

13.INSPIRE ESPUS školenie „Kvalita údajov a služieb“

INSPIRE kvalita údajov



Operačný program
**Efektívna
verejná správa**



Európska únia
Európsky sociálny fond



MINISTERSTVO

ŽIVOTNÉHO PROSTREDIA
SLOVENSKEJ REPUBLIKY



ESPUS

Efektívna správa priestorových údajov a služieb

13. INSPIRE ESPUS školenie „Kvalita údajov a služieb“ INSPIRE kvalita údajov

30.11.2022



Európska únia
Európsky sociálny fond

Prehľad

Motivácia

- Prečo sa venovať kvalite údajov
- Koho sa týka kvalita údajov

Legislatívne požiadavky a technické odporúčania

- Kvalita údajov v INSPIRE
- Dokumenty, pracovné skupiny

Komponenty kvality údajov

- Elementy kvality údajov





ESPUS

Efektívna správa priestorových údajov a služieb

Motivácia

Prečo sa venovať kvalite údajov

- Vyššia hodnota údajov = inovatívne riešenia
- Väčšie možnosti využitia/znovupoužitia údajov (interoperabilita)
- Jednoduchšia práca s údajmi ak poznáme ich kvalitu
- Veľké množstvo voľne dostupných údajov (stále narastá)
- Informácia o kvalite údajov by mala byť súčasťou každého datasetu

Kvalitné údaje = kvalitný GIS (IS)

Transformácia údajov podľa INSPIRE môže zlepšiť niektoré aspekty kvality dát :)

Koho sa týka kvalita údajov?

Poskytovateľ

- Plnenie legislatívnych požiadaviek
- Zvyšovanie kredibility poskytovaných dát

Ten kto vytvára (národný) SDI

- Schopnosť plniť požiadavky interoperability

Používateľ

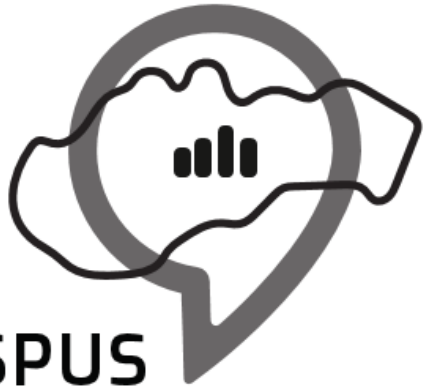
- Komfort využívania údajov
- Podpora rozhodovania



Kvalita údajov v INSPIRE

- Kvalita údajov, metaúdajov, sieťových služieb
- Požiadavky, súlad so smernicou (súradnicové systémy, povinné názvy, ...)
- Aplikačné schéma
- Kvalitné údaje = kvalitná služba
- Validné vs. kvalitné údaje





Efektívna správa priestorových údajov a služieb

Legislatívne požiadavky a technické odporúčania

Kvalita údajov GIS - štandardy



Definícia kvality podľa ISO (19157)

- *Stupeň, do akého súbor základných charakteristík spĺňa požiadavky*
- *Degree to which a set of inherent characteristics fulfils requirements*

ISO norma 19157 je orientovaná na výrobcu, čo sa definície kvality informácie týka (nie na spotrebiteľa)

Definícia kvality podľa INSPIRE

- *Súhrn charakteristík produktu, ktoré majú vplyv na jeho schopnosť uspokojiť stanovené a predpokladané potreby v súlade s EN ISO 19101*
- *Totality of characteristics of a product that bear on its ability to satisfy stated and implied needs in accordance with EN ISO 19101*



Kvalita údajov GIS v INSPIRE



Väzba na ISO v smernici INSPIRE

- ISO 19105:2005 Geographic information — Conformance and testing
- ISO 19106:2006 Geographic information — Profiles
- ISO 19115:2005 Geographic information – Metadata
- ISO 19131:2008 Geographic information — Data product specifications
- ISO 19157:2013 Geographic information — Data quality
- ISO 19101 Geographic information — Reference model

Kvalite údajov v INSPIRE sa venovali

- **INSPIRE Data Specification Drafting Team**
- **JRC Data Specification Support Team**
- **INSPIRE Thematic Working Groups**
- **INSPIRE the Data Quality Expert Group** - zložený zo špecialistov z jednotlivých štátov - *DQ contact points*, *DQCP* mali vytvoriť diskusiu na národnej úrovni so zapojením príslušných orgánov a výsledky z národnej úrovne boli zahrnuté v technical guidelines

