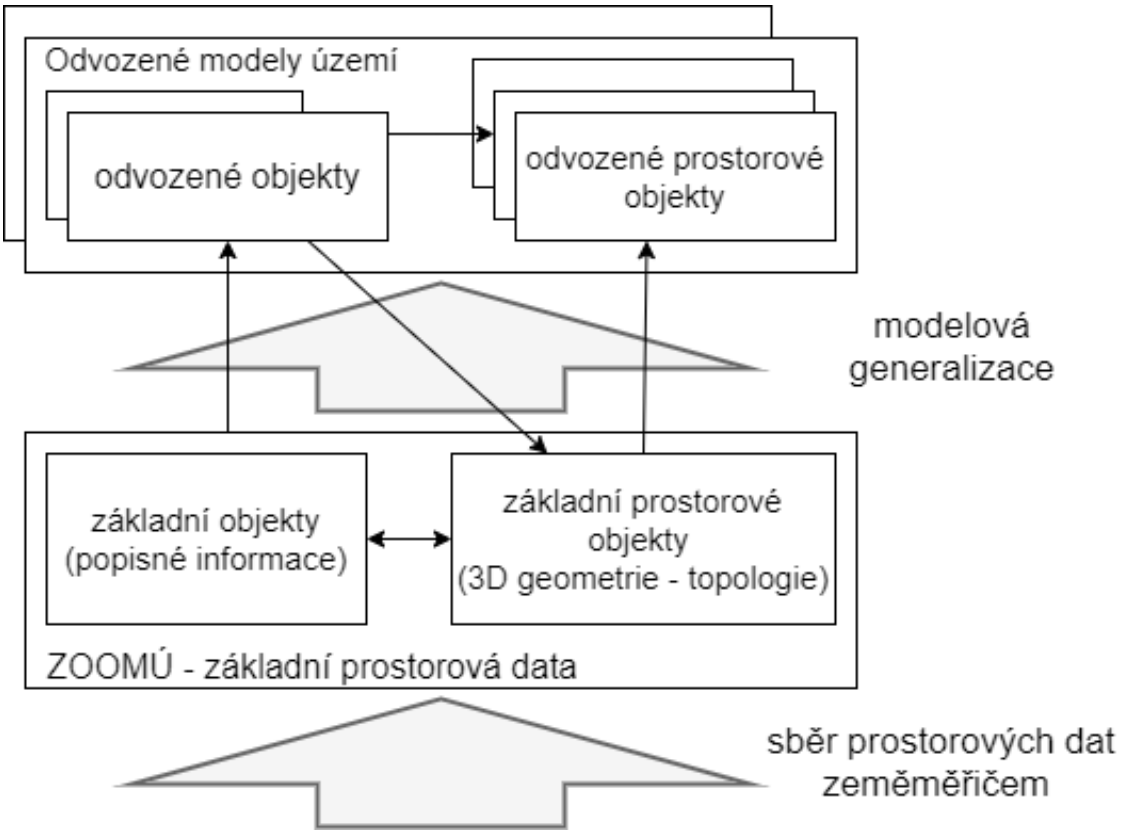
V rámci České republiky existuje velké množství infrastruktur pro prostorové informace, které mají rozdílné obory zájmu, jsou spravovány a vyvíjený pro konkrétní téma a účel, využívají různé technologie, standardy a pojmy pro popis aspektů vybraných jevů reálného světa. Cílem NIPI je zajistit efektivní interoperabilitu a harmonizaci dat především na úrovni veřejné správy a soukromoprávních subjektů, které produkují prostorové informace nutné pro výkon agend veřejné správy. V rámci budování NIPI musí být nalezen vhodný referenční model, který zajistí homogenitu prostorových informací a efektivní model pořizování prostorových informací umožňující harmonizaci prostorových dat jednotlivých subjektů.

Popisu referenčního modelu ISO norem geografické informace a výhodám modelování založeného na jeho využívání jsou věnovány předešlé kapitoly. Následující text se zabývá především návrhem modelu pořizování prostorových informací, který vychází z klasifikace datových modelů uvedené v úvodu této kapitoly.

Národní infrastrukturu pro prostorové informace tvoří souhrn jednotlivých infrastruktur zabývajících se určitou odbornou oblastí v nichž jsou na pořizování svých prostorových dat o jevech reálného světa využívány veřejné finanční prostředky. Přestože jsou obory zájmu jednotlivých infrastruktur odlišné, tedy popisují různé aspekty vybraných jevů reálného světa, často se vztahují na tytéž jevy reálného světa. Příkladem může být typ jevu „budova“, pro který katastr spravuje informace o vlastnictví, stavební úřad o konstrukčních aspektech, městský úřad o využití, atd. Na získávání prostorových dat v jednotlivých infrastrukturách jsou vynakládány stamilionové částky z veřejných zdrojů. Ať už na sběr základních prostorových dat pořizovaných geodetickými a fotogrammetrickými metodami sběru geografických dat, šetřením v terénu nebo na odvozování prostorových dat modelovou generalizací z poskytnutých dat. Cílem NIPI musí být zajištění efektivního vynakládání těchto veřejných zdrojů. Pořizování odvozených

prostorových dat v rámci NIPI lze výrazně zefektivnit aplikováním automatizované generalizace a optimalizací IT technologií. Zefektivnění vlastního sběru dat z terénu, to je zeměměřických činností, však bude po všech stránkách proces mnohem náročnější. Pro tyto účely lze aplikovat na strukturu NIPI výše uvedenou klasifikaci digitálních modelů (viz obrázek 29).

Pro zajištění efektivního pořizování základních prostorových dat je doporučováno v rámci NIPI vytvořit centrální základní digitální model území. S ohledem na současný vývoj v oblasti prostorových informací se jeví jako nutnost zaměřit se na vytvoření uceleného objektově orientovaného 3D geometricko-topologického modelu území. Tento model můžeme pracovně nazývat „Základní objektově orientovaný digitální model území“ a označovat akronymem ZOOMÚ. Tento základní digitální model bude naplňován na základě výsledků sběru prostorových dat zeměměřičem, podle jasně definovaných pravidel (metodik). Požadavky na obsah, podrobnosti a kvalitu základních prostorových dat musí být stanoveny tak, aby vyhovovaly požadavkům subjektů VS, které tato data využívají jako referenční data pro získání odvozených prostorových dat modelovou generalizací.



Obrázek 29: Generalizace v modelu pořízení prostorových dat v rámci NIPI

Oborem zájmu vytvořeného základního digitálního modelu budou především všeobecně využitelná elementární data pořízená geodetickými nebo fotogrammetrickými metodami. Rozsah popisných informací v základních prostorových datech reflektuje především možnosti získání informací zeměměřičem v terénu. Podrobnost a kvalita základních prostorových objektů, tedy geometrických a topologických objektů, musí být stanoveny tak, aby vyhovovaly pro odvození prostorových objektů (geometrie) nižší úrovně podrobnosti a stanovené kvality v rámci VS. Pokud je struktura základních prostorových dat vhodně zvolena, mohou je jednotlivé infrastruktury v rámci NIPI využít jako referenční pro odvození prostorových dat vhodných pro jejich specifické účely. Jejich odvozené objekty obsahující popisné informace mohou přímo odkazovat na základní prostorové objekty, pokud je pro ně základní 3D prostorová reprezentace vyhovující. V případě, že některá z infrastruktur pro prostorové informace v rámci NIPI potřebuje pro své účely jinou podrobnost prostorové reprezentace (například 2D, CityGML LoD 1, CityGML LoD 2) lze využívat automatizovanou modelovou generalizaci ze základních prostorových objektů.

