

# Upravljanje s podatki v nacionalni infrastrukturi odprtega dostopa

Milan Ojsteršek, Janez Brezovnik, Marko Ferme  
Laboratorij za heterogene računalniške sisteme  
Fakulteta za elektrotehniko, računalništvo in informatiko,  
Univerza v Mariboru

[milan.ojstersek@um.si](mailto:milan.ojstersek@um.si); [janez.brezovnik@um.si](mailto:janez.brezovnik@um.si); [marko.ferme@um.si](mailto:marko.ferme@um.si)

# Podatki

8	Javne VSŠ	2174	78	2252	1432	1158	2590	4842	563	938	1501	1501	6343
9	Biotehniški center Naklo, Višja strokovna šola	77		77	46	52	98	175	9	21	30	30	205
10	hortikultura	17		17	8	7	15	32	1	7	8	8	40
11	naravovarstvo	20		20	16	20	36	56	2	4	6	6	62
12	upravljanje podeželja in krajine	20		20	22	25	47	67	4	10	14	14	81
13	živilstvo in prehrana	20		20				20	2		2	2	22
14	Biotehniški izobraževalni center Ljubljana, Višja strokovna šola	156	5	161	93	99	192	353	20	52	72	72	425
15	živilstvo in prehrana	81	2	83	46	41	87	170	10	38	48	48	218

# Podatki v človeku razumljivi obliki

Višje strokovne šole												
	redni študij						Skupaj redni študij	izredni študij			Skupaj izredni študij	Skupaj
	1. letnik			2. letnik				vpis				
	prvi vpis	pon. ali ponov. vpis	Skupaj 1. letnik	prvi vpis	pon. ali ponov. vpis	Skupaj 2. letnik		prvi vpis	drugi vpis	Skupaj vpis		
☐ Javne VSŠ	2174	78	2252	1432	1158	2590	4842	563	938	1501	1501	6343
☐ Biotehniški center Naklo, Višja strokovna šola	77		77	46	52	98	175	9	21	30	30	205
hortikultura	17		17	8	7	15	32	1	7	8	8	40
naravovarstvo	20		20	16	20	36	56	2	4	6	6	62
upravljanje podeželja in krajine	20		20	22	25	47	67	4	10	14	14	81
živilstvo in prehrana	20		20				20	2		2	2	22
☐ Biotehniški izobraževalni center Ljubljana, Višja strokovna šola	156	5	161	93	99	192	353	20	52	72	72	425
živilstvo in prehrana	81	2	83	46	41	87	170	10	38	48	48	218
gostinstvo in turizem	75	3	78	47	58	105	183	10	14	24	24	207



# Viri raziskovalnih podatkov

- Naprave
- Instrumenti
- Senzorji
- Programska oprema
- Ljudje



- Opazovanja
- Eksperimenti
- Simulacije
- Emulacije
- Ankete
- Intervjuji
- Analiza besedil
- Tekstovno rudarjenje
- .....



# Tipi podatkov

- Numerični (tabele, rezultati meritev, števci, podatki iz senzorjev...).
- Tekstovni (knjige, zapiski, ankete...).
- Avdio vizualni podatki (slike, zvočni zapisi, video, animacija).
- Prostorski.
- Specifični glede na znanstveno disciplino.
- Specifični glede na merilni instrument ali napravo.

# Formati datotek

- Multipurpose Internet Mail Extensions (MIME type) tipi podatkov – standard [RFC 6838 - https://www.iana.org/assignments/media-types/media-types.xhtml](https://www.iana.org/assignments/media-types/media-types.xhtml)
- Domensko specifični formati raziskovalnih podatkov - [http://justsolve.archiveteam.org/index.php/Scientific Data formats](http://justsolve.archiveteam.org/index.php/Scientific_Data_formats)
- Formati podatkov, ki jih priporoča Library of Congress za trajno hranjenje - <https://www.loc.gov/preservation/resources/rfs/TOC.html>

Delimo jih na **odprte**, **delno zaprte** in **zaprte** formate podatkov.

# Priporočeni formati datotek

- Slike: JPEG, JPG-2000, PNG, TIFF
- Tekst: plain text (TXT), HTML, XML, RDF, PDF/A
- Avdio: AIFF, WAVE, MP3
- Video: MPEG4, JPEG-2000
- Stiskanje in združevanje datotek: TAR, GZIP, ZIP
- Podatkovne baze: RDF, XML ali CSV namesto binarnih formatov aplikacij



# Občutljivi podatki

- Občutljivi podatki vsebujejo potencialno identifikacijske informacije o udeležencu raziskave. Gre za podatke o ljudeh ali druge vrste občutljivih podatkov. Te podatke je potrebno spremeniti, preden jih delimo z javnostjo. Pomembno je, da se te spremembe izvedejo zaradi zaščite zaupnosti udeležencev, lokacije ogroženih prostoživečih živali ali iz drugih pomembnih razlogov ( npr. nacionalna varnost, varstvo intelektualne lastnine). **Vendar pa lahko te spremembe vplivajo na podatke do te mere, da ponovljivost ali dodatne nadaljnje raziskave drugih raziskovalcev nad temi podatki niso več možne.**
- Morda bi razmislili o ohranitvi več različic podatkov: tiste, ki so primerni za javno objavo, in tiste, ki je primerni za nadaljnje raziskave. Zadnji so na voljo zelo omejenemu številu raziskovalcev, ki imajo za njihovo uporabo ustrezno dovoljenje lastnikov podatkov ali etične komisije iz določenega znanstvenega področja.

# Posredni in neposredni identifikatorji oseb

- **Neposredni identifikatorji:** ti podatki kažejo neposredno na posameznika in so običajno odstranjeni iz podatkovnih naborov, preden jih njihovi lastniki delijo z javnostjo. Ti lahko vključujejo: ime, začetnice, poštni naslov, telefonsko številko, e-poštni naslov, edinstvene identifikacijske številke, kot so EMŠO, davčna številka ali številke vozniškega dovoljenja, identifikatorji vozil, identifikatorji medicinskih pripomočkov, spletni ali IP naslovi, biometrični podatki, fotografije osebe, zvočni posnetki, imena sorodnikov, datumi, specifični za posameznika, na primer datum rojstva, poroke itd.
- **Posredni identifikatorji:** ti se lahko sami po sebi zdijo neškodljivi, vendar lahko kažejo na posameznika v kombinaciji z drugimi podatki. Priporočeno je, da nabore podatkov, ki vsebujejo tri ali več posrednih identifikatorjev, pregleda neodvisni raziskovalec ali odbor za etiko, da oceni tveganje identifikacije. Vse posredne informacije, ki niso potrebne za analizo, je treba odstraniti. Morda bi bilo smiselno nekatere od teh vrst podatkov predložiti v agregirani obliki (na primer razponi letnih dohodkov namesto natančnih števil). Posredni identifikatorji lahko vključujejo: kraj zdravljenja ali ime zdravnika, spol, redko bolezen ali zdravljenje, občutljive podatke, kot je uporaba prepovedanih drog ali druga "tvegana vedenja", kraj rojstva, socialno-ekonomske podatke, kot so delovno mesto, poklic, letni dohodek, izobrazba, itd., splošni geografski kazalniki, kot so poštna številka prebivališča, sestava gospodinjstva in družine, etnična pripadnost, leto ali starost rojstva, dobesedni odgovori ali prepisi

# Datatags – share sensitive data with confidence

Tag Type	Description	Security Features	Access Credentials	
Blue	Public	Clear storage, Clear transmit	Open	Non-confidential information
Green	Controlled public	Clear storage, Clear transmit	Email- or OAuth Verified Registration	Non-confidential information
Yellow	Accountable	Clear storage, Encrypted transmit	Password, Registered, Approval, Click-through DUA	Potentially harmful personal information
Orange	More accountable	Encrypted storage, Encrypted transmit	Password, Registered, Approval, Signed DUA	Sensitive personal information
Red	Fully accountable	Encrypted storage, Encrypted transmit	Two-factor authentication, Approval, Signed DUA	Very sensitive personal information
Crimson	Maximally restricted	Multi-encrypted storage, Encrypted transmit	Two-factor authentication, Approval, Signed DUA	Maximum sensitive personal information

Vir: Sweeney L, Crosas M, Bar-Sinai M. Sharing Sensitive Data with Confidence: The Datatags System. *Technology Science*. 2015101601. October 15, 2015. <https://techscience.org/a/2015101601/>

# Tabelarični podatki

- Ne dajajte v eno preglednico več kot eno tabelo.
- V prvi vrstici preglednice naj bodo definirane spremenljivke, da jih lahko razumejo ljudje.
- Grafe dodajte v nove zavihke. Ne dajajte jih v isti zavihek.
- Pretvorite tabelarične podatke iz delno odprtih formatov (npr. XLS) v odprte formate ([CSV](#) ali [TSV](#)).

