

@ 11001101010011
1010100101101110
1011INFORMATIKA01
101V LJAVNI0101001010
011011010UPRAVI02016

Informatika v javni upravi:
"Digitalna preobrazba javne uprave - GaaS"

Kongresni center Brdo pri Kranju | 5. - 6. december 2016



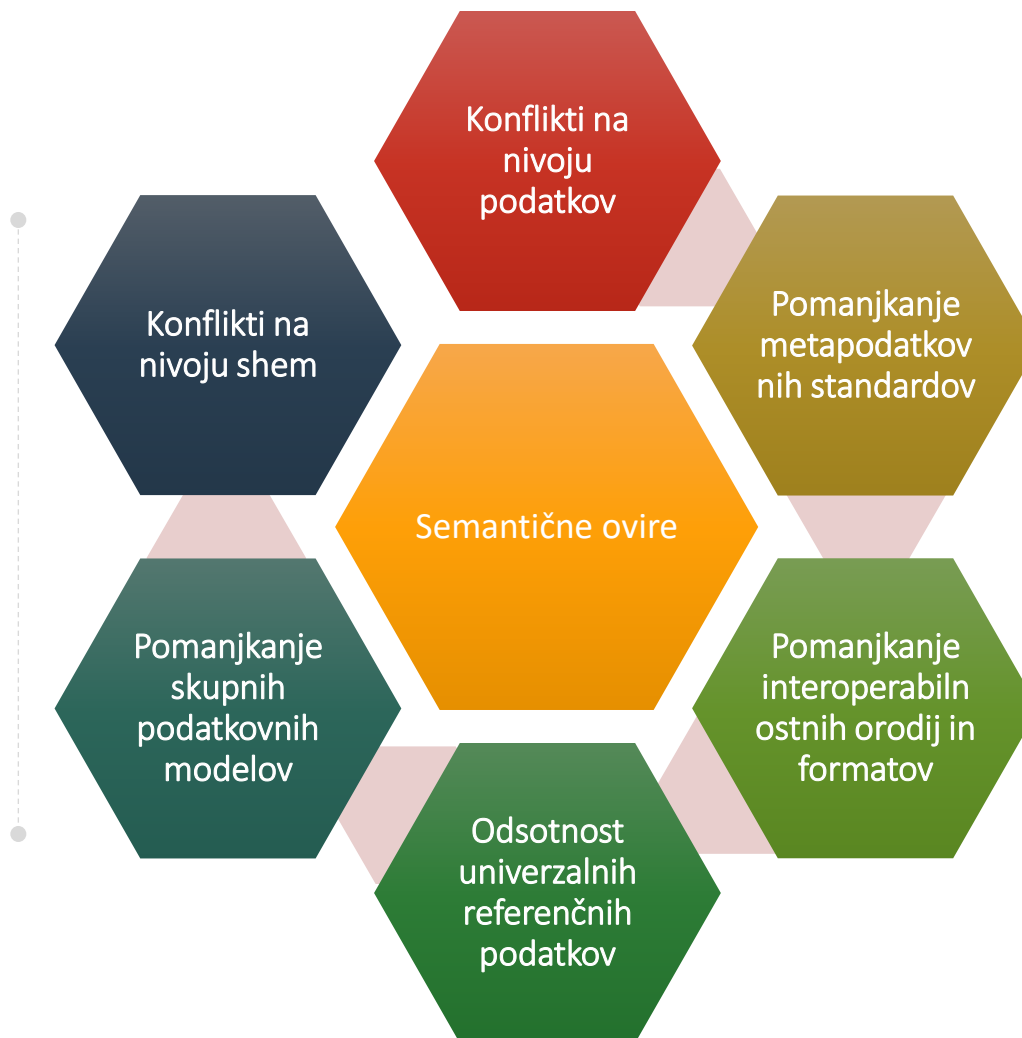
REPUBLIKA SLOVENIJA
MINISTRSTVO ZA JAVNO UPRAVO
DIREKTORAT ZA INFORMATIKO

ZAGOTAVLJANJE SEMANTIČNE INTEROPERABILNOSTI

Igor Tričkovič - Rifelj
Direktorat za informatiko
Ministrstvo za javno upravo

Opredelitev problema

Okolje, v katerem se izmenjujejo podatki med organi javne uprave, je lahko kompleksno in vključuje številne semantične ovire.



Konflikti na nivoju podatkov

Konflikti na nivoju podatkov so:

- **Različne definicije konceptov,**
- **različne vrednosti,**
- **različni formati,**
- **različne enote,**
- **različna natančnost merskih lestvic.**

Konflikti na nivoju shem

Konflikti na nivoju shem so:

- Isto ime za več konceptov ali več imen za isti koncept,
- izomorfizen shem (isti koncept je opisan z različnimi atributi ali z različnim številom atributov),
- različna generalizacija,
- različna agregacija.

