

Predstavitev evropskega standarda za metapodatke

Izvleček

Metapodatki ali podatki o podatkovnih nizih so izpeljane informacije o pomenu, zgradbi, vsebini, kakovosti, predstavitvi, obsegu (geometričnem in časovnem), referenčnem sistemu, administraciji in uporabi shranjenih podatkov. Formalno oziroma minimalno opredelitev metapodatkov podaja predlog evropskega standarda (EN) za metapodatke (CEN prEN 12657:1996). Metastandard za geografske podatke določa osnovna vsebinska in formalna pravila za opis prostorskih podatkov.

Ključne besede: CEN TC 287, geografski podatek, metapodatek

Abstract

Metadata or data about datasets is specific information about the context, structure, content, quality, presentation, extent (both geometric and temporal), spatial reference, administration and usage of stored geographic data. Formal and minimal specification of metadata is defined by the new European standard for metadata (CEN prEN 12657:1996). The metastandard for geographic data specifies the basic content and the formal rules for describing spatial data.

Keywords: CEN TC 287, geographic data, metadata

1 UVOD

Načeloma spadajo vsi geografski podatki v skupno kategorijo, vendar jih lahko vsebinsko oziroma funkcionalno razdelimo v dve skupini:

- ☐ geografski (ali prostorski) podatki
- ☐ metapodatki (ali podatki o geografskih podatkovnih nizih).

Geografske informacije so informacije o fenomenih, ki so neposredno ali posredno povezani z lokacijo, nanašajočo se na površino Zemlje. Geografski podatki so računalniško obdeljiva oblika geografskih informacij. Geografske (ali prostorske) podatke lahko opredelimo kot podatke o opisnih, časovnih in kartografskih lastnostih ter odnosih med geografskimi objekti, katerih lokacija je podana v enotnem georeferenčnem sistemu. Geografski podatkovni niz je razpoznavna zbirka prostorskih podatkov, ki je lahko sestavljena iz raznih vsebinskih podnizov.

Možnost posredovanja geografskih podatkov med nepoznanimi uporabniki po omrežju se je pojavila v zadnjih letih predvsem zaradi naglega tehnološkega razvoja komuniciranja. Povečane potrebe po izmenjavi, porazdeljevanju, delitvi in povezovanju geografskih podatkov so sprožile veliko zanimanje tudi za metapodatke. Metapodatki opisujejo tehnične značilnosti geografskih podatkov in njihove poslovne vidike. Metapodatki so tudi potrebni, ker zagotavljajo možnost ponovne oziroma

večkratno uporabo istih podatkovnih nizov. Metapodatki vsebujejo pomembne informacije, ki dovoljujejo uporabnikom, da ocenijo primernost geografskih podatkov za njihove lastne tehnične, pravne, administrativne in poslovne zahteve. Metapodatki so razdeljeni v različne kategorije, kot so na primer opisi:

- ☐ identifikacije in lastništva
- ☐ podatkovne vsebine in zgradbe
- ☐ dostopnosti in dostave.

2 METAPODATKI

Metapodatki so izpeljane informacije o pomenu, zgradbi, vsebini, obsegu, kakovosti, zgodovini, organizaciji, dostopnosti, vrednosti in uporabi shranjenih podatkov.

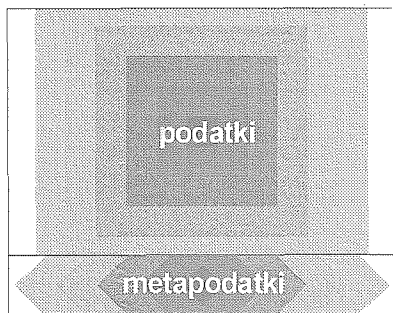
Metapodatki predstavljajo standardni opis značilnosti geografskih podatkovnih nizov. Metapodatki se lahko razdelijo na administrativno skupino in tehnično skupino elementov. Prva skupina metapodatkov, ki tvori tako imenovano zunanjo raven, je predvsem pomembna pred nakupom podatkovnega niza. Uporaba skupine tehničnih elementov metapodatkov, ki predstavlja njihovo interno raven, pa prevladuje po dejanskem nakupu geografskih podatkov.

Metapodatki torej opisujejo geografske podatkovne nize na način, ki omogoča oceno primernosti uporabe določenega podatkovnega niza za načrtovano vrsto rabe.