# XML

```
<prirediv>  
  <naziv>Glasbeni september</naziv>  
  <termin>18.09.2000 do 29.09.2000</termin>  
  <lokacija>Območje mesta Maribor</lokacija>  
  <organizator>SNG Maribor</organizator>  
  <zvrst>koncert</zvrst>  
</prirediv>
```

# Shema XML

```
<?xml version="1.0"?>
  <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

    <xs:element name=",,prireditev">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="naziv" type="xs:string"/>
          <xs:element name="termin" type="xs:string"/>
          <xs:element name="lokacija" type="xs:string"/>
          <xs:element name="organizator" type="xs:string"/>
          <xs:element name="zvrst" type="xs:string"/>
        </xs:sequence>

        </xs:sequence>
      </xs:complexType>
    </xs:element>

  </xs:schema>
```

# Format RSS (Rich Site Summary) - format XML za izmenjavo novic

```
<?xml version="1.0" encoding="windows-1250" ?>
- <rss version="2.0">
- <channel>
  <title>UM FERI - Novice na oglasni deski</title>
  <link>http://www.feri.uni-mb.si/</link>
  <description>Novice na oglasni deski</description>
  <category>Novice</category>
  <copyright>Copyright UM FERI</copyright>
  <managingEditor>potocnik@uni-mb.si</managingEditor>
  <generator>potocnik@uni-mb.si</generator>
  <lastBuildDate>04 Jan 2008 16:45:35 +0100</lastBuildDate>
- <item>
  <title>RSS Resources</title>
  <link>http://www.webreference.com/authoring/languages/xml/rss/</link>
  <description>Defined in XML, the Rich Site Summary (RSS) format has quietly become a dominant
    format for distributing headlines on the Web. Our list of links gives you the tools, tips and tutorials
    you need to get started using RSS. 0323</description>
</item>
- <item>
  <title>MK 1 UN MATEMATIKA I, P 4/1 odpadejo</title>
  <pubDate>03 Jan 2008 13:19:35 +0100</pubDate>
  <link>http://www.feri.uni-mb.si/odeska/brnj2.asp?id=15011&oce=7</link>
  <description>Danes 4. 1. 2008 odpadejo predavanja iz Matematike</description>
  <author>Mateja Osvald</author>
</item>
</channel>
</rss>
```

Primer RSS formata najdete na: <https://dk.um.si/rss.php?o=3&v=dip&lang=slv>

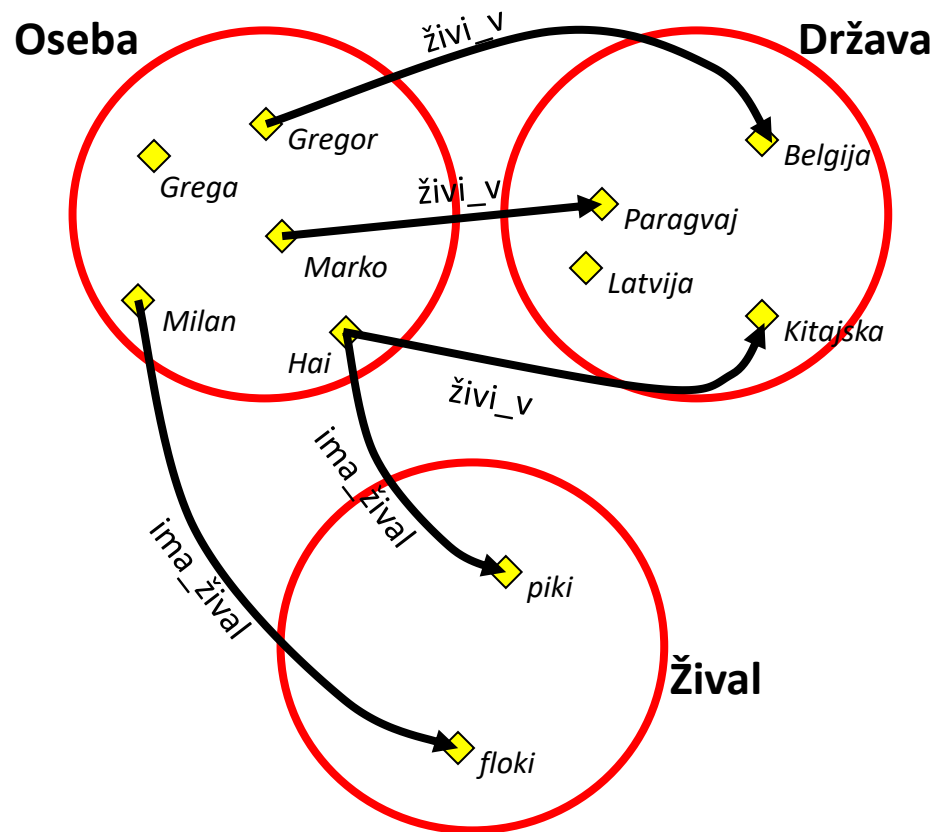
# JSON – Javascript object notation

```
{  
  "glossary": {  
    "title": "example glossary",  
    "GlossDiv": {  
      "title": "S",  
      "GlossList": {  
        "GlossEntry": {  
          "ID": "SGML",  
          "SortAs": "SGML",  
          "GlossTerm": "Standard Generalized Markup Language",  
          "Acronym": "SGML",  
          "Abbrev": "ISO 8879:1986",  
          "GlossDef": {  
            "para": "A meta-markup language, used to create markup languages such as DocBook.",  
            "GlossSeeAlso": ["GML", "XML"]  
          },  
          "GlossSee": "markup"  
        }  
      }  
    }  
  }  
}
```

} - veĉ primerov najdete na <https://json.org/example.html>



# Ontologija

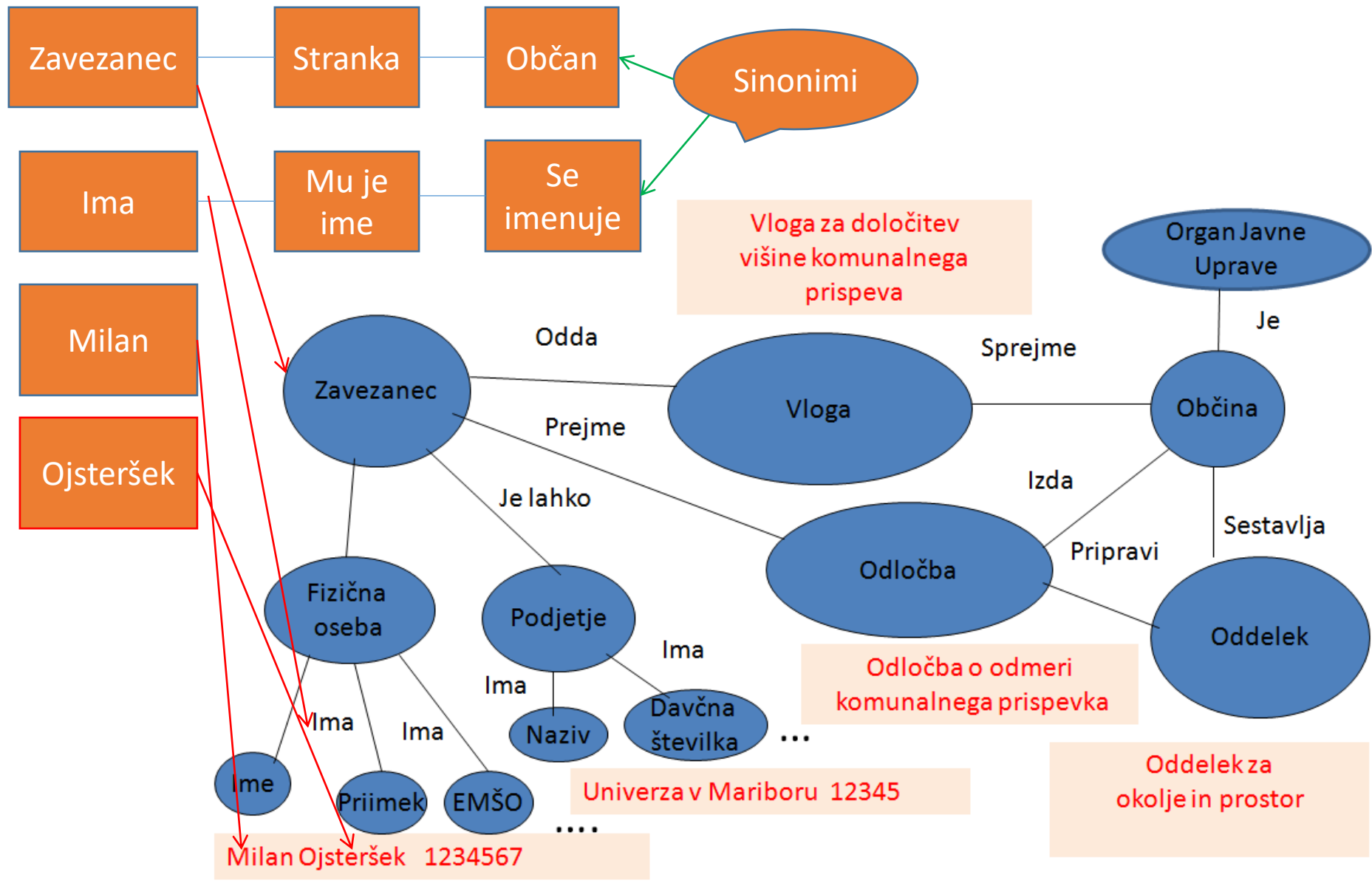


○ Razred , koncept. kategorija  
(angl. class concept,  
category, type)