Kvalita údajov GIS v INSPIRE - dokumenty



- **Smernica**
- **Dátové špecifikácie (kapitoly 7 a 8)**
- **JRC Technical Report - Data Quality in INSPIRE: Balancing Legal Obligations with Technical Aspects**

https://inspire.ec.europa.eu/documents/INSPIRE_/JRC83209_Online_Data_quality_in_INSPIRE.pdf

- **Data Quality training module**

<https://inspire.ec.europa.eu/training/data-quality>

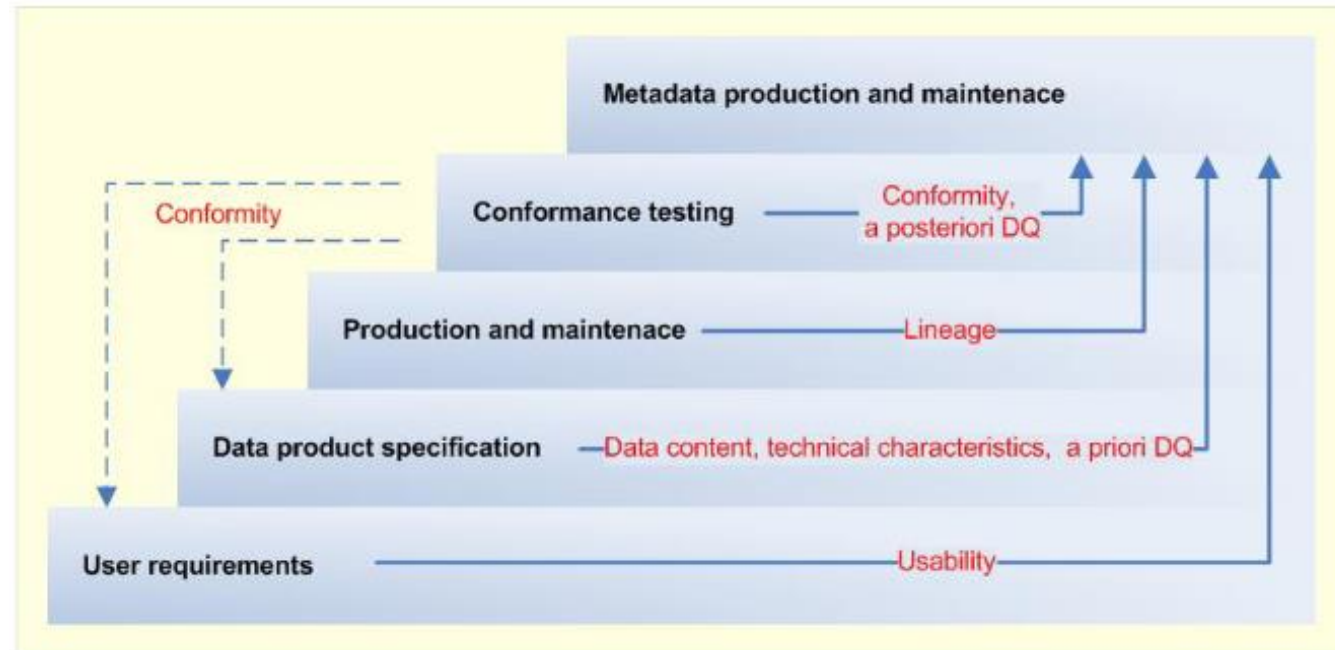
- **Data quality in INSPIRE: from requirements to metadata**

<https://docplayer.net/19206301-Data-quality-in-inspire-from-requirements-to-metadata.html>