Metapodatki zato podajajo predvsem naslednje značilnosti geografskih podatkov v obravnavanem nizu:

- ☐ referenčni sistem
- ☐ geografski in časovni obseg podatkovnega niza
- ☐ ažurnost oziroma aktualnost podatkov v nizu
- ☐ povzetek kakovostnih elementov za podatkovni niz (poreklo, predhodna uporaba, natančnost, popolnost, logična usklajenost, ažurnost itd.)
- ☐ administrativne podatke (avtorstvo, lastništvo, prenosljivost, dostopnost, cena itd.).

3 METAPODATKOVNI STANDARD IN METABAZA



Formalna (minimalna) opredelitev metapodatkov mora biti prav tako standardizirana v splošnem standardu za geografske podatke (SDTS, CEN TC 287 itd.) ali pa v posebnem metapodatkovnem standardu, kot je denimo SDGM (Standard for Digital Geospatial Meta-data), ki ga je kot prvi metastandard razvil FGDC (Federal Geographic Data Committee). Metastandard za geografske podatke opredeljuje vsebinska in formalna pravila za opis geografskih podatkov.

Metastandard služi za standardizacijo metapodatkov (standardizacija standarda).

Namen metastandarda je torej poenotenje metapodatkovnih opisov s ciljem informiranja uporabnikov pri izbiri, uporabi in obdelavi geografskih podatkov.

Metabaza je podatkovna baza, ki vsebuje metapodatke in posreduje informacije o drugih podatkovnih bazah. Namen metabaze je, da potencialnim uporabnikom

podatkov posreduje pregled razpoložljivih baz podatkov, njihovo vsebino in zgradbo. Metabaza je torej neka vrsta "rumenih strani" o bazah geografskih podatkov.

4 EVROPSKI STANDARD ZA METAPODATKE

4.1 Uvod

Evropski standard za metapodatke, uradno CEN prEN 12657:1996 Geografske informacije – Opis podatkov – Metapodatki, določa konceptualno shemo metapodatkov za geografske podatkovne nize. Metapodatki so tehnični podatki o podatkovnih nizih in njihovih poslovnih vidikih. Vsebujejo informacije o vsebini, predstavi, obsegu (geometričnem in časovnem), prostorskem referenčnem sistemu, kakovosti in administraciji podatkovnega niza.

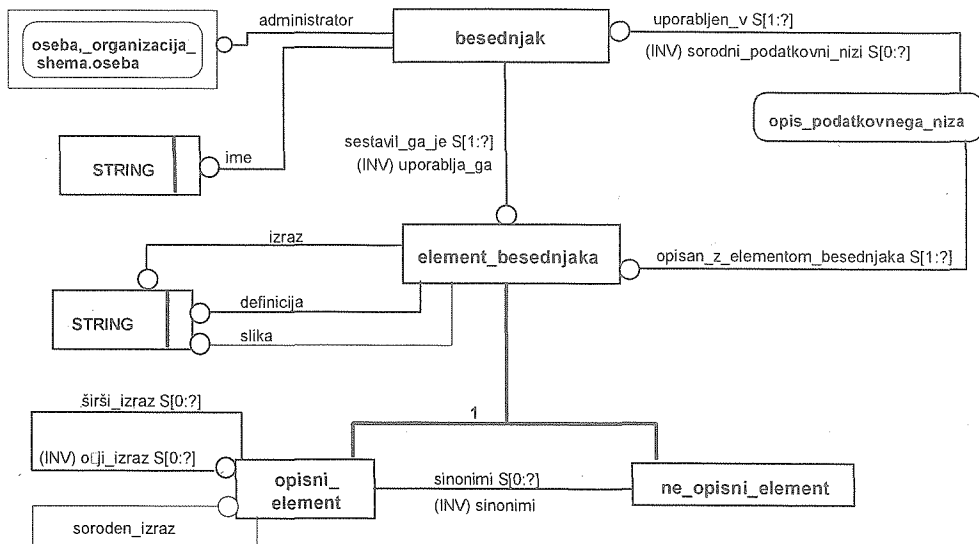
Evropski standard za metapodatke določa tiste postavke, ki so obvezne za opis podatkovnih nizov in te predstavljajo minimalni niz metapodatkov. Standard podaja tudi dodatne elemente in primere, kako se lahko uporabljajo, vendar pa se ne ukvarja z izvedbenimi podrobnostmi oziroma z zgradbo baz podatkov za metapodatke. Načrtovan je predvsem za uporabo z digitalnimi geografskimi nizi podatkov, vendar pa se sama načela lahko uporabijo tudi za opis geografskih informacij v drugih oblikah, kot so papirne karte ali sezname.