Vytvoření modelu pořízení dat NIPI na základě centrálního geometricko-topologického modelu tak logicky umožní harmonizaci prostorové reprezentace společných jevů napříč infrastrukturami NIPI, ušetří veřejné prostředky vynakládané na pořízení prostorových dat zeměměřičem a ve výsledku umožní rychlejší aktualizaci prostorových dat, které tak budou více odpovídat skutečnému stavu v terénu.

Při definování jednotlivých typů objektů v NaSaPO musí být nutně brán zřetel především na současné metody a technologie mapování a na úroveň znalostí zeměměřiče, které je nutné certifikovat. Po stanovení metodik sběru dat a zajištění kvality základních prostorových dat je nutné se zaměřit na certifikaci procesu sběru prostorových dat. Jednou z možností je například revidování rozsahu odborných znalostí požadovaných při současném udělování úředního oprávnění k provádění zeměměřických činností (ÚOZI) a navazujících předpisů.

6. Závěr terminologického a věcného vymezení objektů NaSaPO

Z revizí textů mezinárodních a evropsky harmonizovaných norem vyplývá jednoznačný závěr, že pokud je správně nastavena konstrukce aplikačních částí NaSaPO, lze spolehlivě mapovat české odborné termíny (koncepty) do ukotvených mezinárodně odsouhlasených termínů. S ohledem na problematiku objektů NaSaPO je třeba upozornit, že každý typ třídy objektů NaSaPO by měl mít svoji malou převodní terminologickou API a orchestrací těchto služeb pak může být obsloužena celá terminologická problematika projektu.

Toto tvrzení se opírá o zásadní principy a jedno z pravidel projektových prací nad prostorovou informací, konkrétně o využívání Obecného modelu vzhledu jevu: Tento model umožňuje, že na úrovni konkrétních prací v terénu (tj. na úrovni instancí vzhledů jevů) nejsou výše popsané postupy terminologického ošetření a ontologického mapování nijak podřízeny vůči jednotné mezinárodní terminologii. V této části řešení projektu NaSaPO se tedy může dále pokročit v návaznosti na konkrétní objekty, které budou do NaSaPO zařazeny.

K přípravě a provedení tohoto mapování konceptů k Obecnému modelu vzhledu jevu je potřebné prakticky zvládnout především zásady pro tvorbu definic nových pojmů podle ČSN ISO 704 (01 0505) Terminologická práce – Principy a metody.

Tyto termíny budou doplňovány postupně a jako jeden z vedlejších výsledků řešení projektu NaSaPO bude vytvořen a průběžně doplňován Terminologický výkladový slovník projektu NaSaPO (Příloha A). Tento slovník bude veden jako vícejazyčný, aby bylo zabráněno nekvalitním a zavádějícím ad-hoc překladům.

Výsledky provedené rešerše relevantních terminologických slovníků a vyhodnocení stávajícího stavu terminologie v oblasti geografické informace v České republice i v zahraničí potvrdily zásadní význam jednotné závazné terminologie vycházející z ISO norem řady 19100, potvrdily opodstatněnost ustanovení Aktualizovaného Akčního plánu GeoInfoStrategie [2] i požadavků zadání tohoto projektu NaSaPO ohledně používání terminologie a zcela jednoznačně hovoří ve prospěch důsledného využití konzistentní terminologie ISO norem řady 19100 geografické informace v rámci všech součástí NIPI.

Obecná úvaha o identifikaci, lokalizaci a sémantické interoperabilitě

Schopnost identifikovat a lokalizovat relevantní informace uprostřed rozsáhlé sbírky prostorových dat nebo jiných zdrojů (sbírky map, sbírky hornin) je hlavní a naléhavou výzvou moderního využívání nejrozmanitějších zdrojů. Po právu se stalo předmětem výzkumu v rámci projektu NaSaPO.

Pokud toto využívání nejrozmanitějších zdrojů je mainstreamem aplikací propojených na prostorovou informatiku, je to vše dostatečným vysvětlením potřeby sémantické interoperability. Sémantická interoperabilita je důsledkem vývoje webových služeb a jiných nástrojů, včetně registrů slovních zásob (slovníčků) a schémat metadat, úložišť slovníků a metadatových schémat a služeb obsluhujících vzájemně se překlenující cíle.

Pokud prozatímně vynecháme registry a úložiště, je možné si představit situaci, kdy se zabýváme postupem, jak nakonec ustanovit konkrétní překlenující služby. Ty obvykle spočívají na vzájemném přiřazení mezi různými schématy a slovními zásobami.

Základ pro diskuzi o nejlepší možné výměně konceptů prostorových dat, tj. o interoperabilitě, je položen v ISO 25964-1, která popisuje klíčové charakteristiky tezaurů a poskytuje návod nejlepší praxe. Bohužel srovnatelná mezinárodní norma neexistuje pro jiné typy slovníků, se kterými tezaurus potřebuje interoperovat. Pro tento účel poskytuje tato část ISO 25964 některé elementární popisy z jiných slovníčků, jako jsou klasifikační schémata, schémata s vůdčím subjektem, atd. předtím, než doporučí mapování mezi tímto přístupem a tezaurem. To poskytuje návod na konstrukci slovníčků jiných než tezaurů.

V zásadě budou práce probíhat ve dvou tematických částech:

První část se zaměří na nakládání s principy a praxí interoperability, zejména mapování, což vyžaduje většina slovních zásob (slovníčků) a speciálně tezaury. Dodatečně se pozornost zaměří na návod pro výměnu předběžně koordinovaných tříd, které se vyskytují v klasifikačních schématech a jiných slovních zásobnících, používajících klasifikační přístup.

Druhá část se zaměří na jiné typy slovních zásob. Ačkoliv mají různý účel, terminologie, ontologie a synonyma jsou často zahrnuta. Budeme zkoumat jak rychlý popis informace slovníkových klíčových charakteristik, které kontrastují s jejich sémantickými komponentami, je poskytován. To umožní prozkoumat kontext pro požadavky interoperability a navrhnout doporučení, jak pokračovat.