Tvorba kvality údajov



Postupná akumulácia informácií o kvalite dát pri tvorbe údajov

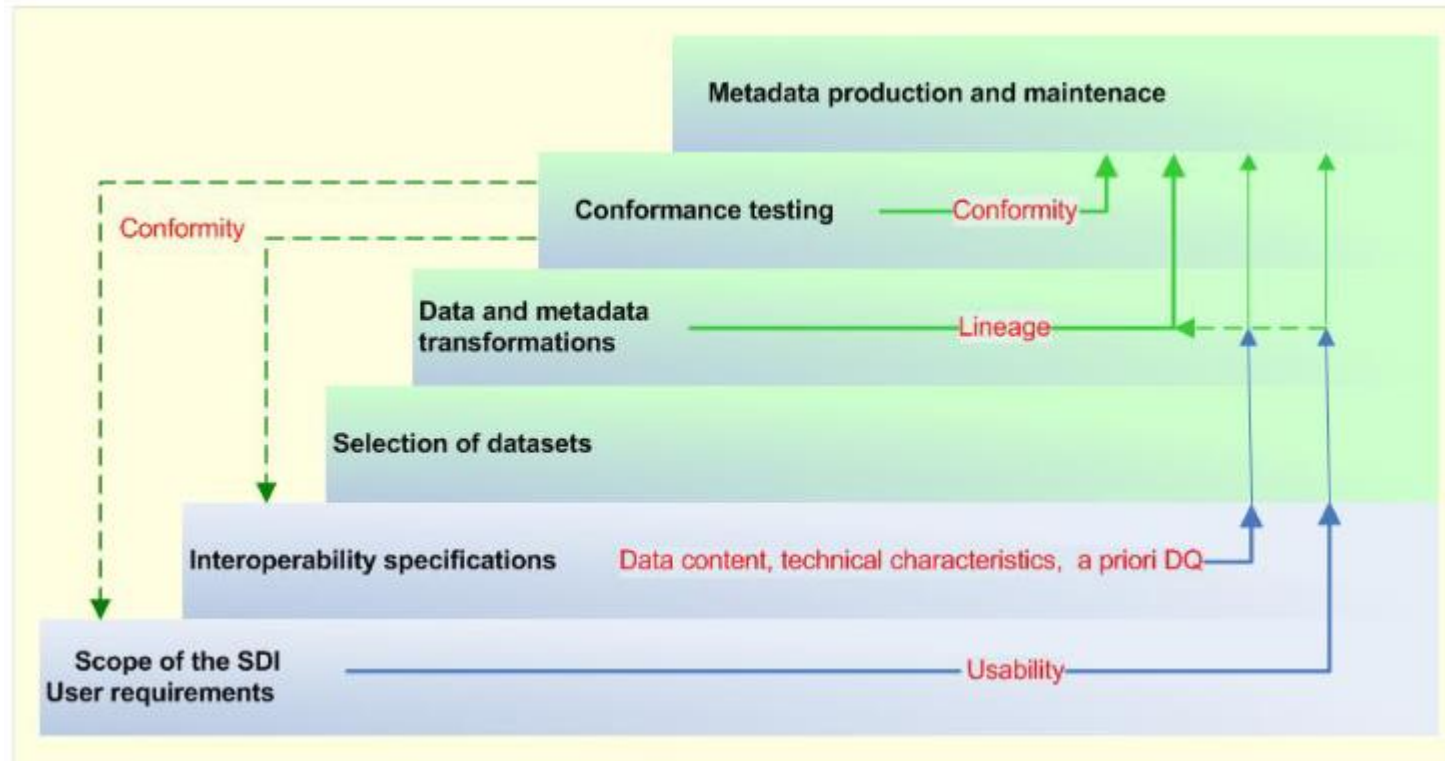


Štandardný postup získavania údajov. Každý krok by mal sprevádzať postup zabezpečenia kvality údajov, ktorý je dokumentovaný ako metaúdaje. Ideálne metaúdaje o kvalite údajov inkrementálne s každou fázou narastajú.

Tvorba kvality údajov v SDI



Publikovanie údajov v SDI



INSPIRE ako SDI

Tvorba kvality údajov v SDI



User requirements - optimálny pomer požiadavky vs. náklady

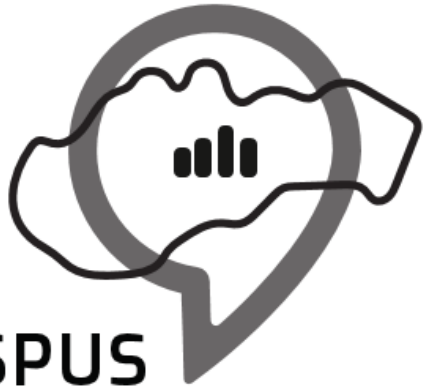
Interoperability specifications - špecifikácia obsahu, technických požiadaviek, ... (ciele interoperability), *a priori* požiadavky