Metapodatki opisujejo strukturo geografskih podatkov za podporo iskanju in urejanju podatkov. To omogoča izdelovalcu podatkovnih nizov izdevo njihovega opisa na dosleden način ter omogoča uporabnikom ocenitev uporabnosti in primernosti za njihove posebne namene. Pomanjkanje zavesti o obstoječih podatkovnih nizih lahko vodi do premajhne ali pa napačne uporabe geografskih informacij. Osnovni razlog za razvoj metastandarda je povečati uporabo geografskih podatkov. Končni cilj metapodatkov je zagotoviti učinkovite ter zanesljive mehanizme za prenos geografskih podatkov, in sicer brez izgube informacij.

4.2 Sestava evropskega standarda za metapodatke

Predlagani evropski standard za metapodatke sestavlja šest delov. V prvih treh poglavjih so opis obsega celotnega razvojnega projekta, napotki na druge sorodne standarde in pregled definicij za uporabljeno terminologijo. Za razumevanje metapodatkovnega standarda je treba podrobno poznati nekatere sorodne standarde iz iste skupine, kot so zlasti Pregled, Referenčni model, Prostorska shema, Kakovost, Položaj in tudi Prenos. Potrebno je tudi predznanje jezika Express, ki je namenjen za informacijsko modeliranje. Osnovni sestavini informacijskega modeliranja v jeziku Express sta predvsem naslednji:

- a) Leksikalni jezik Express (ISO 10303-11) je formalni jezik, ki služi za opredelitev konceptualne sheme oziroma uporabniškega modela.
- b) Grafična notacija (Express-G) se uporablja za slikovno predstavitev konceptualnih shem podatkovnega modela in je usklajena z jezikom Express.



Osrednji del metastandarda tvori podroben opis sestave metapodatkov, ki so razdeljeni v devet skupin. Za vsako pomembno skupino metapodatkov je podana tudi ustrezna slika v Express-G diagramu. Skupine metapodatkov pa so naslednje:

- ☐ identifikacija podatkovnega niza
- ☐ pregled podatkovnega niza
- ☐ elementi kakovosti podatkovnega niza
- ☐ prostorski referenčni sistem
- ☐ obseg
- ☐ definicija podatkov
- ☐ klasifikacija
- ☐ administrativni metapodatki
- ☐ metapodatkovna referenca.

Vsi metapodatkovni elementi so opredeljeni z imenom in s kratkim opisom pomena. Za primer je na zgornji sliki prikazan Express-G grafični prikaz strukture in povezav skupine metapodatkovnih elementov z imenom Klasifikacija. Sledi opredelitev te skupine elementov v jeziku Express.

4.3.7 Klasifikacija

4.3.7.1 Besednjak

Če se v podatkovnem nizu poda besednjak, potem naj se za takšen besednjak podajo naslednji podatki:

- ☐ Ime besednjaka naj podaja ime obstoječega standardnega ali pa uporabniško specifičnega besednjaka.
- ☐ Upravitelj besednjaka naj podaja ime organizacije ali osebe, ki je odgovorna za vzdrževanje besednjaka.
- ☐ Elementi besednjaka: (glej poglavje 4.3.7.2).

Naslednja preglednica prikazuje izvleček iz Dodatka A za skupino elementov z imenom Klasifikacija.

metapodatki	opis	omejitev	kardinalnost	podatkovni tip
Classification		klasifikacija		P
Thesaurus		besednjak		O
Name of thesaurus	Ime besednjaka podaja naziv slovarja, bodisi obstoječega standardnega ali pa posebnega uporabniškega besednjaka.	O	1	podatkovni niz
Thesaurus administrator	Upravitelj besednjaka je ime organizacije ali osebe, ki je odgovorna za vzdrževanje besednjaka.	O	1	podatkovni niz
Thesaurus element		element besednjaka		O
Term	Izraz je objektni, opisni ali asociacijski tip na katerikoli stopnji generalizacije.	O	1	podatkovni niz
Definition	Opredelitev opredeljuje pomen izraza.	O	1	podatkovni niz
Synonyms	Sinonimi so pomensko enakovredni izrazi.	O	m	podatkovni niz
Related term	Soroden izraz podaja pomensko zelo podoben izraz (denimo letališče in letališka steza).	O	m	podatkovni niz
Broader term	Širši izraz podaja splošnejši pomen (denimo cestni promet in promet).	O	m	podatkovni niz
Narrower term	Ožji izraz podaja bolj specializiran pomen (denimo avtocesta kot vrsta ceste).	O	m	podatkovni niz
Picture	Slika prikazuje grafično ilustracijo izraza.	P	1	vektorska ali rastrska datoteka