Namen in cilji

Namen in cilji zagotavljanja semantične interoperabilnosti so:

- znižanje visoke stopnje heterogenosti in kompleksnosti informacijskih sistemov,
- doseganje višje stopnje standardizacije in interoperabilnosti na podatkovnem sloju,
- omogočanje večjega izkoristka informacijskih virov,
- poenostavitev modeliranja podatkovnega sloja informacijskih sistemov,
- skrajšanje časa, potrebnega za načrtovanje podatkovnega sloja informacijskih sistemov,
- dosledno uveljavljanje načela enkratnega zapisa podatkov,
- omogočanje uporabe standardiziranih gradnikov v okviru generičnih tehnoloških zahtev,
- omogočanje naprednega iskanja in pridobivanja podatkov s portalov organov javne uprave.

Programi evropske unije za zagotavljanje semantične interoperabilnosti

Program in dejavnosti za zagotavljanje semantične interoperabilnosti, ki potekajo na Ministrstvu za javno upravo Republike Slovenije, se navezujejo tudi na sorodne programe Evropske unije. Ti programi so:

- **ISA² (krovni program)**
- **SEMIC,**
- **DCAT-AP,**
- **CORE vocabulary,**
- **INSPIRE,**
- **ltd.**

Program ISA²/SEMIC

Program ISA2/SEMIC je namenjen promoviranju operativne uporabe obstoječih interoperabilnostnih rešitev in razvoju novih struktur na semantičnem sloju, ki so namenjene zagotavljanju načela enkratnega zapisa podatkov.

V okviru programa ISA2/SEMIC poteka tudi nadaljnji razvoj odprtokodne aplikacije za upravljanje tezavrov in ontologij VocBench.

Aplikacija VocBench je večjezično spletno orodje za upravljanje nadzorovanih besednjakov, tezavrov in ontologij. Zagotavlja funkcionalnosti za:

- kreiranje in vzdrževanje nadzorovanih besednjakov, tezavrov in ontologij,
- upravljanje delovnih tokov pri vzdrževanju nadzorovanih besednjakov, tezavrov in ontologij,
- objavo nadzorovanih besednjakov, tezavrov in ontologij kot povezanih odprtih podatkov,
- upravljanje semantičnih relacij z drugimi nadzorovanimi besednjaki,
- administracijo in upravljanje skupin.

RDF, SKOS in OWL

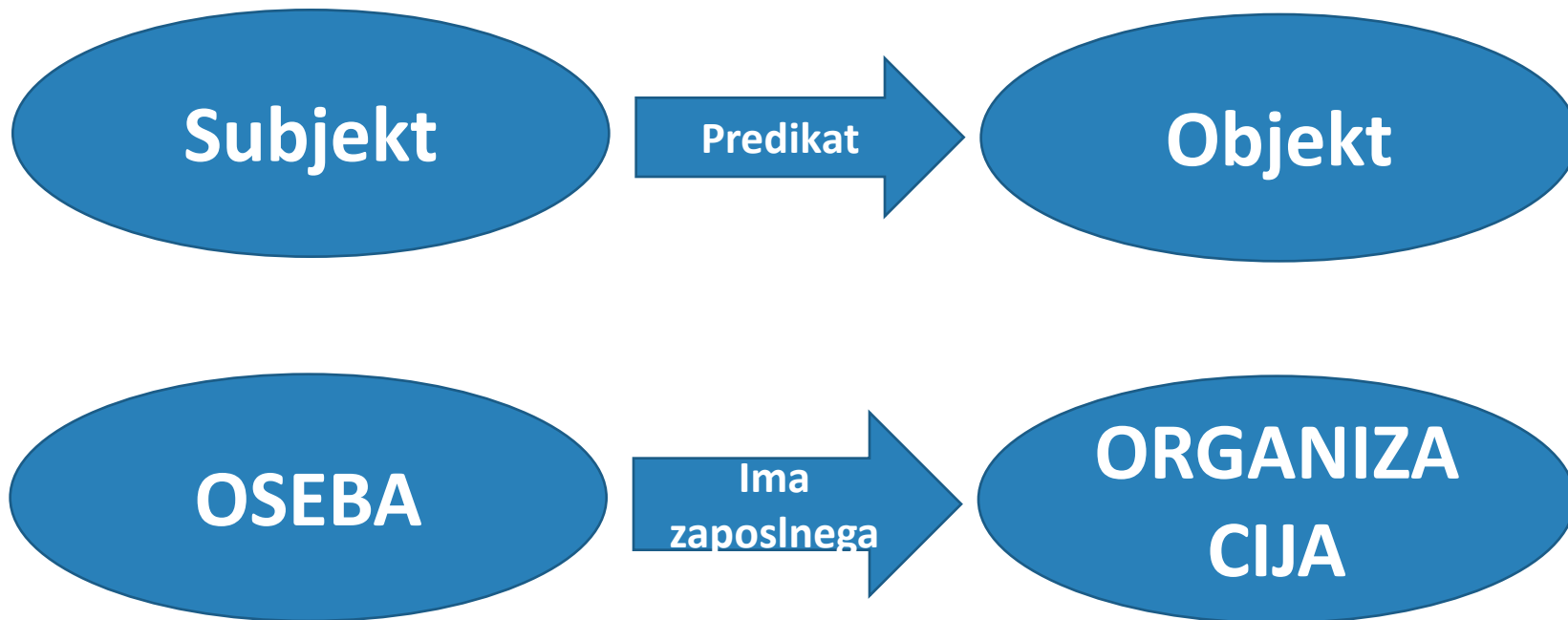
Aplikacija VocBench podpira standarde RDF, SKOS in OWL.