Seznam příloh

A. Terminologický výkladový slovník projektu NaSaPO
B. Terminologie znalostních knihoven
C. Přehled ISO norem geografické informace
D. Zásady tvorby definic
E. Vznik objektově orientovaného přístupu
F. Termíny GeoInfoStrategie v projektu NaSaPO

Literatura

[1] Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020 (GeoInfoStrategie). Ministerstvo vnitra aj. Praha 2014

[2] Akční plán Strategie rozvoje infrastruktury pro prostorové informace v České republice do roku 2020. Ministerstvo vnitra aj. Praha 2016

[3] ČSN EN ISO 19101-1, Geografická informace – Referenční model – Část 1: Základy; Český normalizační institut. Praha 2018

[4] ČSN P ISO/TS 19103, Geografická informace – Jazyk konceptuálního schématu; Český normalizační institut. Praha 2006

[5] ČSN P ISO/TS 19104, Geografická informace – Terminologie; Český normalizační institut. Praha 2010

[6] ČSN EN ISO 19107, Geografická informace – Prostorové schéma; Český normalizační institut. Praha 2005

[7] ČSN EN ISO 19109, Geografická informace – Pravidla pro aplikační schémata; Český normalizační institut. Praha 2020

[8] ČSN EN ISO 19115-1, Geografická informace – Metadata, Část 1: Základy; Český normalizační institut. Praha 2020

[9] ČSN ISO/IEC 2382-17. Informační technologie – Slovník – Část 17: Databáze. Český normalizační institut. Praha 1999

[10] UN-GGIM Integrated Geospatial Information Framework (IGIF), Part 2: Implementation Guide – Strategic Pathway 6 Standards. Dostupné z: https://ggim.un.org/IGIF/documents/SP6-Standards-7Apr2020-GLOBAL-CONSULTATION.pdf

[11] OMG/UML UML Semantics, version 1.3, 1997.

[12] Kresse, W., Danko, M., Fadaie, K. in Handbook of Geographic Information (ed. W. Kresse, M. Danko), Part B, Chapter 13, Standardization, pages: 393-565, Springer, 2012, ISBN: 978-3-540-72678-4

[13] Mildorf, T.: Modelová generalizace pozemkového datového modelu, ZČU – doktorská disertační práce, 2012, Plzeň

[14] Grünreich, D., Počítačově podporovaná generalizace in Session: Data Manipulation, CERCO Cartography Course 21. – 29. 10. 1985, Frankfurt a. M.

[15] Hake, G. (1982) Kartographie I. Sammlung Göschen, Band 2165, Walter de Gruyter-Verlag, Berlin-New York

[16] Hardy, P.: Multi-scale database generalization for topographic mapping, hydrography and web-mapping, using active object techniques. International archives of photogrammetry and remote sensing, vol. XXXIII, part B4. Amsterdam 2000.

[17] Sarjakoski, L. T.: Conceptual Models of Generalization and Multiple Representation, in Generalisation of Geographic Information: Cartographic Modelling and Applications. Elsevier, Oxford 2007, ISBN-13:978-0-08-045374-3

[18] Kresse, W. International Standards for Geographic Information, In: EduServ 14, 2016, University of Applied Sciences, Neubrandenburg

[19] https://github.com/ISO-TC211/UML-Best-Practices/blob/master/Reference%20material/11-107_OWS-8_Domain_Modelling_Cookbook.pdf
[20] Neumann, J. „Vzhled jevu jako ekvivalent termínu feature v českém překladu mezinárodních norem geografické informace“, GEO Informace č. 1/2003, pp. 10–11
[21] Dokument EU COM(2017) 134 final ANNEX 2 Evropský rámec interoperability – Strategie provádění, Brusel, 2017
[22] Neubauer, Z. Smysl a svět. Hermeneutický pohled na svět. Praha, Moraviapress, a. s., 2001, 231 s.
[23] Zemek, J. a kol., Kvalita dat v geoprostorových databázích, VÚGTK Zdiby 2011
[24] Updated Guideline for Implementing the ISO 19100 Geographic Information Quality Standards in National Mapping and Cadastral Agencies, v. 2.0, EuroGeographics QKEN, 2018, 63 s
[25] Data Quality in INSPIRE: Balancing Legal Obligations with Technical Aspects. EC JRC Technical Reports, Ispra, 2013
[26] The Defence Geospatial Feature Concept Dictionary (DGFCD) Description and Content, v. 2.0.0, DGIWG 2017. Dostupné na: https://www.dgiwg.org/dgiwg-standards/206
[27] Buttenfield, B., McMaster R., Map generalization: Making rules for knowledge representation, Longman Scientific & Technical, Longman Group Limited, 1991. Dostupné na: https://www.researchgate.net/publication/242517193_Map_Generalization_Making_Rules_for_Knowledge_Representation/link/0deec530271a5e855c000000/download
[28] Sester M., Cartographic generalization, Journal of spatial information science, Number 21 (2020), pp. 5 - 11. Dostupné na: http://www.josis.org/index.php/josis/article/viewFile/716/296
[29] Multilingual Dictionary of Technical Terms in Cartography, International Cartographic Association, 1973

Tento projekt je financován se státní podporou
Technologické agentury ČR v rámci
Programu BETA2.

www.tacr.cz
Výzkum užitečný pro společnost.

MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

II. část

Metodika pro budování otevřené a rozšířené ontologické struktury (NMet – certifikovaná metodika)

Zpracovali: doc. Ing. Václav Čada, CSc., Západočeská univerzita v Plzni
doc. Ing. Mgr. Otakar Čerba, Ph.D., Západočeská univerzita v Plzni

7 Metodika pro budování otevřené a rozšířené ontologické struktury

Cílem metodiky pro budování otevřené a rozšířené ontologické struktury (tříd) v projektu NaSaPO je navržení postupu, který umožní vytvoření ontologické struktury používané pro účely konceptuálního modelování tříd objektů v projektu NaSaPO. Ontologická struktura se bude skládat ze systému tříd ontologie, instancí ontologie, relevantních objektových, datových (datotypových) i anotačních vlastností a příslušných metadat. Ontologická struktura bude obsahovat konstrukce umožňující vazby na prvky ontologií v externích zdrojích (podpora propojených dat). Součástí metodiky budou také postupy zaměřené na rozšiřování, aktualizaci, údržbu a správu ontologické struktury.

Ontologie pro potřeby NaSaPO bude sloužit jako základní nástroj konceptuálního modelování. Jejím účelem bude propojení tříd a objektů reálného světa a tříd objektů modelového světa (konkrétně datového modelu NaSaPO). Ontologie by měla umožňovat pomocí vazeb slovníku SKOS (Simple Knowledge Organisation System)[5] propojení také s prvky jiných nomenklatur, tezaurů a ontologií publikovaných jako propojená data. Ontologie bude sloužit jako most mezi subjekty pořizujícími data a jednotlivci i institucemi, které se zabývají zpracováním a ukládáním dat.

Metodika je určená pro tvůrce ontologických systémů popisujících třídy, které reprezentují typy prostorových objektů, včetně NaSaPO. Skupina uživatelů nezahrnuje pouze experty vytvářející kód vlastní ontologie, ale především širokou skupinu odborníků, která zahrnuje manažery celého projektu a správce (osoby reprezentující) jednotlivé poskytovatele a uživatele dat. Z tohoto důvodu byla metodika koncipována jako otevřený a flexibilní proces, který umožňuje problematiku tvorby ontologie uchopit konceptuálně bez ohledu na konkrétní sadu modelovaných objektů.