4.3.7.2 Element besednjaka

Če se v podatkovnem nizu poda besednjak, potem naj se za vsak element besednjaka podajo naslednji podatki:

- ☐ izraz lahko podaja objektni tip ali katerikoli drugo raven posploševanja, na primer cesta, cestišče, promet itd.
- ☐ opredelitev – opredeljuje enolični pomen izraza
- ☐ sinonim podaja izraz, ki je pomensko enakovreden
- ☐ soroden izraz podaja pomensko zelo podoben izraz
- ☐ širši izraz podaja splošnejši pomen
- ☐ ožji izraz podaja bolj podroben pomen.

O elementu besednjaka se lahko podajo tudi naslednji dodatni podatki:

- slika prikazuje grafično ilustracijo izraza.

Metastandardu so dodani tudi štirje dodatki. Dodatek A je preglednica, ki sistematično podaja vse metapodatkovne elemente s kratkim opisom, omejitvami, kardinalnostjo in podatkovnim tipom. Opis podaja kratko definicijo metapodatkovnega elementa. Omejitve podajajo, kateri podatkovni tipi so obvezni oziroma izborni. Kardinalnost podaja pojavnost vrednosti elementa in sicer kot ena (1) ali mnogo (m). Podatkovni tip je lahko število, datum, znakovni niz, slika, naslov, našeta vrednost, rastrska ali vektorska datoteka itd.

Dodatek B podaja podrobno specifikacijo metapodatkovnih elementov v jeziku Express. V nadaljevanju je primer kode za skupino elementov z imenom Klasifikacija.

```
ENTITY Thesaurus;                -- Classification (see clause 4.3.7)
name_of_thesaurus                : STRING;
thesaurus_administrator          : person;
built_up_by                      : SET [1:?] OF Thesaurus_element;
INVERSE
referenced_datasets              : SET OF Dataset_description FOR referred_by_
                                thesaurus;

UNIQUE
url: name;
END_ENTITY; -- Thesaurus
ENTITY Thesaurus_element SUPERTYPE OF (ONEOF(descriptive_element,
non_descriptive_element));
term                             : STRING;
definition                      : STRING;
picture                         : OPTIONAL External_graphic_file;
INVERSE
used_by                         : Thesaurus FOR built_up_by;
END_ENTITY; -- Thesaurus_element
ENTITY descriptive_element SUBTYPE OF (Thesaurus_element);
synonyms                       : SET OF non_descriptive_element;
related_term                   : OPTIONAL descriptive_element;
broader_term                   : SET OF descriptive_element;
INVERSE
narrower_term                  : SET OF descriptive_element FOR broader_term;
END_ENTITY; -- descriptive_element
ENTITY non_descriptive_element SUBTYPE OF (Thesaurus_element);
INVERSE
synonyms                       : SET OF descriptive_element FOR synonyms;
END_ENTITY; -- non_descriptive_element
```

Dodatek C podaja pregled celotne Express-G sheme za vse metapodatkovne elemente standarda. Dodatek D prikazuje uporabo metastandarda v dveh enostavnih primerih.

5 SKUPINA STANDARDOV ISO 10303 (STEP)

ISO 10303:1994 je mednarodni standard (uradni naslov: Industrial automation systems and integration – Product Data Representation and Exchange), ali krajše STEP. ISO 10303 je dejansko obsežna skupina standardov za avtomatizacijo in povezovanje industrijskih sistemov. Namen skupine standardov je zagotoviti računalniško podprto opredelitev, predstavitev in izmenljivost podatkov o izdelkih. Osnovni cilj skupine standardov je zagotoviti nevtralne mehanizme, ki so primerni za opisovanje digitalnih podatkov o izdelkih skozi njihov celotni življenjski cikel in so hkrati neodvisni od kateregakoli računalniškega sistema. Značilnosti in lastnosti takega opisa zagotavljajo ne samo primernost za neodvisno izmenjavo podatkov, temveč so tudi osnova za delitev podatkov in organizacijo ustreznih baz podatkov.