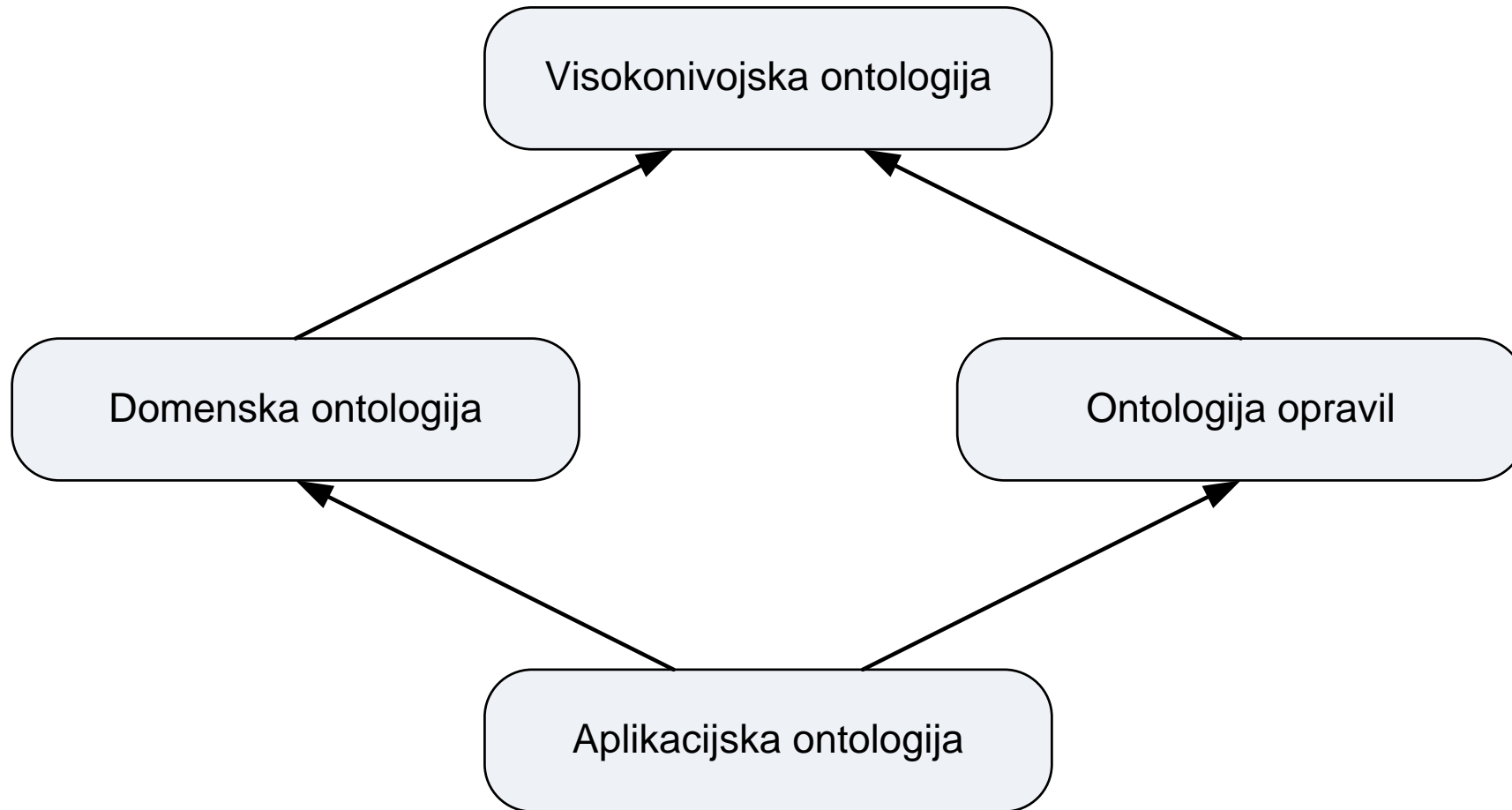
◆ Primerek, posameznik, instanca  
razreda (angl. individual.  
instance)

↑ puščica = relacija, odnos (angl.  
relationship)

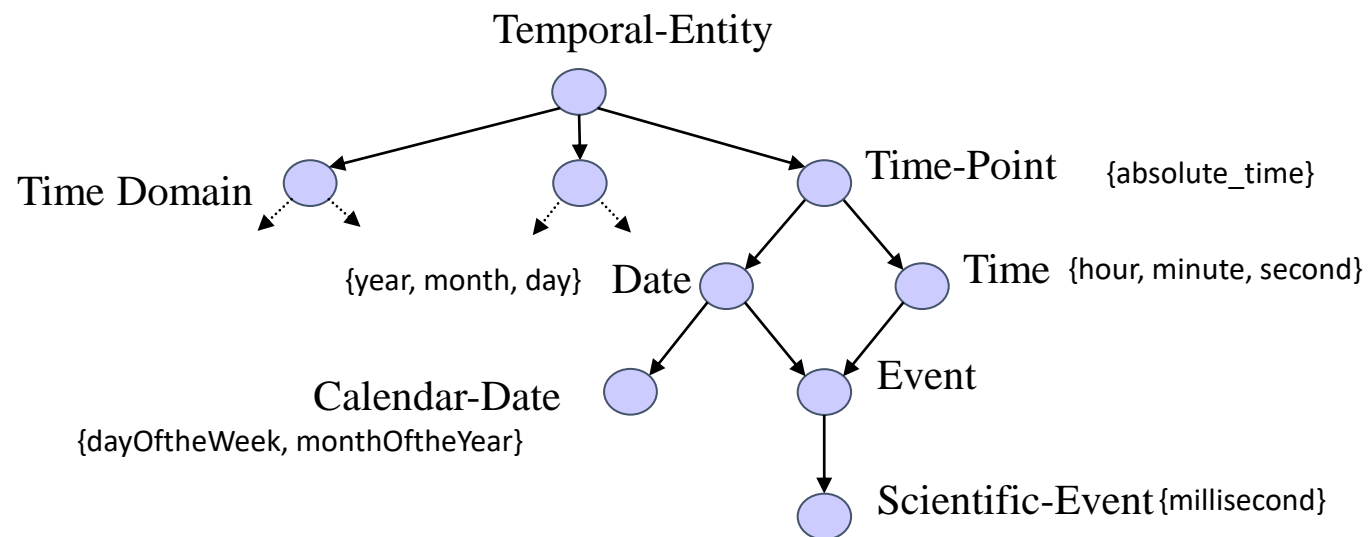
↑ oznaka = lastnost (angl.  
property, slot, relationtype,  
relation, role, semantic link  
type, attribute)



# Vrste ontologij



# Primer visokonivojske ontologije



Primeri so [SUMO](#), [DOLCE](#), [CYC](#),  
[UMBEL](#), [BFO](#), [GFO](#), [COSMO](#),  
[PROTON](#), [MILO](#)



# Skrajšan in standardni opis vira z RDF

Skrajšan opis vira:

```
<rdf:RDF
  <rdf:Description about="http://www.w3.org"
    s:Publisher="World Wide Web Consortium"
    s:Title="W3C Home Page"
    s:Date="1998-10-03T02:27"
  />
</rdf:RDF>
```

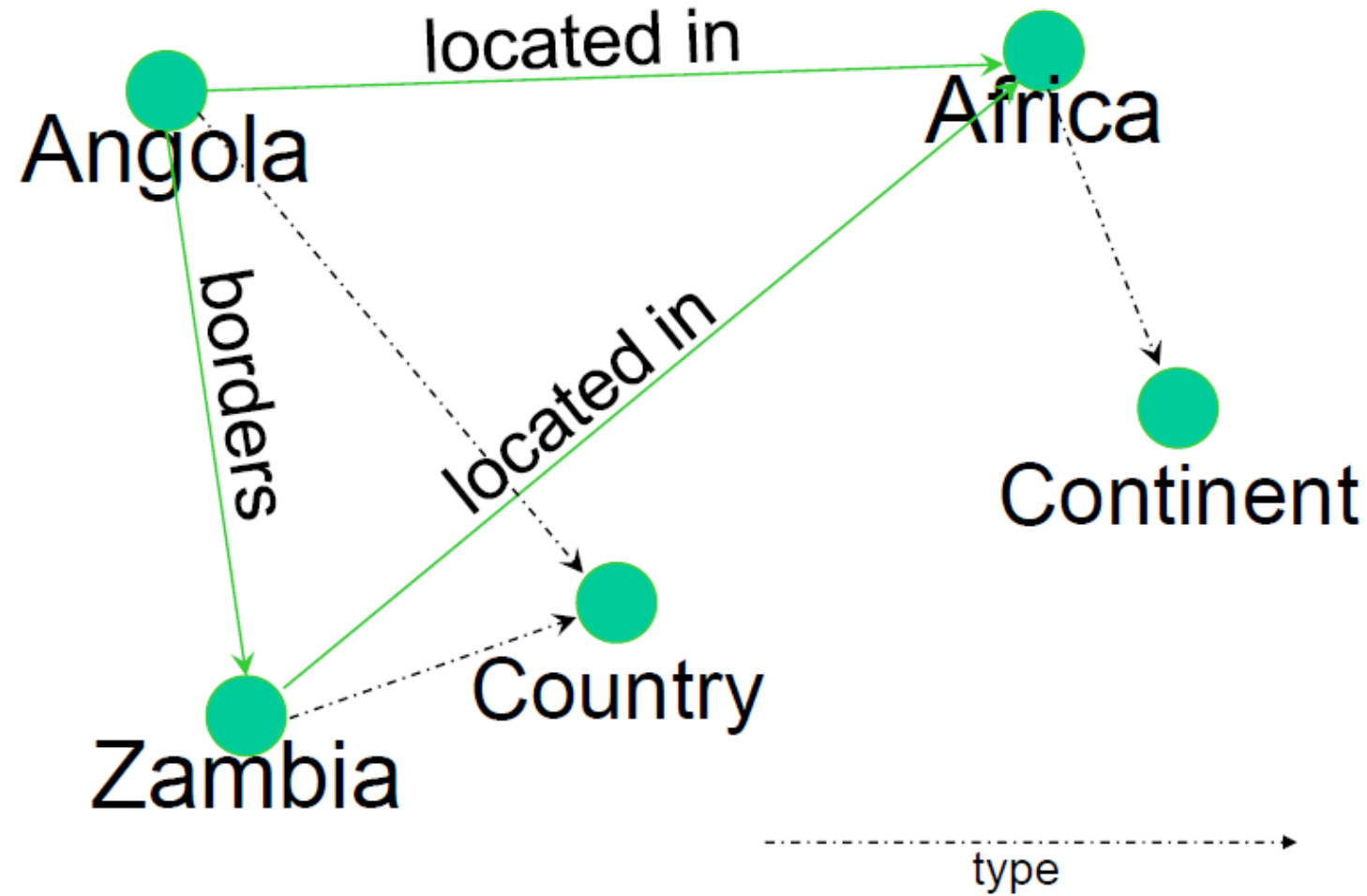
Standardni opis vira:

```
<rdf:RDF
  <rdf:Description about="http://www.w3.org">
    <s:Publisher> World Wide Web Consortium </s:Publisher>
    <s:Title> W3C Home Page </s:Title>
    <s:Date> 1998-10-03T02:27 </s:Date>
  </rdf:Description>
</rdf:RDF>
```