7.1 Úvod do metodiky pro tvorbu ontologie

Thomas Gruber (Gruber, 1993) definuje ontologie jako „explicitní specifikace konceptualizace“. Tato definice byla rozšířena Willemem Nico Borstem (Borst, 1997), který ontologii považuje za „formální specifikace sdílené konceptualizace“. Pokud bychom se obě výše uvedené definice pokusili vyjádřit českými slovy, pak můžeme ontologii charakterizovat jako „jasný, zřetelný a přímo vyjádřený přesný výčet pojmů daného systému, včetně vztahů mezi těmito pojmy“. Klíčový termín konceptualizace (použitý v obou definicích) je ovšem mnohem širší a kromě jiného znamená „způsob jak člověk chápe svět a jak ho vyjadřuje“, přičemž koncepty mohou představovat prvky jako například parcela, dálnice nebo jezero (Bittner & Winter, 2004). V rámci ontologie pro účely NaSaPO bude termín „koncept“ používaný v kontextu ontologií a v souladu s terminologií používanou v ontologickém inženýrství. Výraz „koncept“ bude souhrnně označovat objekty, jevy nebo třídy ve smyslu, jak jsou nebo budou definovány pro účely NaSaPO. Z pohledu informačních technologií je nesporně důležitý také přívlastek „formální“, který vyjadřuje určitou preciznost, explicitnost a jednoznačnost, přičemž tyto vlastnosti jsou velice důležité z hlediska možnosti strojového (automatizovaného) zpracování.

[5] W3C Recommendation SKOS je kompatibilní s relevantními ISO standardy, zejména ISO 2788 a ISO 5964.

Ontologie jsou považovány za produkt informačních technologií, proto by při jejich návrhu a tvorbě měly být využity standardizované postupy (metodiky). Těmi se v případě ontologií a znalostních systémů zabývá obor nazývaný ontologické inženýrství. Standardizace v rámci ontologického inženýrství však naráží na fakt existence příliš velké závislosti dílčích aktivit (například volbě podrobnosti ontologie) při tvorbě ontologií na individuálních pohledech, potřebách a přístupech. Existuje téměř nekonečné množství různých pohledů na jednu doménu a také je nutné počítat s individuálním přístupem každého odborníka zapojeného do procesu návrhu a tvorby ontologií. Proto obvykle není možné nalézt jediný, nejlepší způsob nebo postup, ve smyslu nejvhodnější a ideální metodiky pro vytvoření ontologického systému. Během posledních třiceti let bylo vyvinuto velké množství různých metodik – popisů, jakým způsobem co možno nejefektivněji, za použití existujících struktur zkonstruovat nové ontologie. Tyto metodiky byly většinou vytvořeny během vývoje nějakého konkrétního systému nebo projektu (například Cyc nebo projekt TOVE), a tím pádem došlo k praktickému ověření jejich funkcionality.

Podle (Gaševič et al. 2006) se metodika pro vývoj ontologií skládá z principů, procesů, metod, aktivit vedoucích k návrhu, konstrukci, evaluaci a dalšímu rozvoji ontologie. Obecně lze říci, že všechny uvedené metodiky spojuje snaha o co nejvěrnější a nejúplnější zachycení pojmového aparátu (konceptů) ontologie a vztahů mezi jednotlivými pojmy. Také proto jsou si jednotlivé metodiky velmi podobné, neboť téměř všechny obsahují klíčové aktivity jako například testování, identifikace základních informací apod.

Na základě rešerše vybrané literatury bylo zjištěno, že zřejmě nejvíce zmiňovanou metodikou je metodika s názvem METHONTOLOGY. Zároveň je zapotřebí brát v úvahu výsledky průzkumu, který prováděl J. Cardoso (Cardoso, 2007). Ty paradoxně uvádí, že nejvíce ontologií je vytvářeno bez jakékoli metodiky (víceméně živelným způsobem, celkově 60 %) a pokud byla nějaká metodika použita, pak se na prvním místě jednalo o On-To-Knowledge (13,9 %).

V článku (Čerba, 2019) je detekováno osm metodik vhodných pro tvorbu ontologií obsahujících třídy a objekty tak, jak jsou nebo v průběhu projektu budou definovány pro potřeby NaSaPO. Výčet byl vytvořený na základě mnoha srovnávacích studií (Jones et al., 1998, Fernández-López, 1999, Corcho et al., 2003, Cristani & Cuel, 2005, Iqbal et al., 2013) s přihlédnutím k pětici tzv. zralých metodik podle Cardoso (2007). Tato množina vhodných metodik obsahuje následující položky: TOVE, Enterprise Model Approach, METHONTOLOGY, Methodology publikovaná v Bernaras et al. (1996), On-To-Knowledge Methodology, DOLCE, Ontology Development 101 a UPON. Kromě těchto ontologií je možné při návrhu ontologií využít i další metodiky, jako Sensus (popsáno v Corcho et al., 2003), KACTUS (Schreiber et al., 1995, Corcho et al., 2003), Sugumaran and Storey (Sugumaran & Storey, 2002), Unified Methodology (Uschold, 1996), MSPM Ontological Methodology (Cheah, 2007) nebo návrh metodiky popsaný v publikaci (Öhgren & Sandkuhl, 2005).

Protože se na tvorbě a především plnění ontologie pro účely NaSaPO bude podílet mnoho velmi autonomních subjektů z různých oborů, je vhodné do metodiky začlenit v maximální možné míře participativní faktory. Ty by umožňovaly efektivní komunikaci tvůrců ontologické struktury, osob podílejících se na plnění ontologie a další zainteresovaných subjektů (tzv. stakeholders). Hodnocení metodik z pohledu participace externích institucí a expertů bylo provedeno v publikaci (Čerba, 2019). Velké množství uvedených metodik uvádí jako ideální případy participativního zapojení uživatelů následující kroky: **souhrn a specifikaci požadavků na ontologii, identifikaci cílů a tzv. kompetenčních dotazů, evaluaci a testování**. Kromě prvků z metodik ontologického inženýrství metodika pro tvorbu ontologií pro potřeby NaSaPO obsahuje také části technik a procesů jako jsou Design Thinking (Brown, 2008), Lean Startup (Ries,

2011; Eisenmann et al. 2012) a Agile development (Beck et al. 2001; Schwaber & Beedle, 2002). Právě ty umožní praktičtější monitoring projektu a propojení na reálné uživatele.

7.2 Metodika pro tvorbu ontologie

Konkrétní metodika pro tvorbu ontologie pro účely NaSaPO sestavená z výše uvedených zdrojů se skládá z následujících kroků:

1. Specifikace
2. Vývoj úvodní verze ontologické struktury
3. Vylepšování úvodní verze ontologie, její další členění a doplňování obsahu
4. Správa ontologie
5. Implementace ontologie
6. Evaluace ontologie
7. Tvorba dokumentace

Tato metodika svým rozsahem překračuje projekt TITIMV706IP01. Proto některé kroky bude potřeba realizovat nebo dokončovat i v dalších navazujících aktivitách, včetně reálného provozu celého produktu NaSaPO.

7.3 Specifikace ontologie

Specifikace ontologie definuje cíl, důvod vytvoření a rozsah ontologie, případně specifické požadavky (ty budou průběžně doplňovány na základě zpětné vazby uživatelů). **Cílem** ontologie je vytvořit konceptuální model pro potřeby NaSaPO. Model bude popisovat objekty a třídy reálného světa, které budou v rámci NaSaPO modelované prostřednictvím datových struktur (**rozsah ontologie**), a jejich vztah k datovým modelům NaSaPO. Ontologie by měla sloužit jako nástroj pro komunikaci mezi pořizovateli dat sbíranými pro účely NaSaPO a jejich zpracovateli, potažmo uživateli (**důvod vytvoření ontologie**).