Selection datasets - reálny výber údajov podľa špecifikácií a cieľov interoperability

Data and Metadata transformations - najväčší dôraz na kvalitu údajov v zmysle jej zachovania (výber nástrojov a postupov pre transformáciu)

Conformance testing - testovanie a porovnanie *a priori* požiadaviek s výslednými údajmi, výsledok porovnania je evidovaný ako *a posteriori* metaúdajové elementy

Metadata production and maintenance - kompletný súbor metaúdajov prístupný používateľom ako súčasť SDI



ESPUS

Efektívna správa priestorových údajov a služieb

Komponenty kvality údajov

Komponenty kvality údajov

Kvalita údajov

- Vyjadrená prvkami kvality údajov
- Vykazovaná v správach o kvalite údajov alebo v metaúdajoch
- Týka sa konkrétnej geografickej oblasti a konkrétneho súboru údajov

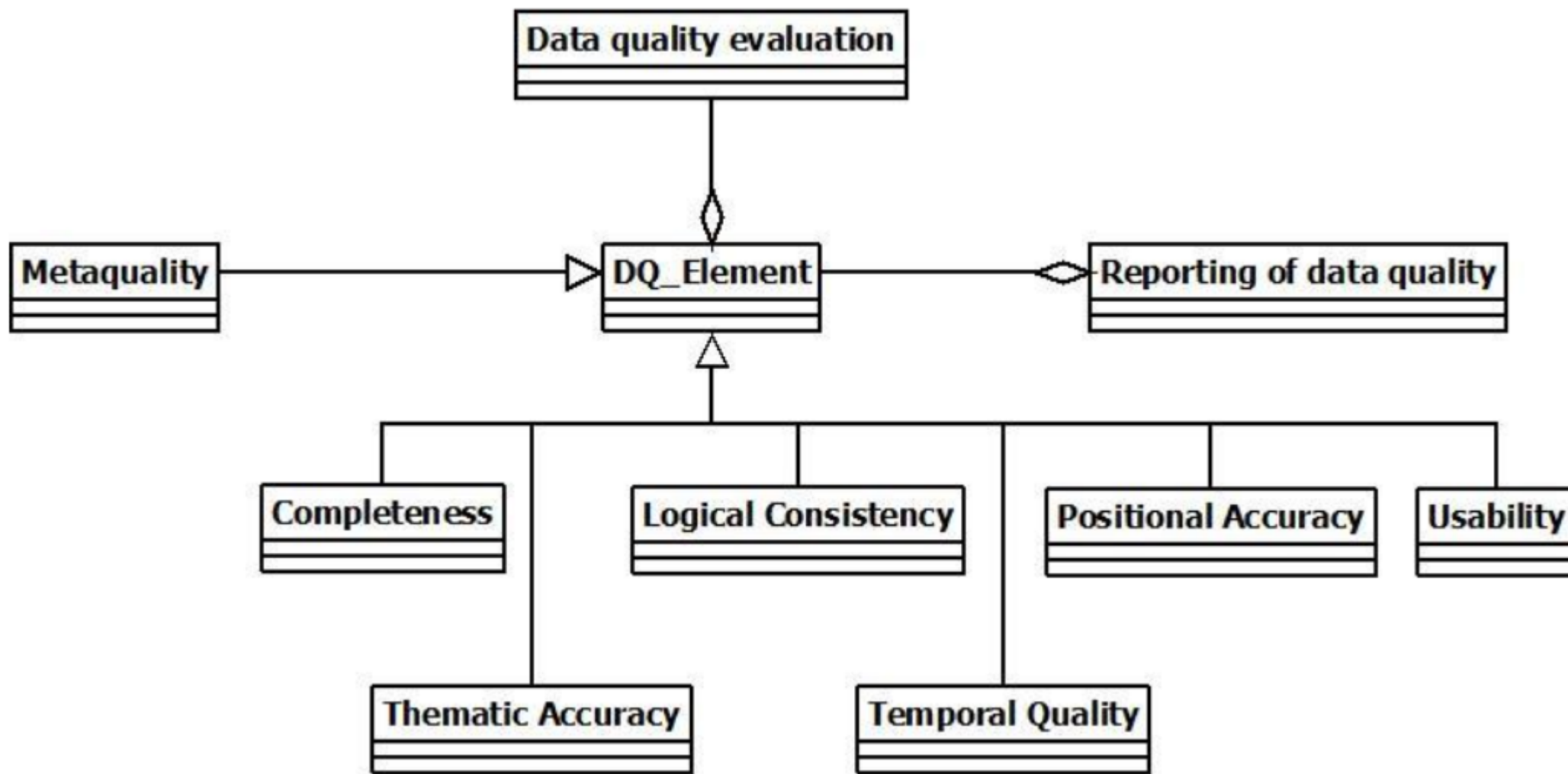
Element kvality údajov je popísaný:

- Definíciou
- Meradlami kvality údajov
- Výsledkami kvality údajov
- Metakvalitou

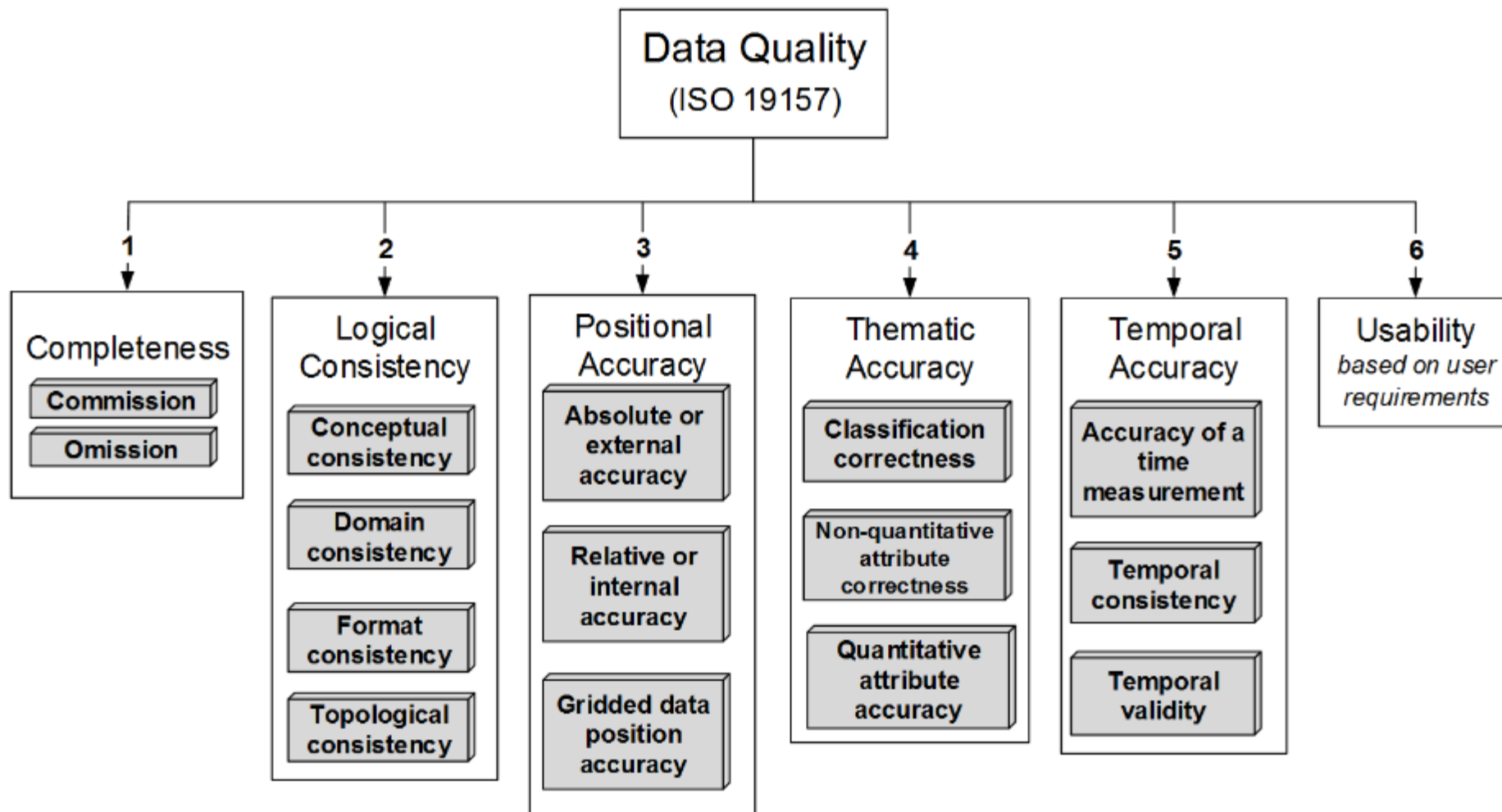
Elementy kvality údajov:

- *Úplnosť*
- *Logická konzistencia*
- *Polohová presnosť*
- *Tematická presnosť*
- *Časová kvalita*
- *Použitelnosť*

Komponenty kvality údajov



INSPIRE Technical Guidelines use ISO 19157 *Geographic Information-Data quality*



Úplnosť

Definícia: *Prítomnosť a absencia objektov, ich atribútov a vzťahov.*

Úplnosť sa často meria ako percento vynechaných položiek, napr.:

V súbore údajov je zastúpených 95 % všetkých chránených mokradí

70 % údajov o chránených mokradiach má platného vlastníka pozemku

Subelementy:

Nadbytočnosť

- Prítomnosť nadbytočných údajov v datasete.

Vynechanie

- Chýbajúce údaje v datasete.

Logická konzistencia

Definícia: *Miera dodržiavania logických pravidiel údajovej štruktúry, priradených atribútov a vzťahov (štruktúra údajov môže byť konceptuálna, logická alebo fyzická).*

Subelementy:

Konceptuálna konzistencia

- Dodržiavanie pravidiel konceptuálnej schémy.

Doménová konzistencia

- Dodržiavanie doménových hodnôt.

Konzistencia formátu

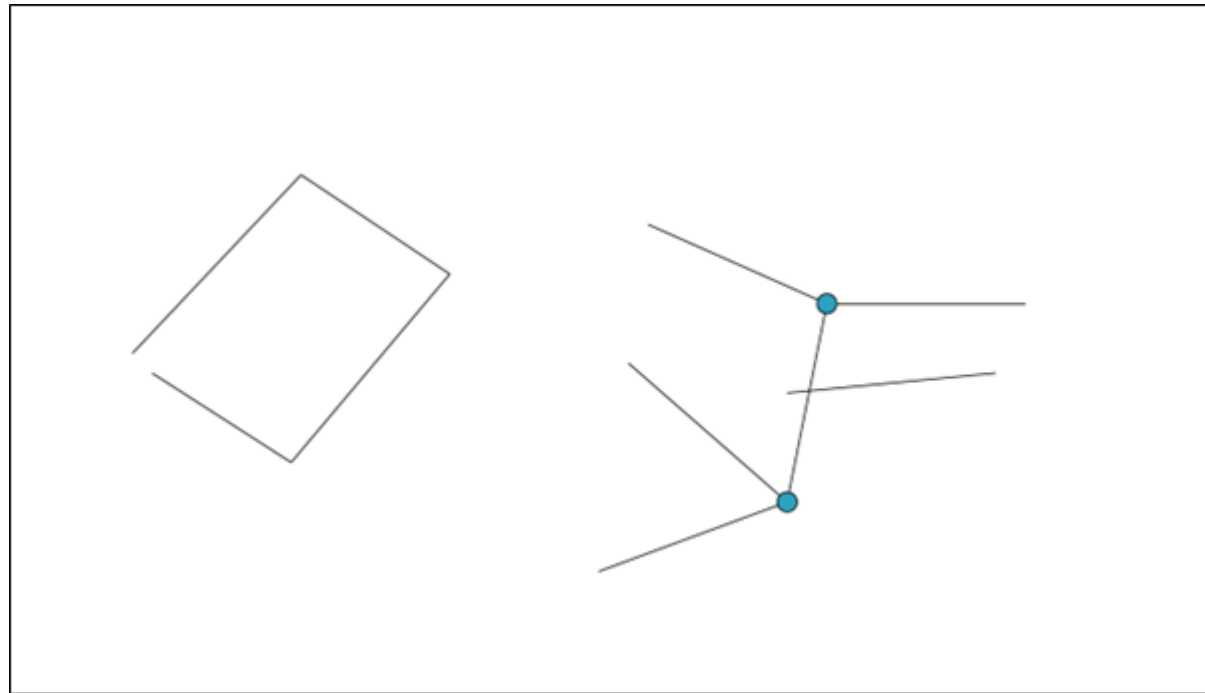
- Miera, do akej sú údaje v súlade s fyzickou štruktúrou datasetu.