# Primer zapisa primerka v RDF za primerek Angola

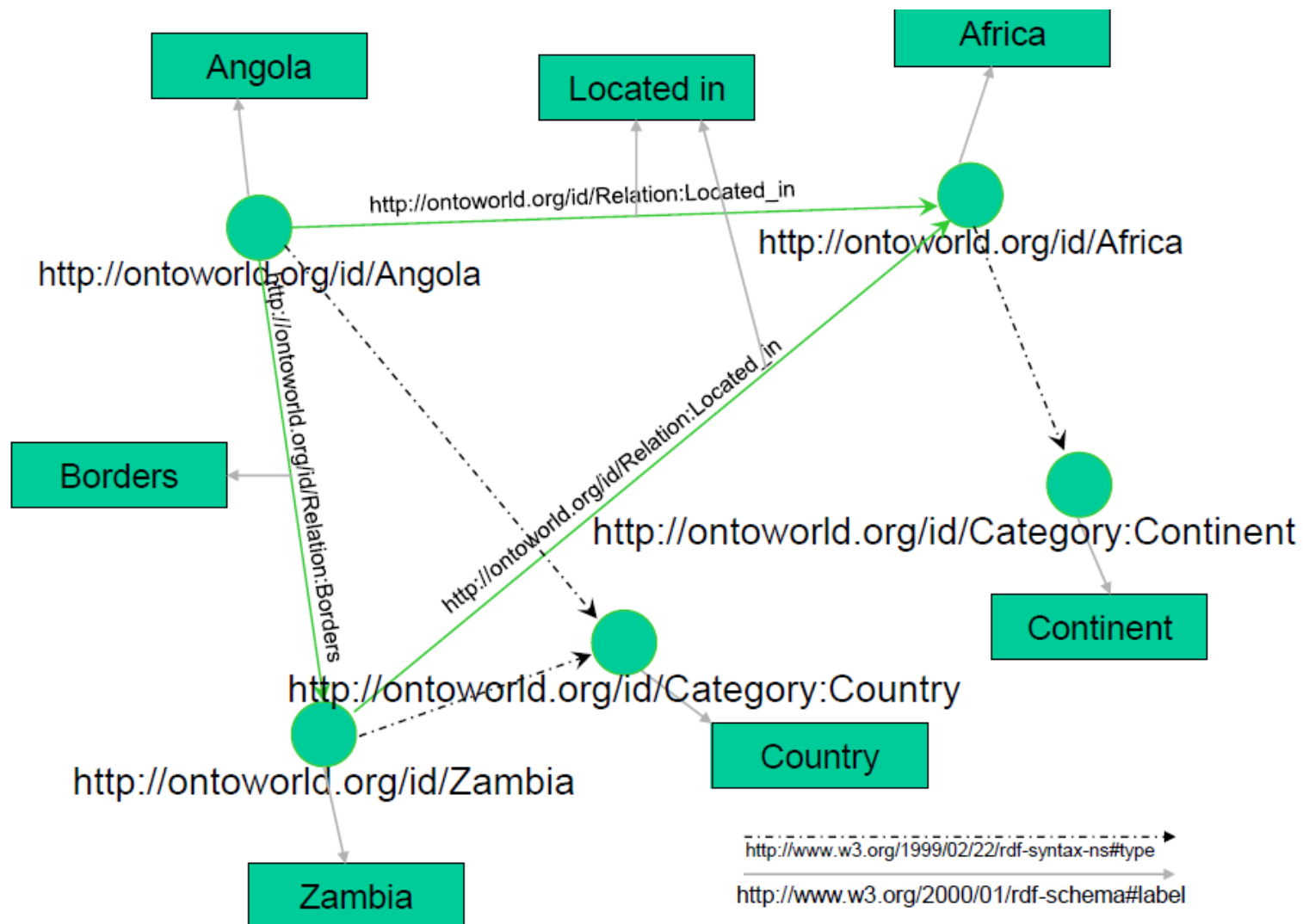
```
<smw:Thing rdf:about="&thing;Angola">
  <rdfs:label>Angola</rdfs:label>
  <smw:hasArticle rdf:resource="&wikiurl;Angola"/>
  <rdfs:isDefinedBy rdf:resource="&wikiurl;Special:ExportRDF/Angola"/>
  <rdf:type rdf:resource="&thing;Category-3ACountry"/>
  <relation:Borders rdf:resource="&thing;Atlantic_Ocean"/>
  <relation:Borders rdf:resource="&thing;Zambia"/>
  <relation:Borders rdf:resource="&thing;Democratic_Republic_of_the_Congo"/>
  <relation:Borders rdf:resource="&thing;Namibia"/>
  <relation:Located_in rdf:resource="&thing;Africa"/>
  <attribute:Area-23km-C2-B2 rdf:datatype="http://www.w3.org/2001/XMLSchema#de
  <attribute:Population rdf:datatype="http://www.w3.org/2001/XMLSchema#integer'
</smw:Thing>
```