Ontologie pro účely NaSaPO bude mít charakter doménové ontologie. Bude obsahovat objekty a třídy spadající do oblasti působnosti NaSaPO. Tyto prvky budou v ontologii popsány pomocí tříd ontologie (prvek class). Dále bude ontologie obsahovat relace (objektové, datové a anotační) a logická omezení.

Úkolem tohoto projektu je vytvořit strukturu ontologie, která musí být natolik robustní a flexibilní, aby bylo možné reagovat na jakékoliv požadavky z oblasti resortů, které budou do ontologie poskytovat své informace a znalosti. Projekt musí vytvořit takové mechanismy, aby organizace odpovědná za ontologii (její správu, údržbu i naplňování) s ní mohla komfortně pracovat a zároveň, aby tam existovaly mantinely zabraňující svévolným změnám a živelným změnám (rozšiřování).

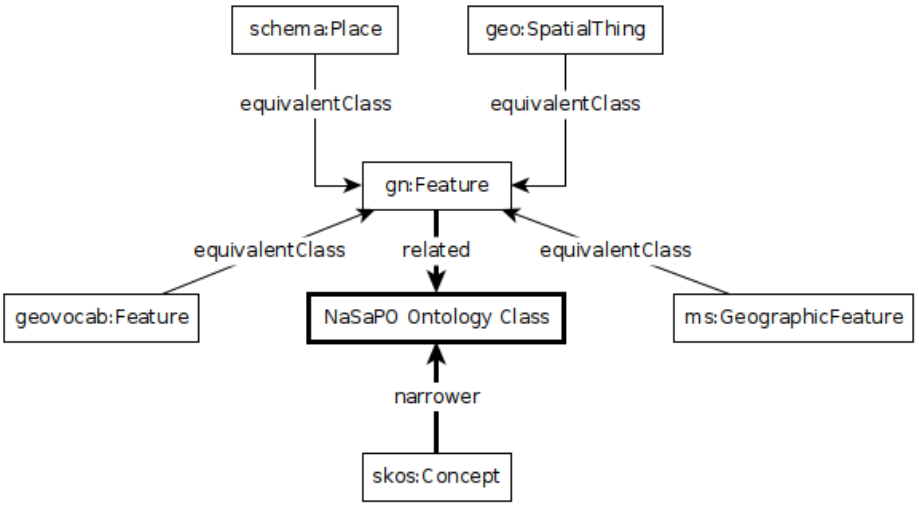
Použité jmenné prostory[6]:

- GeoNames Ontology – *gn: – http://www.geonames.org/ontology/*
- OWL – *owl: – http://www.w3.org/2002/07/owl#*
- RDF – *rdf: – http://www.w3.org/1999/02/22-rdf-syntax-ns#*
- RDFS – *rdfs: – http://www.w3.org/2000/01/rdf-schema#*
- SKOS – *skos: – http://www.w3.org/2004/02/skos/core#*
- SKOS XL – *skosxl: – http://www.w3.org/2008/05/skos-xl#*
- XML – *xml: – http://www.w3.org/XML/1998/namespace*
- XSD – *xsd: – http://www.w3.org/2001/XMLSchema#*

7.4 Jádrová ontologie

Kořenovým prvkem ontologie pro účely NaSaPO je třída (ve smyslu jazyka OWL, v tomto smyslu bude termín třída používán v této metodice i dále, pokud nebude uvedeno jinak) NaSaPO Ontology Class (kvůli srozumitelnosti navrhuji anglická jména pro univerzální třídy, která budou doplněná o české popisky; zkratka NOC). Třída NOC představuje zastřešující (do jisté míry abstraktní) kontejner pro jakékoliv další třídy a individuály ontologie pro účely NaSaPO.

Tato třída je sémanticky propojená (*skos:related*) s třídou gn:Feature z GeoNames Ontology (schéma na obrázku 30). Zároveň jde o specifikaci (*skos:narrower*) třídy *skos:Concept* ze standardu SKOS (Miles & Bechhofer, 2009). Následující schéma ukazuje vazby tříd *gn:Feature* na další prvky světa propojených dat (slovníky Schema.org, Basic geo (wgs84 lat/long) vocabulary nebo GeoVocab). Tučné linie ve schématu představují vazby a třídy definované pro ontologii pro účely NaSaPO. Naopak tenké linie ukazují třídy a relace v existujících znalostních bázích, ontologiích a dalších sémantických nástrojích.



Obrázek 30: Vazba kořenového objektu na externí sémantické nástroje

[6] Jmenné prostory jsou mechanismy, které podporují jednoznačnou identifikaci. Umožňují seskupování identifikátorů do skupin (zpravidla rozdělených na základě tématu), přičemž kombinace jmenného prostoru a identifikátorů by měla být unikátní.

Diskutována byla také vazba na standard GeoSPARQL (Perry & Herring, 2012). Ten však má vazbu na na UML class GFI_Feature definovanou v ISO 19156. Proto vazba na GeoSPARQL bude v případě potřeby začleněna do pozdější verzí.

Identifikátor tříd ontologie bude kromě URI ontologie (bude definováno před publikováním ontologie ve spolupráci s výsledkem V07) obsahovat název třídy (v češtině, s počátečním velkým písmenem, bez diakritiky a s podtržítky místo mezer). Popisky tříd (jako jejich anotační vlastnosti) budou kódovány pomocí standardní vlastnosti *rdfs:label*, včetně příslušného označení jazyka (*@cs*[7]). Pojmenování bude v češtině, s počátečním velkým písmenem, s diakritickými znaménky a s mezerami mezi slovy.

Pokud bude jedna třída ontologie vyžadovat více pojmenování, bude se tento případ řešit pomocí vlastností standardu SKOS eXtension for Labels (SKOS-XL). Ten umožňuje ke každé třídě ontologie doplnit doporučovaný (preferovaný) název (*skosxl:prefLabel*) a případná alternativní pojmenování (*skosxl:altLabel*). V ontologii pro účely NaSaPO by se mělo doporučené pojmenování (*skos:prefLabel*) shodovat s hlavním označením třídy (*rdfs:label*).

Kromě identifikátoru a popisků jsou jednotlivé typy objektů a jejich instance popsány pomocí dalších standardizovaných anotačních vlastností:

- *skos:definition* – definice typu objektu, včetně zdroje. Jeden typ objektu může obsahovat více různých definic. Zdroje jsou k definicím doplňovány formou volného textu (optimálně omezeného vhodně zvolenou citační konvencí). Doplnění zdrojů k definicím je realizováno pomocí „anotace anotace“, kdy je vlastnost skos:definition popsána (anotována) pomocí anotační relace *rdfs:isDefinedBy*.

- *skos:note* – poznámka k definici. Podobně jako v předchozím případě, může být poznámka použita jako rozšíření definice nebo jiné anotační vlastnosti.