Topologická konzistencia

- Správnosť explicitne kódovaných topologických charakteristík datasetu.

Logická konzistencia

Príklad: topologická konzistencia, kontrola v GIS softvéroch



Polohová presnosť

Definícia: *Presnosť polohy prvkov v rámci priestorového referenčného systému.*

Presnosť = blízkosť zhody medzi výsledkom testu a akceptovanou referenčnou hodnotou [ISO 3534-1]

Subelementy:

Absolútna alebo externá presnosť

- Blízkosť nahlásených súradnicových hodnôt k hodnotám, ktorú sú akceptované ako pravdivé alebo sú pravdivé.

Relatívna alebo interná presnosť

- Blízkosť relatívnych pozícií prvkov v datasete k ich príslušným relatívnym pozíciám, ktoré sú akceptované ako pravdivé.

Polohová presnosť gridových dát

- Blízkosť hodnôt priestorovej polohy gridových údajov k hodnotám, ktoré sú akceptované ako pravdivé alebo sú pravdivé.

Polohová presnosť

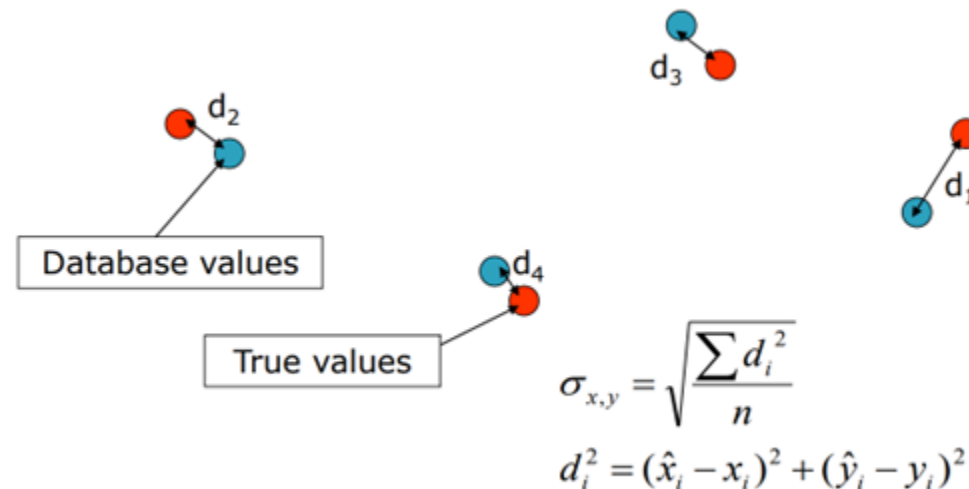
Príklad: Polohová presnosť sa často vyjadruje ako stredná kvadratická hodnota.

$$\sigma = \sqrt{\frac{\sum (\hat{x}_i - x_i)^2}{n}}$$

kde:

\hat{x}_i – sú hodnoty v databáze

x_i – sú zodpovedajúce „pravdivé“ hodnoty



Tematická presnosť

Definícia: Presnosť kvantitatívnych atribútov a správnosť nekvantitatívnych atribútov a klasifikácií objektov a ich vzťahov.

Subelementy:

Správnosť klasifikácie

- Porovnanie tried priradených k objektom alebo ich atribútom.

Správnosť nekvantitatívnych atribútov

- Miera, či je nekvantitatívny atribút správny alebo nesprávny.

Presnosť kvantitatívneho atribútu

- Blízkosť hodnoty kvantitatívneho atribútu hodnote, ktorá je akceptovaná alebo známa ako pravdivá.

Tematická presnosť

Príklad: Tematická presnosť sa často vzťahuje na nominálne údaje (napr. typ objektu) alebo binominálne údaje (áno/nie, zapnuté/vypnuté, ...). Tematická presnosť binominálnych údajov býva zvyčajne vyjadrená ako chybovosť. Pri nominálnych údajoch sa vyjadruje buď ako chybovosť alebo pomocou matice chybenej klasifikácie.