# Primer povezav med primerki

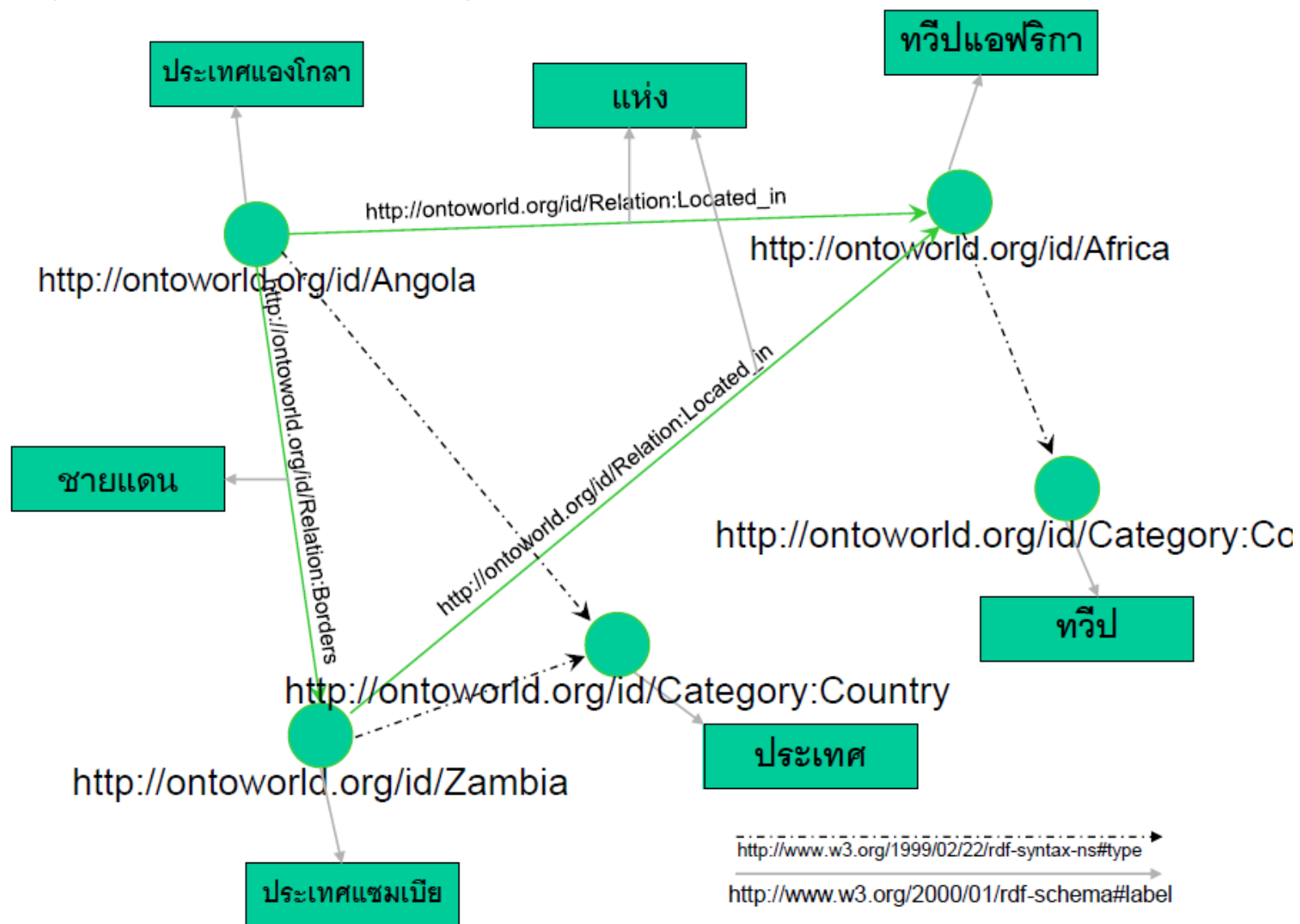


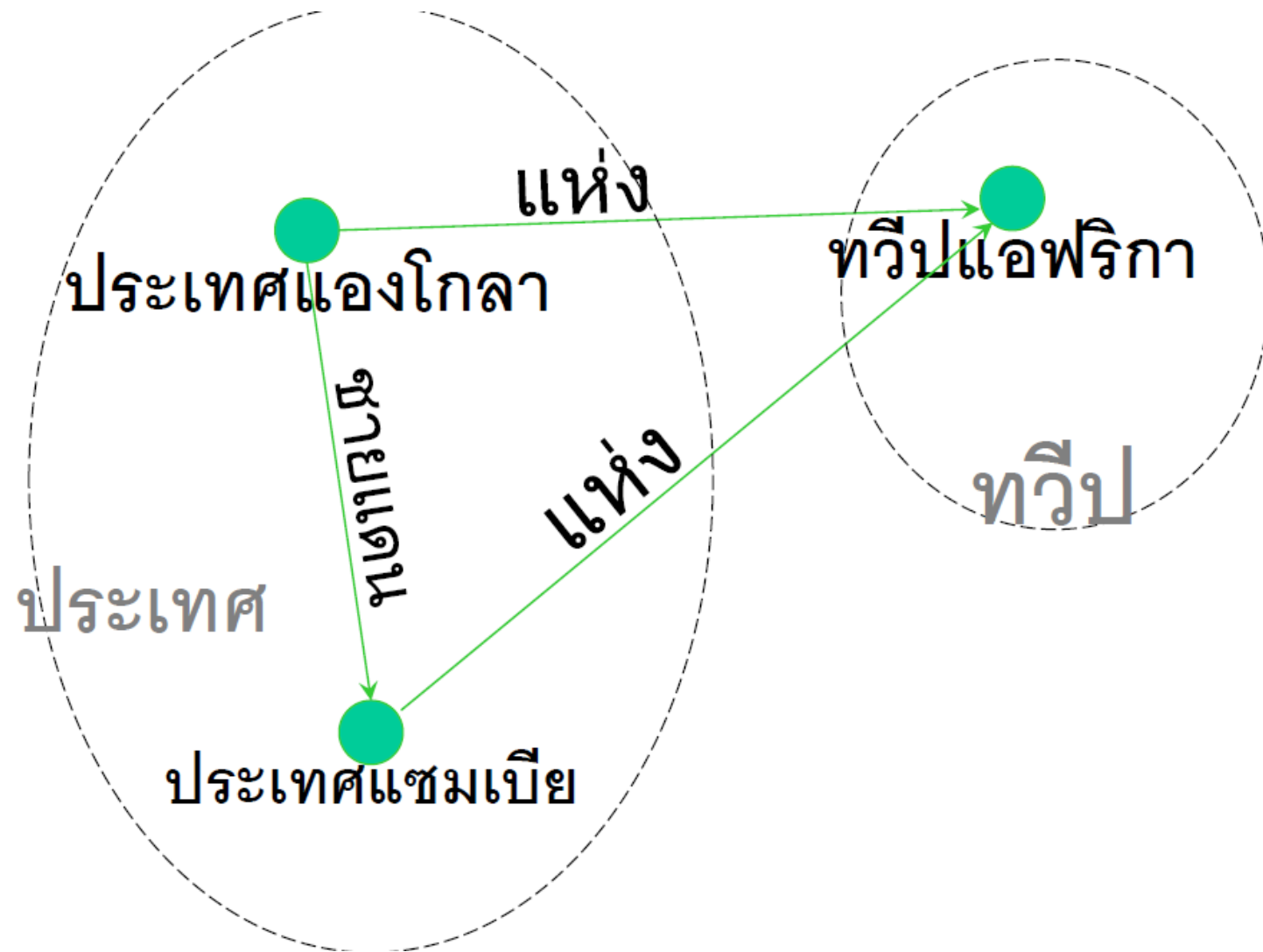


# Primer povezav med rdf trojčki

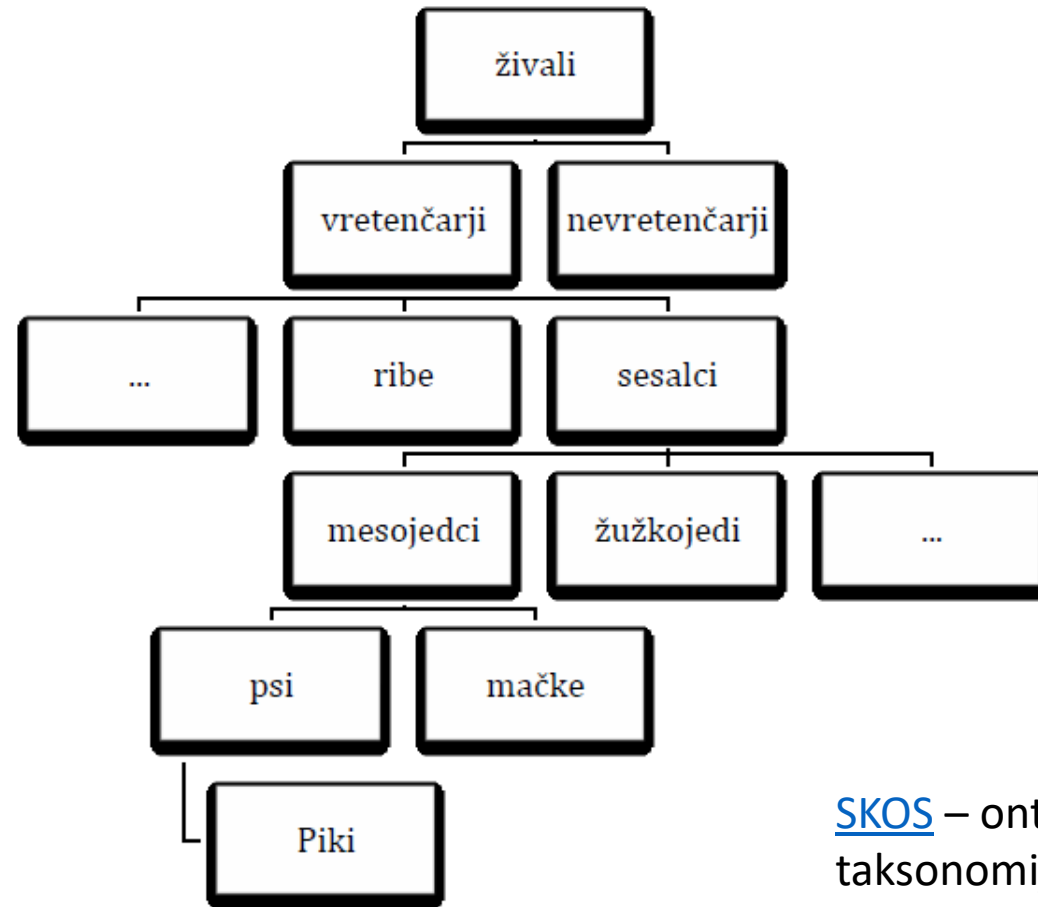


# Primer uporabe večjezičnosti





# Organiziranje informacij - primer taksonomije živali

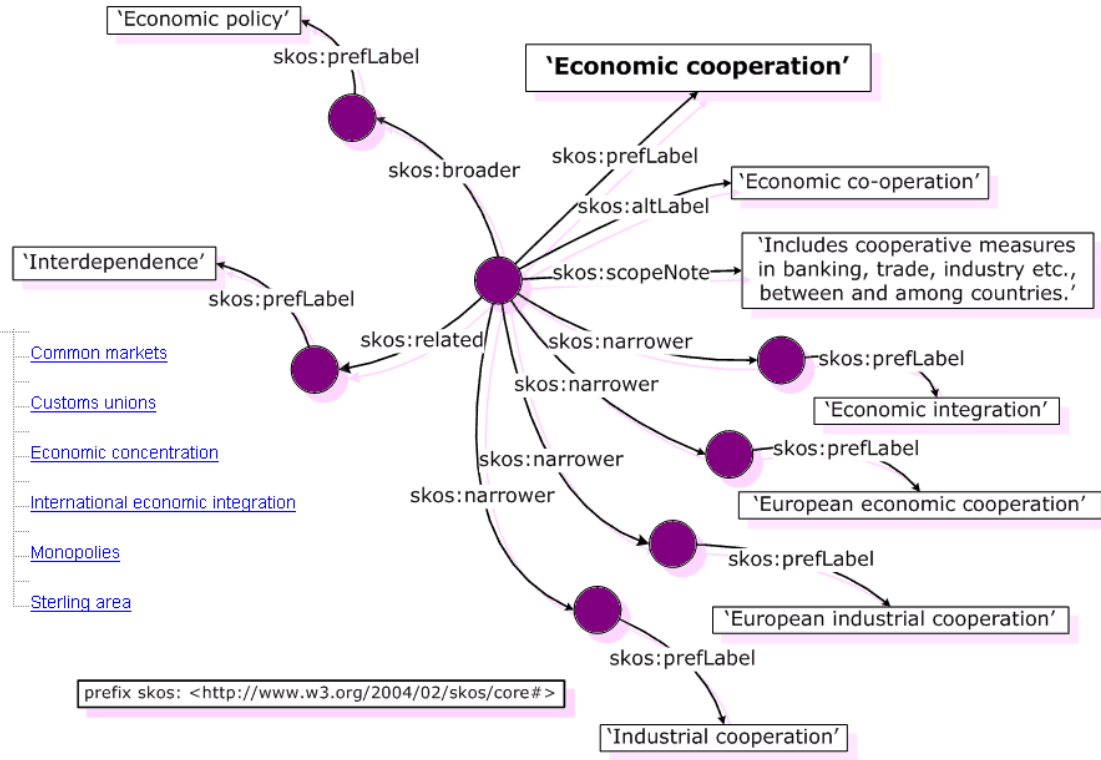


[SKOS](#) – ontologija za opis taksonomij

# UKAT

## 6.25 Economics

- [Business cycles](#)
- [Economic conditions](#)
- [Economic policy](#)
  - [Counter-inflation policy](#)
  - [Economic cooperation](#)
    - [Economic integration](#)
      - [European economic cooperation](#)
      - [European industrial cooperation](#)
      - [Industrial cooperation](#)
  - [Economic legislation](#)
  - [Economic planning](#)
  - [Economic reform](#)
  - [Incomes policy](#)
  - [Nationalization](#)
  - [Price policy](#)
  - [Privatization](#)
  - [Structural adjustment](#)
- [Economic systems](#)
- [Economic theory](#)
- [Economics](#)
- [Income and wealth](#)
- [Interdependence](#)
  - [Economic relations](#)
- [National accounting](#)