Třídy ontologie jsou propojené pomocí sémantických vlastností definovaných ve standardu SKOS (Miles & Bechhofer, 2009). Základní sémantické vztahy mezi třídami jsou popsány pomocí objektových vlastností *skos:related*, *skos:broaderTransitive*, *skos:broader*, *skos:narrowerTransitive* a *skos:narrower*. V názvech vlastností i v jejich popisu (součást publikace Miles & Bechhofer, 2009) se vyskytují tři základní pojmy:

- broader – odkazuje na třídu (skos:Concept) s obecnějším významem (například *„Silnice 1. třídy“ skos:broader „Silnice“*)

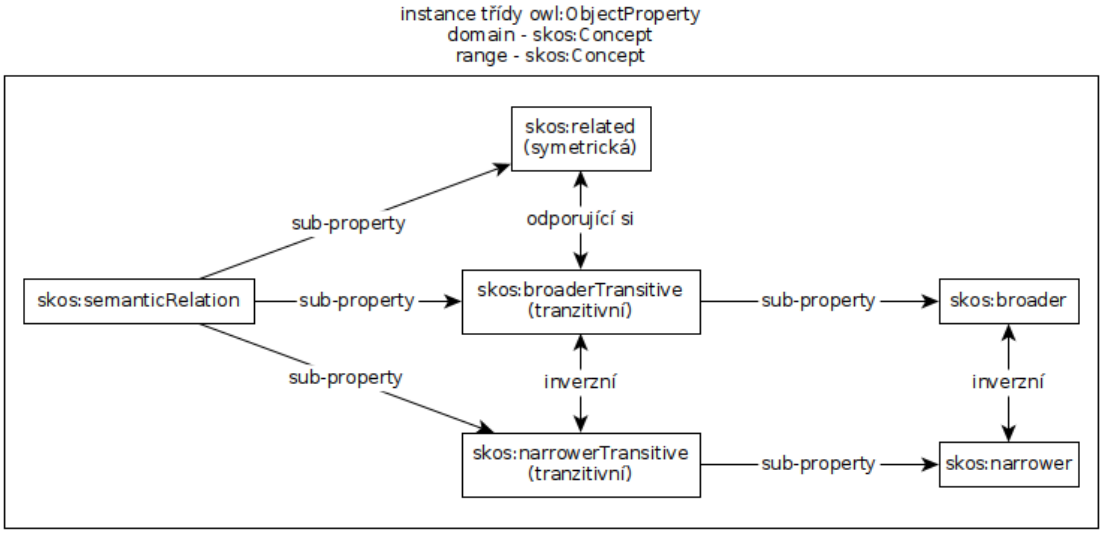
[7] Jazykový atribut *xml:lang* by měl být součástí každé anotační vlastnosti, zvláště pokud ontologie bude obsahovat vícejazyčnou terminologii.

- narrower – odkazuje na třídu (skos:Concept) s užším (specifičtějším) významem (například *„Silnice“ skos:narrower „Silnice 1. třídy“*)

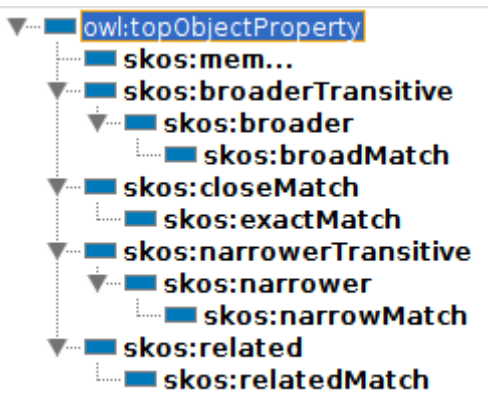
- related – se používá pro vyjádření asociace (spojení volného charakteru)

Vzájemné vazby a omezení sémantických vlastností ukazuje následující schéma (obrázek 31) vytvořené podle Miles & Bechhofer (2009).

Mapovací vlastnosti standardu SKOS (obrázek 32) slouží k propojení konceptů náležejících do různých ontologií (konceptuálních schémat). V případě ontologie pro účely NaSaPO budou využívány pro vytvoření vazeb mezi třídami ontologie a prvky z jiných sémantických nástrojů (slovníků, tezaurů a podobně).



Obrázek 31: Vazby a omezení sémantických vlastností



Obrázek 32: Objektové vlastnosti ontologie určené mapování konceptů

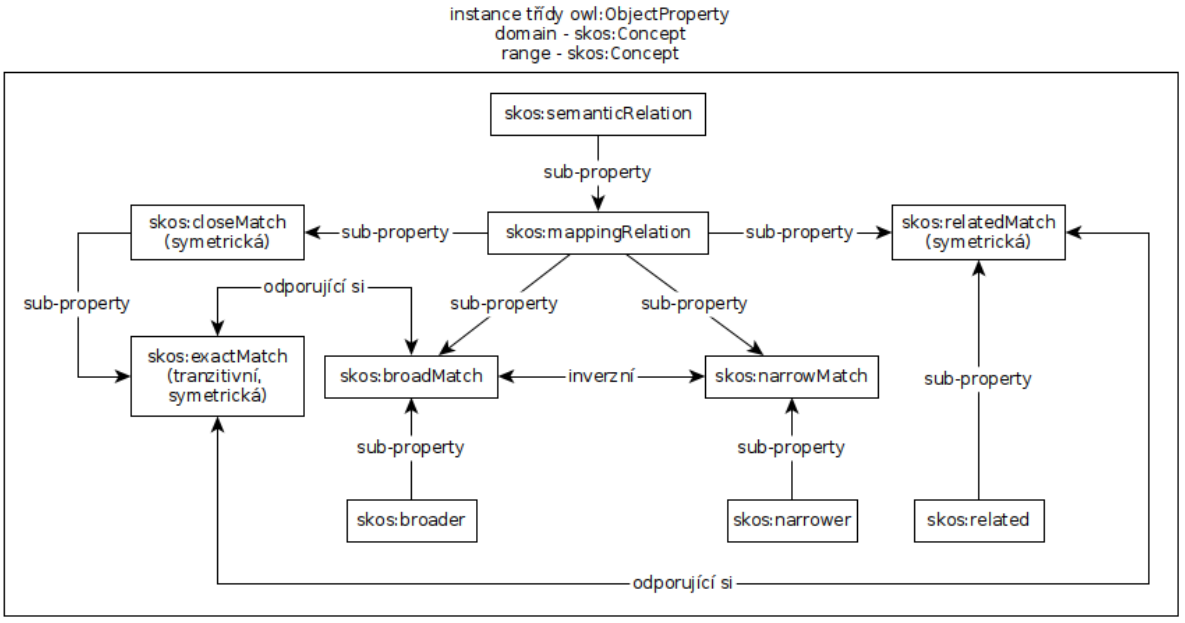
V ontologii budou implementované tyto vlastnosti (jejich využití pak závisí na konkrétních potřebách, požadavcích a objektech zařazených do ontologie):

- mapování na třídy, které vykazují vysoký stupeň podobnosti a některých specifických případech být používány jako ekvivalentní (mapování na podobné třídy) – *skos:closeMatch*
- mapování na třídy, které mohou být s vysokým stupněm spolehlivosti používány jako synonyma (mapování na identické třídy) – *skos:exactMatch*
- hierarchické mapování – *skos:broadMatch*, *skos:narrowMatch*
- asociační mapování – *skos:relatedMatch*

Následující schéma (obrázek 33) ukazuje vzájemné vazby a omezení mapovacích vlastností i jejich propojení na sémantické vlastnosti standardu SKOS.

Doporučení pro implementaci: Ontologie pro účely NaSaPO by se měla v maximální možné míře vyhýbat vlastnostem, které jsou ve standardu definovány vágním způsobem. Jedná se především o vlastnost skos:closeMatch, případně vlastnosti typy „related“.