- **RDF (angl. Resource Description Framework) je družina specifikacij konzorcija W3C ([World Wide Web Consortium](http://www.w3.org/)).**
- **SKOS (angl. Simple Knowledge Organization System) je model za izražanje osnovne strukture in vsebine konceptualnih shem, kot so nadzorovani besednjaki, taksonomije in tezavri.**
- **OWL (angl. Web Ontology Language) je družina jezikov za predstavljanje znanja v tezavrih in ontologijah.**

Temeljni koncept jezika RDF



Tako imenovani trojčki (angl. Triplets), ki so sestavljeni iz subjekta, predikata, in objekta.



Opredelitev pojmov nadzorovan besednjak, taksonomija, tezaver in ontologija

Poznamo različne stopnje semantične definiranosti podatkov in sicer:

1. **nadzorovan besednjak (angl. controlled vocabulary),**
2. **taksonomija,**
3. **tezaver,**
4. **ontologija.**

Nadzorovan besednjak

- Je osnovna stopnja semantične definiranosti podatkov.
- Je eksplicitno določen seznam besed, ki ga upravlja določen organ za registracijo.
- Vse besede v nadzorovanem besednjaku morajo imeti jasno, nedvoumno in neredundantno definicijo.
- Najmanj, kar mora nadzorovan besednjak zagotavljati, je:
 - če se nek termin uporablja za različne koncepte v različnih kontekstih, mora biti njegovo ime v določenem kontekstu eksplicitno kvalificirano,
 - če se uporablja več terminov, ki opredeljujejo isti koncept, mora biti eden izmed teh določen kot prednostni, preostali pa kot sinonimi, psevdonimi in podobno.

Taksonomija

- **Taksonomija je zbirka terminov, ki je v osnovi enaka nadzorovanemu besednjaku, vendar so v taksonomiji termini organizirani še v hierarhično strukturo.**
- **Vsak termin v taksonomiji je v enem ali več odnosih nadrejenosti ali podrejenosti do drugih terminov.**
- **V taksonomiji je dovoljena tudi polihierarhija, kar pomeni, da ima lahko termin več nadrejenih terminov.**

Tezaver

- Predstavlja naslednjo, višjo stopnjo semantične definiranosti.
- Tezaver je mrežno strukturiran nadzorovani besednjak, ki poleg hierarhičnih odnosov med koncepti vključuje tudi asociativne (nehierarhične) odnose.
- Izraznost asociativnih odnosov med termini je lahko različnih stopenj.

Ontologija

- **Beseda ontologija ima danes različne interpretacije. Uporablja se jo za poimenovanje podatkovnih slovarjev, tezavrov, taksonomij, shem, podatkovnih modelov, itd.**
- **Gre za nadzorovan besednjak, ki je oblikovan z določenim ontološkim predstavitvenim jezikom. Ta jezik pri definiranju konceptov uporablja določena pravila, s čimer izraža njihove pomene v neki domeni.**

Pričakovani rezultati programa zagotavljanje semantične interoperabilnosti

- Nadzorovani besednjaki, tezavri, ontologije, referenčni podatki in sheme XML, pripravljene za uporabo pri razvoju informacijskih sistemov javne uprave,
- manj interoperabilnostnih konfliktov pri integraciji sistemov in zagotavljanju skupnih storitev organov javne uprave,
- medsebojna usklajenost besednjakov, šifrantov in metapodatkovnih modelov,
- nižji stroški priprave specifikacij zaradi lažjega podatkovnega modeliranja,
- dosledno uveljavljanje načela enkratnega zapisa podatkov in zagotavljanja pravnega okvira zanj.

Prednosti pri iskanju informacij

Semantične strukture, ki so zapisane v formatih RDF, OWL in SKOS omogočajo:

- **Odgovore na vprašanja in ne zgolj iskanje po ključnih besedah.**
- **Pri ekstrakciji informacij ni potrebno ročno brskanje po dokumentih. Odgovori na poizvedbe so predstavljeni v človeku prijazni obliki.**
- **Odgovori na poizvedbe se oblikujejo na osnovi več dokumentov ali spletnih virov.**

Vloga Ministrstva za javno upravo

Centralno vodenje nadzorovanih besednjakov, tezavrov in ontologij za javno upravo.

Vzpostavitev mreže nosilcev za posamezna vsebinska področja.

Usklajevanje vsebin med posameznimi vsebinskimi nosilci.

Strokovna pomoč in uvajanje.

Zagotavljanje zakonske podlage dejavnosti

Informatika v javni upravi 2016 "Digitalna preobrazba javne uprave - GaaS"



Igor Tričkovič - Rifelj
Direktorat za informatiko
Ministrstvo za javno upravo

***Hvala za
pozornost!***