# Schema RDF - Primer

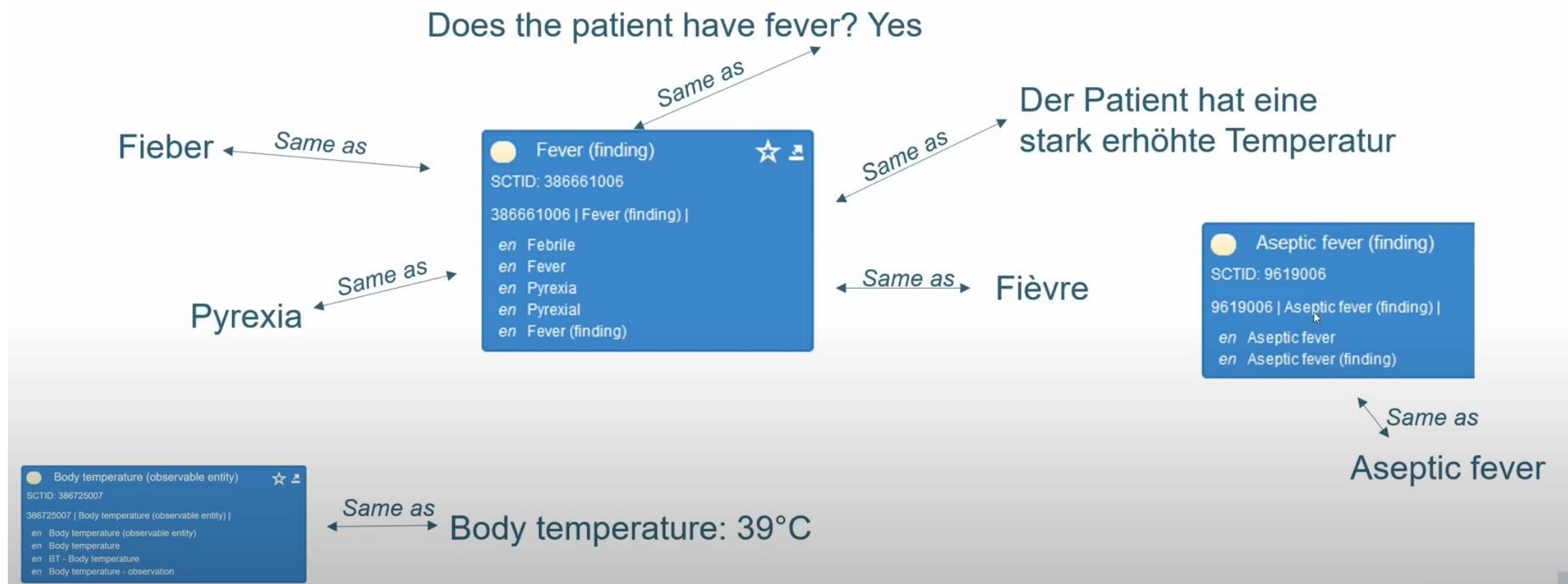
## Daljši zapis:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xml:base="http://www.animals.fake/animals#">
  <rdf:Description rdf:ID="animal">
    <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  </rdf:Description>
  <rdf:Description rdf:ID="horse">
    <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
    <rdfs:subClassOf rdf:resource="#animal"/>
  </rdf:Description>
</rdf:RDF>
```

## Skrajšan zapis:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xml:base="http://www.animals.fake/animals#">
  <rdfs:Class rdf:ID="animal" />
  <rdfs:Class rdf:ID="horse">
    <rdfs:subClassOf rdf:resource="#animal"/>
  </rdfs:Class>
</rdf:RDF>
```

# Primer pomenskega opisa vročine pacienta

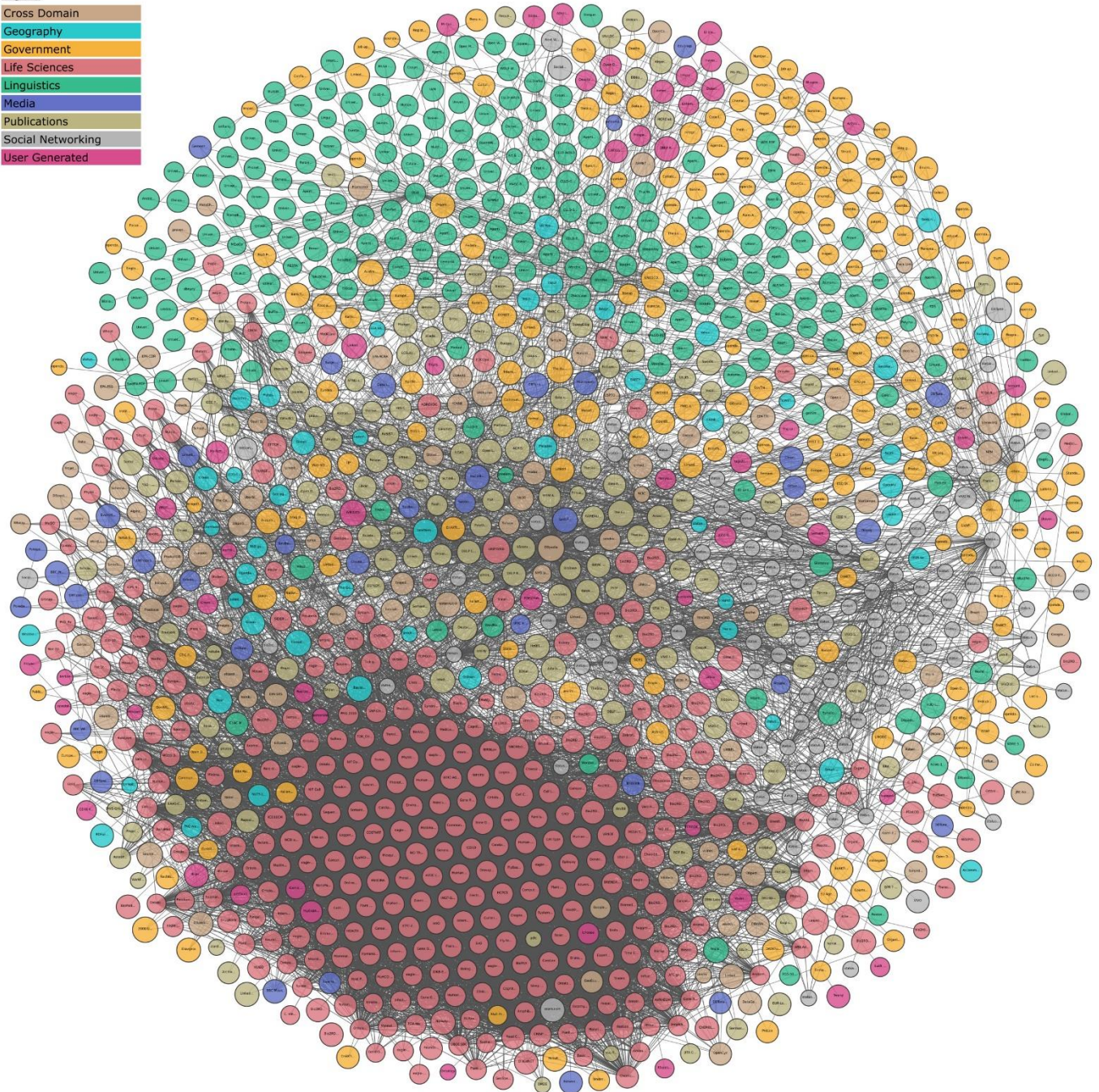


# Povezave na kontrolirane besednjake

- DDI Controlled Vocabularies: <https://ddialliance.org/controlled-vocabularies>
- CESSDA vocabularies: <https://vocabularies.cessda.eu/>
- COAR vocabularies: <https://www.coar-repositories.org/news-updates/what-we-do/controlled-vocabularies/>
- Research vocabularies Australia: <https://vocabs.ardc.edu.au/search#!?q=&pp=15&p=1&activeTab=vocabularies>
- Unified Medical Language System (UMLS): <https://www.nlm.nih.gov/research/umls/index.html>
- Getty's vocabularies: <https://www.getty.edu/research/tools/vocabularies/index.html>
- Data privacy vocabulary: <https://harshp.com/dpv/dpv/>
- Backbone thesaurus: <https://www.backbonethesaurus.eu/>
- NERC: <https://vocab.nerc.ac.uk/>
- Finto: <https://finto.fi/en/>
- EU vocabularies: <https://op.europa.eu/en/web/eu-vocabularies>
- Linked open vocabularies: <https://lov.linkeddata.es/dataset/lov/>
- Open BioMedical Ontologies (OBO) - <https://obofoundry.org>
- IVOA Vocabularies: <https://ivoa.net/rdf/> (astronomy, despite the name, it's not only RDF)
- Scientific Evidence Code System (SEVCO) in development for study design, statistics, and risk of bias terms <https://fevir.net/resources/Project/27845> and <https://fevir.net/resources/CodeSystem/27270>