Standard ISO 10303 (STEP) je organiziran kot serija neodvisno razvitih in objavljenih standardov. Skupine standarda 10303 so naslednje: opisne metode, integrirani viri, aplikacijski protokoli, testna garnitura, izvedbene metode in testiranje prilagojenosti. Za aplikacije tehnologije GIS-ov so praktično pomembni predvsem naslednji deli družine standardov ISO 10303:

- 1. del (ISO 10303-1): pregled osnovnih načel
- 11. del (ISO 10303-11): opisne metode (priročnik za formalni opisni oziroma leksikalni jezik Express)
- 21. del (ISO 10303-21): služi za kodiranje v jeziku Express podanih besedil v ustrezno fizično datotečno strukturo
- 31. del (ISO 10303-31): testiranje kakovosti podatkov
- 42. del (ISO 10303-42): opis in podpora geometrični in topološki predstavitvi.

5.1 Leksikalni jezik Express (standard ISO 10303-11:1994)

Konceptualno modeliranje je namenjeno izdelavi konceptualnega podatkovnega modela stvarnosti, ki se formalno opredeli z ustreznimi konceptualnimi shemami. Te morajo biti strojno in programsko neodvisne od kateregakoli računalniškega sistema, da lahko predstavljajo ustrezno osnovo za sistemsko neodvisno izvedbo. Učinkovit opis vsebine in strukture podatkovnih shem zahteva uporabo ustreznega formalnega jezika, ki ga je mogoče uporabljati v različnih računalniških okoljih. Tak leksikalni jezik mora biti neodvisen od kateregakoli programskega ali strojnega sistema, da lahko zagotavlja celovitost in izvedbeno usklajenost različnih aplikacij in uporabniških zahtev.

Leksikalni jezik Express je bil v zadnji (drugi) različici leta 1994 (ISO/TC 184) opredeljen kot mednarodni standard (ISO 10303-11). Express se je začel razvijati okoli leta 1986 kot jezik za podatkovno modeliranje. Express je osrednji del skupine standardov ISO s formalno oznako 10303 in skrajšanim imenom STEP, ki so namenjeni opredelitvi splošnih pravil, načel in pripomočkov za izmenjavo načelno kakršnihkoli digitalnih podatkov o izdelkih med računalniškimi sistemi. Jezik Express je bil v preteklosti intenzivno v rabi zlasti pri razvijanju različnih modelov za CAD, zlasti v aeronavtiki, strojni in elektronski industriji. Pomen tega leksikalnega jezika se je še posebej povečal, ko je bil leta 1996 izbran kot formalni jezik družine evropskih standardov, ki jih je razvil CEN-ov tehnični odbor 287.

5.2 Značilnosti jezika Express

Jezik Express je namenjen za informacijsko modeliranje. Express je neproceduralni opisni jezik za podajanje podrobnega in celovitega opisa konceptualnega podatkovnega modela, ki je neodvisen od kateregakoli programskega jezika ali tehnologije baz podatkov. Osnovne značilnosti jezika Express so predvsem naslednje:

- neodvisnost od kateregakoli računalniškega sistema, čeprav je mogoče Express praktično uporabljati v vseh računalniških sistemih,
- mogoče je opredeliti vse sestavine konceptualnih shem, kot so objektni (ali entitetni) tipi, opisi, relacije, globalna uskladivena in celovitostna pravila in procesne funkcije,
- kot usklajeno dopolnilo jeziku obstaja tudi posebna grafična notacija (Express-G), ki jo lahko uporabljamo za grafično modeliranje in prikazovanje konceptualnih shem,
- na voljo so tudi komercialna orodja CASE in avtomatski generatorji jezikovnega zapisa Express, ki na podlagi grafičnega prikaza konceptualnih shem samostojno tvorijo ustrezno kodo Express,
- jezik je modularen, kar omogoča nadgradnjo, povezave in sestavljanje konceptualnih shem iz posameznih modulov (shem), ki so ustrezno povezani,
- jezik je primeren in v celoti opremljen tudi s posebnimi mehanizmi za opredeljevanje geometrijskih in topoloških vidikov prostorskih podatkov.