Všechny objektové vlastnosti mají implicitně nastavený definiční obor (Domain) a obor hodnot (Range) kromě následujících výjimek, kde došlo k zpřesnění těchto informací (Tabulka 3).



Obrázek 33: Vazby sémantických a mapovacích vlastností

Vlastnost	Definiční obor (Domain)	Obor hodnot (Range)
skos:Member	*NaSaPO_Context*	*NaSaPO_Context* ∪ *NaSaPO_Ontology_Class*
skos:related	*NaSaPO_Ontology_Class*	*NaSaPO_Ontology_Class*

Tabulka 3: Upřesnění domén a oborů hodnot objektových vlastností

Pro účely mapování na externí zdroje byly vytvořeny třídy *External_Object* a *Vocabulary*. Třída *External_Object* je kontejner seskupující objekty z externích zdrojů (například Sémantický slovník pojmů je nástrojem určeným pro popis otevřených datových sad publikovaných veřejnou správou ČR nebo EuroVoc), na které jsou v rámci NaSaPO Ontologie vytvořené mapovací vazby. Tato třída byla nastavena jako definovaná (což zvýšilo klasifikační potenciál ontologie) pomocí následující podmínky: třída *External_Object* je ekvivalentní s anonymní třídou popsanou jako *skos:closeMatch some NaSaPO_Ontology_Class* (jinými slovy instance třídy *External_Object* musí být propojená pomocí minimálně jedné vazby *skos:closeMatch* s individuálem spadajícím do prvků ve třídě *NaSaPO_Ontology_Class*).

Třída *External_Object* zahrnuje podtřídy jednotlivých zdrojů propojených dat: DTM (ontologie Digitální technické mapy, EuroVoc, Sémantický slovník pojmů a Wikidata). Příslušnost k těmto podtřídám je definována na základě vazby (*skos:related*) externího objektu na slovník (instance třídy *Vocabulary*), v němž je daný objekt definovaný.

Podobně jako třída *External_Object* je dále členěna i třída *NaSaPO_Ontology_Class*. Ta obsahuje dvě specifické definované podtřídy: (1) objekty, které jsou napojené do systému otevřených propojených dat a (2) objekty 1. úrovně hierarchie prvků NaSaPO, která vychází ze struktury aplikačních schémat.

Vzhledem k tomu, že ontologie pro účely NaSaPO se dotýká mnoha resortů a aktivit veřejné správy, je nutné v ontologii vytvořit mechanismus, který bude schopný zajistit popis různých kontextů (perspektiv, pohledů) na jeden objekt. Jinými slovy ontologie bude umožňovat vícenásobný popis jednoho objektu. Podobné principy využívají i renomované světové ontologie a tezaury jako například GEMET (Themes), EuroVoc (microthesauri) nebo UNESCO Thesaurus (Groups). Vzhledem k tomu, že ontologie pro účely NaSaPO, bude pro účely tvorby vazeb na kontexty použita konstrukce *skos:Collection*[8] – *skos:member*[9] [10]. Pro ontologii popisovanou v tomto dokumentu byla vytvořena třída *NaSaPO_Context*, která je podtřídou *skos:Collection*.

Všechny výše popsané třídy (konkrétně *Vocabulary*, *External_Object*, *NaSaPO_Context* a *NaSaPO_Ontology_Class*) byly pomocí vlastnosti disjoint nastaveny jako disjunktní (bez společného průniku). Jinými slovy prvek ontologie nemůže být zároveň součástí nějakého externího zdroje (slovníku, tezauru…)

[8] Jedná o seskupení konceptů, které mají nějakou společnou vlastnost, která není zohledněna v hierarchii.
[9] Objektová vlastnost, která propojuje kolekci s koncepty nebo jinou kolekcí.
[10] Standard SKOS umožňuje používat i uspořádané kolekce, ale ty nemají pro danou ontologii žádný význam.

Jednotlivé kontexty musí být vytvořeny před naplňováním ontologie ve formě otevřeného či uzavřeného seznamu. Ideální variantou je seznam persistentních URI (ale v první fázi se pravděpodobně bude jednat o seznam textových položek). Navrhujeme vytvoření více kontextů, z nichž jeden se bude orientovat na agendu veřejné správy (například gesce jednotlivých ministerstev). Další seznam kontextů by měl vycházet ze standardizovaných tematických nebo multioborových číselníků (INSPIRE, informační systémy knihoven…).

Aktuální verze ontologie obsahuje

- 11 tříd,
- 12 objektových vlastností,
- 43 individuálů (z toho 22 typů objektů NaSaPO),
- 226 logických axiomů.

Současná verze ontologie je publikována ve formátu RDF/XML. Zdrojový kód je možné stáhnout pomocí tohoto odkazu. Ontologie má následující omezení:

- Pracuje s provizorním systémem identifikátorů (http://www.zcu.cz/NaSaPO#). Ten bude v průběhu projektu nahrazený novým systémem, který bude v souladu s koncepcí NaSaPO a tvorbou identifikátorů pro veřejnou správu v ČR.

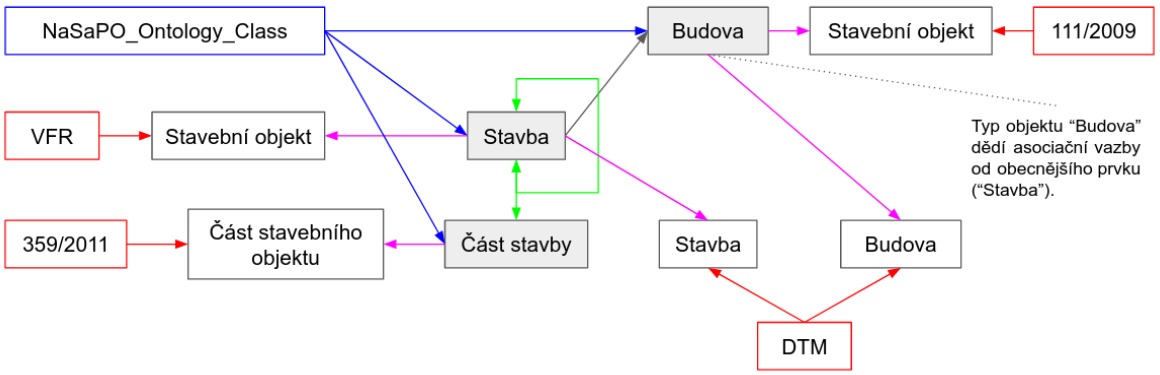
- Je na modelové úrovni. Naplňování ontologie (tzv. populating) ilustračními příklady objektů je součástí dalších kroků projektu. Vytvoření kompletního obsahu ontologie pro účely NaSaPO jde daleko za rámec tohoto projektu.

7.5 Příklad konceptu ontologie – Budova

Aktuální verze ontologického modelu prvku Budova vychází ze dvou základních zdrojů: Ontologie vytvořená pro účely jednotného výměnného formátu Digitální technické mapy a Sémantického slovníku pojmů pro popis otevřených datových sad publikovaných veřejnou správou ČR.