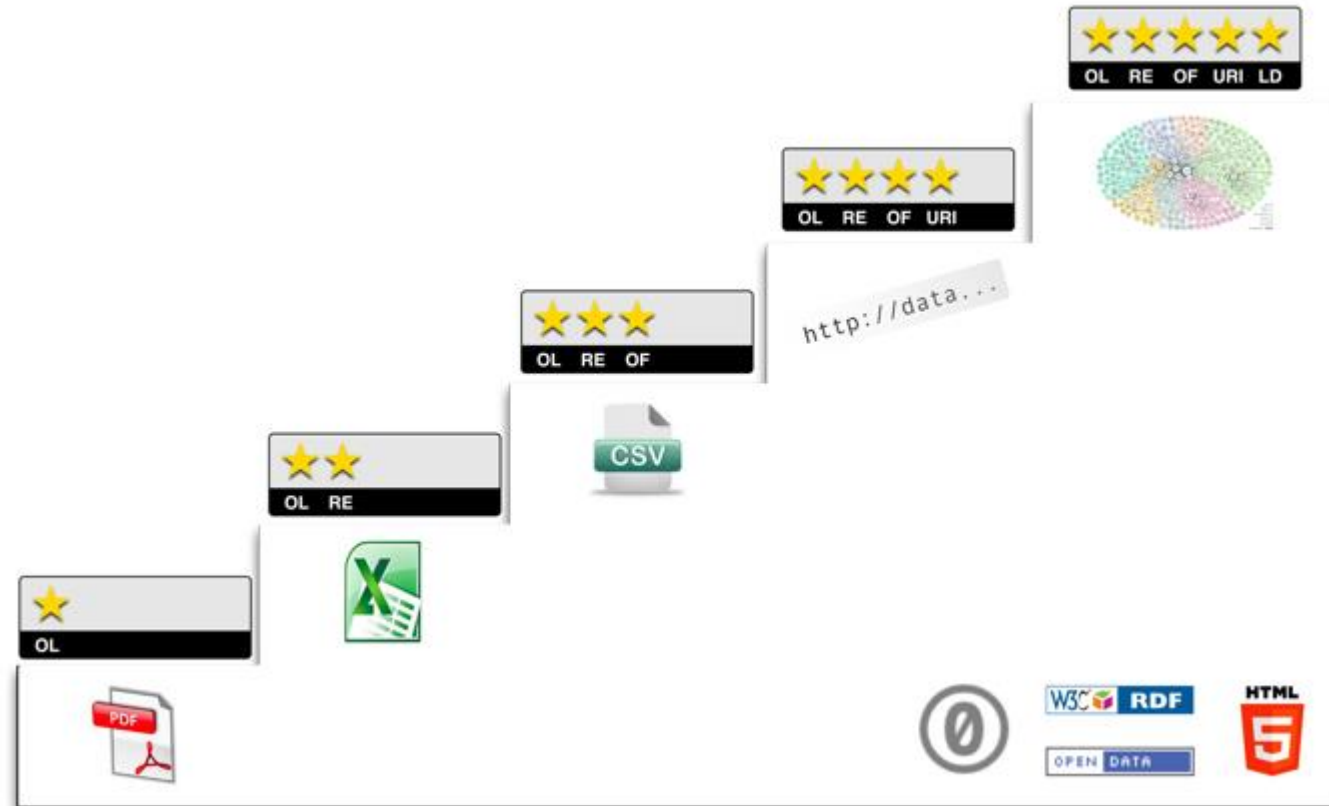
# Linked open data cloud



# Povezave na primere ontologij in standardov za podatke

- EU semantic interoperability community: <https://github.com/SEMICEu>
- OE-GOV: Ontologies for e-government: <http://www.oegov.us/>
- Bioportal: <https://bioportal.bioontology.org/>
- OBO Foundry: <http://www.obofoundry.org/>
- Agroportal: <http://agroportal.lirmm.fr/>
- Spar ontologies: <http://www.sparontologies.net/ontologies>
- OGC standards: <https://www.ogc.org/docs/is>
- BARTOC terminološki registri: <http://bartoc.gbv.de/registries>

# Pet stopenj odprtih podatkov



Več na <http://5stardata.info/>

# Kaj pomenijo te številke?

001	12	01	98
002	36	02	175
003	72	01	94
004	42	01	130
005	18	02	125

# To nam povedo metapodatki!

Respondent ID	Age	Sex	Weight
001	12	01	98
002	36	02	175
003	72	01	94
004	42	01	130
005	18	02	125

**Age is in years as of last birthday**

**Sex: 01 = Male, 02 = Female**

**Weight is in pounds**

Metapodatki

# Kaj je v tej konzervi?



# Sedaj veste!



# Metadapodatki, ki jih dobimo iz fotografije





# Metapodatki

- Vsebinski
  - Strukturni
  - Skrbniški
- 
- Bere jih lahko človek ( angl. human readable)
  - Strojno berljivi ( angl. machine readable)
  - Strojno izvedljivi (angl. machine actionable)

Primeri metapodatkovnih formatov: DC, COMARC, PREMIS, DCAT, FOAF...

# Metapodatkovni standardi - osnovni

## *Dublin Core*

- Title
- Creator
- Subject
- Description
- Publisher
- Contributor
- Date
- Type
- Format
- Identifier
- Source
- Language
- Relation
- Coverage
- Rights

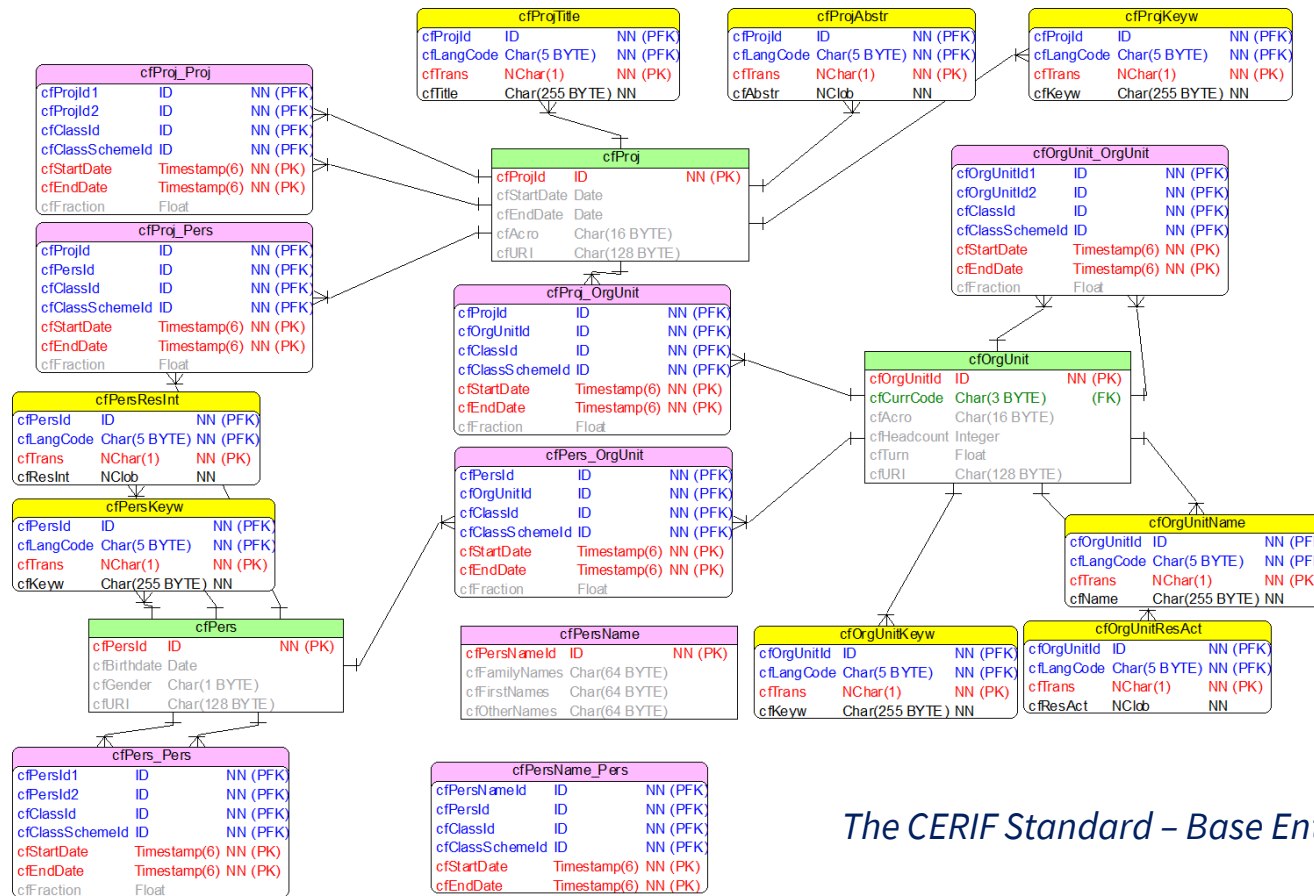
## *DataCite*

- Title
- Creator
- Publisher
- Identifier
- Publication Year
- Resource Type
- Subject
- Contributor
- Date
- Related identifier
- Description
- Geolocation
- Language
- *Alternate identifier*
- *Size*
- *Format*
- *Version*
- *Rights*
- *Funding Reference*

## *EDMI*

- Name
- Description
- Identifier
- url
- Creator
- Date Created
- license
- Data Standard
- Date Modified
- Access URL
- Access Interface
- Structure
- Included In
- Measurement Technique
- Keywords
- Variable Measured
- Format
- Scientific Type
- Includes
- Content Type
- Size
- Authentications
- *Version*
- *Metric*
- *Same as*
- *Spatial Coverage*
- *Temporal coverage*
- *Citation*
- *Reference citation*
- *compression*