	Water	Forest	Urban	Agri		
Water	90	2	0	1	97%	
Forest	5	78	2	8	84%	
Urban	0	1	52	1	96%	
Agri	2	5	0	94	93%	
	93%	91%	96%	90%	92%	

Časová kvalita

Definícia: Kvalita časových atribútov a časových vzťahov objektov.

Subelementy:

Presnosť merania času

- Blížkosť nahlásených meraní času k hodnotám, ktoré sú akceptované ako pravdivé

Časová konzistentnosť

- Správnosť poradia udalostí

Časová platnosť

- Platnosť údajov s ohľadom na čas (napr. 33. marec nie je platný dátum)

Použitelnosť

Definícia: Stupeň súladu datasetu s konkrétnym súborom požiadaviek.

Používa sa na opis špecifických informácií o vhodnosti datasetu pre konkrétnu aplikáciu alebo zhode so súborom požiadaviek.

Súvis s prvým krokom - ***User requirements***

The main requirement for usability is interoperability

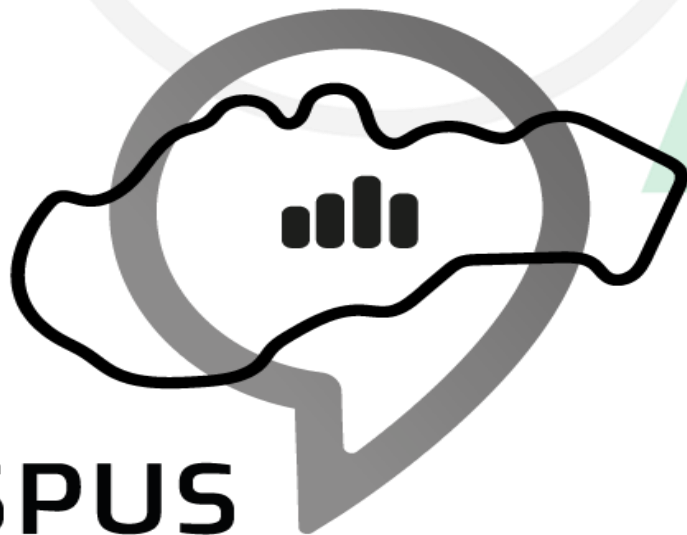
Art.	Citation	Related DQ (sub-)element	DQ evaluation against the
5(2)	Metadata shall include information on the quality and validity of spatial data sets	All relevant to the data set	Application schema of the data theme Implementing Rule / Data specification of the data theme
7(3)	Member States shall ensure that [...] spatial data sets and the corresponding spatial data services are available in conformity with the implementing rules [...]	DQ_ConceptualConsistency	Application schema of the data theme
7(4)	Implementing rules [...] shall cover the definition and classification of spatial objects relevant to spatial data sets related to the themes listed in Annex I, II or III and the way in which those spatial data are geo-referenced.	DQ_ConceptualConsistency DQ_ThematicClassificationCorrectness DQ_TopologicalConsistency	Application schema of the data theme
8(1), (2)	In the case of spatial data sets corresponding [...] the themes listed in Annex I or II [...] the implementing rules shall address the following aspects		
	- framework for the unique identification	DQ_DomainConsistency	GCM
	- the relationship between spatial objects	DQ_ConceptualConsistency	Application schema of the data theme
	- key attributes [...]	DQ_ConceptualConsistency DQ_NonQuantitativeAttributeAccuracy DQ_QuantitativeAttributeAccuracy	Application schema of the data theme
	- information on the temporal dimension	DQ_TemporalConsistency DQ_TemporalValidity	Application schema of the data theme GCM
8(3)	[...] consistency between items of information which refer to the same location	DQ_ConceptualConsistency	Application schema of the data theme (multi-scale representation)
	or between items of information which refer to the same object represented at different scales	(DQ_PositionalAccuracy)	Data specifications of the related data themes
10(2)	In order to ensure that [...] a geographical feature, the location of which spans the frontier between two or more Member States, are coherent, Member States shall, [...] decide by mutual consent on the depiction and position of such common features.	DQ_LogicalConsistency DQ_PositionalAccuracy	Agreement between the interested parties



Ďakujem za pozornosť!

Mgr. Marcel Récky
marcel.recky@zymestic.sk

ESPUS



ESPUS

Efektívna správa priestorových údajov a služieb

<https://inspire.gov.sk/projekty/espus>