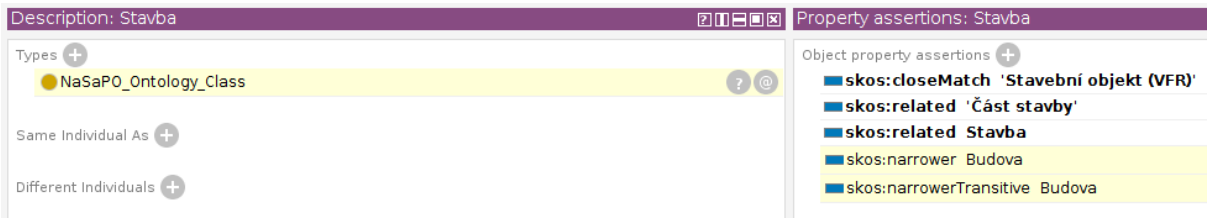
Následující schéma (obrázek 34) ukazuje vztahy mezi základními instancemi Ontologie pro potřeby NaSaPO a externími prvky z výše uvedeného slovníku. Prvky s bílým pozadím reprezentují třídy (modrý obrys – ontologie pro potřeby NaSaPO, šedý obrys – Sémantický slovník pojmů) nebo externí slovníky – součásti Sémantického slovníku pojmů (červený obrys). Instance (individuály) jako klíčové prvky ontologie pro potřeby NaSaPO jsou vyjádřené pomocí obdélníků se šedým pozadím. Všechny výše zmíněné objekty jsou propojené následujícími vazbami (viz obrázek 34):

- Instanční vazby (třída–instance) – modrá spojnice
- Specializační vazby (*skos:narrower*) – černá spojnice
- Asociační vazba (*skos:related*) – zelená spojnice
- Mapovací vazba (prvek ontologie NaSaPO – prvek v externím zdroji) – fialová spojnice
- Slovníková vazby (propojení mezi prvkem v externím zdrojem a relevantním slovníkem) – červená spojnice

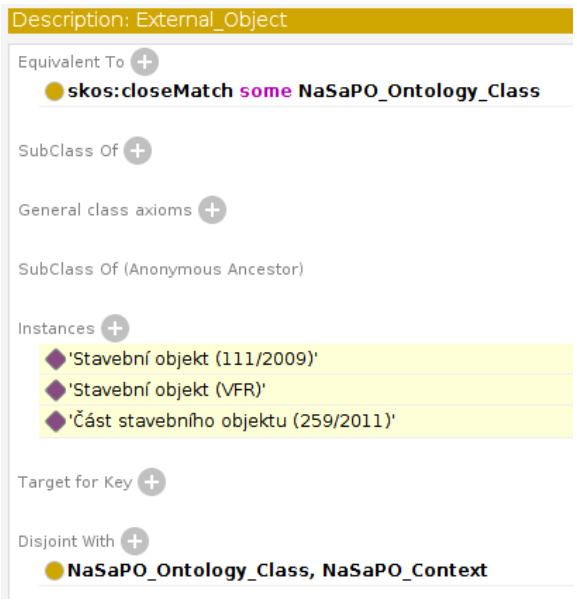


Obrázek 34: Ontologické vazby konceptu *Budova*

Výsledkem je ontologický model, který na základě implicitně dodaných informací (vazby mezi jednotlivými instancemi) a logických omezení vytvořených v rámci ontologického modelu je schopný automaticky odvodit (s využitým reasoneru FaCT++ 1.6.5) začlenění jednotlivých instancí do tříd (obrázky 35 a 36, kde žlutě podbarvené informace byly získány automaticky).



Obrázek 35: Popis třídy Stavba



Obrázek 36: Instance třídy External Object

Literatura

Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., & Kern, J. (2001). Manifesto for agile software development.

Bernaras, A. (1996). Building and reusing ontologies for electrical network applications. Proc. of ECAI 96, 1996, 298-302.

Bittner, T., & Winter, S. (2004). Geo-semantics and Ontology Extended abstract. In Proceedings of the Bentley Empowered Conference, Orlando, Florida.

Borst, W. N. (1999). Construction of engineering ontologies for knowledge sharing and reuse.

Brown, T. (2008). Design thinking. Harvard business review, 86(6), 84.

Cardoso, J. (2007). The semantic web vision: Where are we?. IEEE Intelligent systems, 22(5), 84-88.

Corcho, O., Fernández-López, M., & Gómez-Pérez, A. (2003). Methodologies, tools and languages for building ontologies. Where is their meeting point?. Data & knowledge engineering, 46(1), 41-64.

Cristani, M., & Cuel, R. (2005). A survey on ontology creation methodologies. International Journal on Semantic Web and Information Systems (IJSWIS), 1(2), 49-69.

Čerba, O. (2019). Participatory methodology for development of geo-ontologies: Initial study. In Advances and Trends in Geodesy, Cartography and Geoinformatics II: Proceedings of the 11th International Scientific and Professional Conference on Geodesy, Cartography and Geoinformatics (GCG 2019), September 10–13, 2019, Demänovská Dolina, Low Tatras, Slovakia (p. 171). CRC Press.

Eisenmann, T. R., Ries, E., & Dillard, S. (2012). Hypothesis-driven entrepreneurship: The lean startup. Harvard Business School Entrepreneurial Management Case, (812-095).

Fernández-López, M. (1999). Overview of methodologies for building ontologies. In IJCAI99 workshop on ontologies and problem-solving methods: Lessons learned and future trends (Vol. 430).

Gaševic, D., Djuric, D., & Devedžic, V. (2006). Model driven architecture and ontology development. Springer Science & Business Media.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge acquisition, 5(2), 199–221.

Cheah, C. (2007). Ontological Methodologies–From Open Standards Software Development to Open Standards Organizational Project Governance. IJCSNS, 7(3), 283.

Iqbal, R., Murad, M. A. A., Mustapha, A., & Sharef, N. M. (2013). An analysis of ontology engineering methodologies: A literature review. Research journal of applied sciences, engineering and technology, 6(16), 2993–3000.

Jones, D., Bench-Capon, T., & Visser, P. (1998). Methodologies for ontology development.

Miles A, Bechhofer S (2009). SKOS simple knowledge organization system reference. W3C recommendation.

Öhgren, A., & Sandkuhl, K. (2005). Towards a methodology for ontology development in small and medium-sized enterprises. In IADIS AC (pp. 369–376).

Ries, E. (2011). The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses. Currency.

Schreiber, G., Wielinga, B., & Jansweijer, W. (1995). The KACTUS view on the 'O'word. In IJCAI workshop on basic ontological issues in knowledge sharing (pp. 159–168).

Schwaber, K., & Beedle, M. (2002). Agile software development with Scrum (Vol. 1). Upper Saddle River: Prentice Hall.

Sugumaran, V., & Storey, V. C. (2002). Ontologies for conceptual modeling: their creation, use, and management. Data & knowledge engineering, 42(3), 251–271.

Uschold, M. (1996). Building ontologies: Towards a unified methodology. In 16th Annual Conf. of the British Computer Society Specialist Group on Expert Systems, Cambridge, UK.

Tento projekt je financován se státní podporou Technologické agentury ČR v rámci Programu BETA2.

www.tacr.cz

Výzkum užitečný pro společnost.

MINISTERSTVO VNITRA ČESKÉ REPUBLIKY