V jeziku Express je shema osnovni modul za opis podatkovne strukture. V shemi opredelimo entitetne tipe, ki so glavni objekti pri opisu podatkovne zgradbe in sestave. Entitetne tipe določajo izbrani in imenovani opisi, ki predstavljajo pomembne lastnosti skupine objektov. Vsak opis je lahko vnaprej določenega podatkovnega tipa, kot so: enostavni osnovni podatkovni tipi (cela in realna števila, znaki itd.), sestavljeni osnovni podatkovni tipi (polja, nizi, datumi, sezname itd.) ali pa entitetno sestavljeni podatkovni tip, ki ga podajajo drugi objektni tipi (uporabniško določeni podatkovni tipi).

V sedANJI različici jezika Express opazimo tudi nekatere objektno usmerjene značilnosti. Enostavno dedovanje lahko opredelimo s supertipom in podtipom, kar omogoča sestavo ustreznih hierarhičnih podatkovnih zgradb. Express v obstoječi izvedbi ni v celoti objektno usmerjen jezik, ker ne omogoča uporabe drugih objektno usmerjenih načel (polimorfizem, skrivanje informacij, kapsulacija, pozno povezovanje itd.). Jezik Express kot tak sicer ni povezan z nobeno obstoječo tehniko za konceptualno modeliranje in zato omogoča formalni opis konceptualnih shem, razvitih s katerokoli tehniko za informacijsko modeliranje. V veliko podporo uporabi jezika Express je tudi skladna grafična notacija Express-G, ki je njegov grafični podniz in je v celoti usklajena s samim jezikom. Razvijajo in pripravljajo tudi novo različico jezika Express (Express-C), ki bo vsebovala tudi preostale objektno usmerjene značilnosti.

Literatura:

- CEN prENV 12009:1996, *Geografske informacije – Referenčni model*
CEN prENV 12160:1996, *Geografske informacije – Opis podatkov – Prostorska shema*
CEN prENV 12656:1996, *Geografske informacije – Opis podatkov – Kakovost*

Prispelo za objavo: 1997-11-13

Kako priti do registra stavb in stanovanj? Po najgospodarnejši poti!

1 UVOD

Strokovno posvetovanje o nepremičninah ne sme mimo brez resnega razpravljanja o problematiki vzpostavitve registra nepremičnin, evidentiranju nepremičnin, prav tako pa tudi ne brez določenih konkretnih sklepov in njihove uresničitve. Vsako razmišljanje o tej problematiki pred strokovnjaki s tega področja je zelo zahtevno, saj je veliko navzočih neposredno ali posredno z različnimi nameni posegalo v to področje. Poskusil bom sistemizirati določena dejstva, ugotoviti, kje so stališča enotna (skladna), in na podlagi tega podati svoje mnenje in predloge.

2 KATERA JE NAJGOSPODARNEJŠA POT OZ. KATERI POSTOPEK JE NAJGOSPODARNEJŠI ZA VZPOSTAVITEV REGISTRA?

Razpravo bom začel z enotnimi stališči (to se je dalo zaključiti iz dosedanjih srečanj in aktivnosti), torej z dejstvi, sprejemljivimi za vse. Nato bom naštel dejstva, o katerih so mnenja različna. Dejstva, s katerimi se v celoti (po mojem mnenju) vsi strinjamo, so naslednja:

- ☐ Popolne evidence ali registra zgradb (stavb, stanovanj, poslovnih prostorov) v Republiki Sloveniji nimamo.
- ☐ Evidenco ali register stavb, prav tako stanovanj in poslovnih prostorov v Republiki Sloveniji nujno potrebujemo. To ugotavljajo mnoga ministrstva in njihova telesa, republiška uprava in mnogi uradi, raziskovalne organizacije, občine in upravne enote (kako že samo dobavitelji energije za ogrevanje, elektrike, iščejo lastnike stanovanj).
- ☐ Evidenco je treba vzpostaviti v najkrajšem času. Strinjamo se, da je treba začeti z majhnimi koraki, z določitvijo minimalne vsebine bodoče evidence, tako da bo ta minimalna vsebina zaživel kot sistem (presekanje stanja, tekoče ažuriranje, izhodne informacije – uporaba).
- ☐ Minimum podatkov je treba združiti iz obstoječih virov in jih medsebojno povezati. Obvezen je tudi obisk terena. Postopek povezovanja podatkov, ki