# Metapodatkovni standardi - kompleksnejši



The CERIF Standard - Base Entities

# Dublin Core ( DC)

- Je standarden in splošen format za predstavitev metapodatkov za vire na medmrežju. Specifikacija za Dublin Core (DC) razlikuje dve ravni standarda: *enostavno* in *kvalificirano*. Enostavni DC predpisuje 15 elementov (polj), kvalificirani DC pa vključuje dodatni element Audience in skupino podrobnosti (lastnosti oz. kvalifikatorjev) za posamezne elemente; ki omogočajo boljše odkrivanje virov. Značilnost DC je tudi ta, da se vsak element lahko poljubno krat ponovi; pri tem pa ni noben element obvezen. Za potrebe prenosov podatkov se ponavadi uporablja nekvalificiran (enostaven) DC.
- Več o posameznih metapodatkih iz sheme DC najdete na <http://dublincore.org/documents/dcmi-terms/>
- Poglejte tudi klasifikacije od [Library of Congress](#) , [ACM klasifikacijo](#), [SEARS klasifikacijo](#), [Dewey Decimal klasifikacijo](#), [UNESCO nomenklatura](#), [UDC \(UDK - Universal Decimal Clasification\)](#), [Medical Subject Hedings](#), [National Library of Medicine Classification](#), [Getty Thesaurus of Geographic Names](#), [MARC standard](#), [Metapodatkovni standardi](#)

<?xml version="1.0" ?>

```
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
- <rdf:Description rdf:about="http://dkum.uni-mb.si/lzpisigradiva.php?id=5946 ">
  <dc:title>Placilni procesi multinacionalnega podjetja : diplomsko delo</dc:title>
  <dc:creator>Filipan Kraljic, Biserka (Komentor )</dc:creator>
  <dc:creator>Zbašnik, Dušan (Mentor)</dc:creator>
  <dc:creator>Kutnjak, Matej (Avtor)</dc:creator>
  <dc:subject>plačilni promet</dc:subject>
  <dc:subject>prenosni sistemi</dc:subject>
  <dc:subject>multinacionalne družbe</dc:subject>
  <dc:subject>plačilni instrumenti</dc:subject>
  <dc:subject>kliring</dc:subject>
  <dc:subject>dokumentarni akreditivi</dc:subject>
  <dc:subject>bančno poslovanje</dc:subject>
  <dc:subject>garancije</dc:subject>
  <dc:subject>jamstvo</dc:subject>
  <dc:subject>mednarodne finance</dc:subject>
  <dc:subject>menice</dc:subject>
  <dc:subject>plačilni sistemi</dc:subject>
  <dc:subject>poravnava</dc:subject>
```

```
<dc:subject>mednarodna podjetja</dc:subject>
<dc:subject>cilj</dc:subject>
<dc:subject>centralizacija</dc:subject>
<dc:subject>zakladnica</dc:subject>
<dc:description>[M. Kutnjak]</dc:description>
<dc:date>2006</dc:date>
<dc:date>2007-09-11 14:51:18</dc:date>
<dc:type>Bibliografija</dc:type>
<dc:identifier>5946</dc:identifier>
<dc:identifier>COBISS_ID: 402472</dc:identifier>
<dc:identifier>UDK: 339.727(043.2):334.726(497.4:4)</dc:identifier>
<dc:source>Velenje</dc:source>
<dc:language>sl</dc:language>
</rdf:Description>
</rdf:RDF>
```

Primer DC za  
diplomo iz  
Digitalne  
knjižnice  
Univerze v

# JSON-LD

- JSON-LD je notacija, ki se uporablja za zapis metapodatkov po priporočilu Schema.org (<https://schema.org/> ).
- Primere JSON-LD najdete na <https://jsonld.com>
- [Primer JSON-LD za nabor podatkov.](#)
- Primer API-ja na Openscience.si (<http://openscience.si/OpenData.aspx> )

# Pomembni viri za metapodatke

- RDA [Guidelines for publishing structured metadata on the Web](#)
- DCC list of metadata standards: <https://www.dcc.ac.uk/guidance/standards/metadata/list>
- RDA metadata standards catalog: <http://rd-alliance.github.io/metadata-directory/standards/>
- Fairsharing: <https://fairsharing.org/>
- DCMI Metadata Terms: <https://dublincore.org/specifications/dublin-core/dcmi-terms/>
- DataCite Metadata Schema 4.3: <https://schema.datacite.org/meta/kernel-4.3/>
- DCAT 3.0: <https://www.w3.org/TR/2021/WD-vocab-dcat-3-20210504/>
- DCAT Application profile for data portals in Europe 2.0.1: <https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe>
- GeoDCAT-AP 1.0.1 <https://joinup.ec.europa.eu/solution/geodcat-application-profile-data-portals-europe/distribution/geodcat-ap-101-docx>
- StatDCAT- AP 1.0.1: <https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/statdcat-application-profile-data-portals-europe/release/101>
- DDI Codebook 2.5: [https://ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/field\\_level\\_documentation.html](https://ddialliance.org/Specification/DDI-Codebook/2.5/XMLSchema/field_level_documentation.html)
- Bioschemas: <https://bioschemas.org/>

# Življenski krog ravnanja z raziskovalnimi podatki

- Načrtovanje in iskanje virov podatkov.
- Zbiranje in ustvarjanje.
- Obdelava in analiza.
- Objava in deljenje.
- Trajno ohranjanje.
- Ponovna uporaba podatkov in drugih rezultatov raziskav.
- Več na <https://www.adp.fdv.uni-lj.si/usposobi/ZKG/>



# Načrtovanje in iskanje virov podatkov

- Kaj bomo raziskovali?
- Katere raziskovalne podatke boste zbirali ali ustvarjali?
- Kako bodo raziskovalni podatki zbrani ali ustvarjeni?
  - Metode zbiranja, kakovost podatkov, IPR ponovno uporabljenih podatkov.
- Priprava ustreznih dokumentov:
  - Načrt ravnanja s podatki.
  - Priprava obrazca za obveščeno soglasje
  - Vloga za etično presojo raziskave.

# Vsebina načrta za ravnanje z raziskovalnimi podatki

- Osnovni podatki o projektu.
- Katere raziskovalne podatke boste zbirali ali ustvarjali?
- Kako bodo raziskovalni podatki zbrani ali ustvarjeni?
- Katera dokumentacija bo ustrezno opisala vaše raziskovalne podatke?
- Izbira metapodatkovnega in podatkovnega standarda.
- Kako boste reševali etična vprašanja?
- Kako boste urejali vprašanja glede avtorskih pravic in pravic intelektualne lastnine (IP / IPR)?
- Kako bodo raziskovalni podatki shranjeni in varnostno kopirani med raziskavo?
- Kako boste upravljali dostop in zagotovili varnost raziskovalnih podatkov?

# Vsebina načrta za ravnanje z raziskovalnimi podatki - nadaljevanje

- Kateri raziskovalni podatki so dolgoročno vredni in jih je potrebno hraniti, deliti in/ali ohraniti?
- Izbira ustreznega repozitorija ali podatkovnega arhiva.
- Kako boste raziskovalne podatke delili?
- Ali so potrebne kakršne koli omejitve za izmenjavo raziskovalnih podatkov?
- Kdo bo odgovoren za ravnanje z raziskovalnimi podatki?
- Katere vire boste potrebovali za uresničitev svojega načrta?
- Kako bodo potencialni uporabniki našli vaše raziskovalne podatke?
- Kako narediti raziskovalne podatke odprto dostopne?
- Kako omogočiti interoperabilnost vaših raziskovalnih podatkov?
- Kako zagotoviti ponovno uporabo raziskovalnih podatkov.

# Obveščeno soglasje

- Opis raziskave, metode zbiranja podatkov, opis zbranih podatkov.
- Obveznosti udeleženca raziskave.
- Obseg obveznosti in nadomestilo za udeležbo.
- Tveganja in nevarnosti udeležbe.
- Postopek odstopa od sodelovanja v raziskavi.
- Koristi udeležbe.
- Prostovoljnost udeležbe.
- Varovanje osebnih podatkov, publiciranje rezultatov, arhiviranje in deljenje podatkov.
- Kontaktni podatki raziskovalca, virov financiranja
- Dodatne privolitve in pojasnila glede GDPR, če zajemamo osebne podatke udeležencev raziskave (varovanje osebnih podatkov, čas hranjenja osebnih podatkov, anonimizacija, zagotavljanje zaupnosti)

# Obveščeno soglasje - nadaljevanje

- Obrazec o pristanku mora biti napisan v enostavnem jeziku. **Udeleženec raziskave mora imeti jasno možnost opredeliti se glede naslednjega:**
  - da je prebral in razumel informacije o projektu,
  - da je imel možnost pridobiti dodatne informacije,
  - da se prostovoljno strinja s sodelovanjem v projektu,
  - da razume, da lahko preneha s sodelovanjem brez opredelitve razloga za to in brez kazni,
  - da so pojasnjeni postopki glede zaupnosti (uporaba imena, psevdonimi, anonimiziranje podatkov itd.) in
  - da so razloženi načini uporabe podatkov iz raziskav v publikacijah, za izmenjavo in arhiviranje.
  - Obrazec naj vsebuje podpise in datume podpisa za udeleženca in raziskovalce.
- Udeleženca raziskave je potrebno informirati, da se bodo **po zaključku projekta raziskovalni podatki shranili in bodo dostopni drugim raziskovalcem za potrebe sekundarne analize v raziskovalne in izobraževalne namene.**

# Vloga za etično presojo

- Kontaktni podatki vlagatelja.
- Opis raziskave ( namen, trajanje, opis zbranih podatkov, metode zbiranja, etični vidiki, koristi, tveganja in obremenitve vključenih udeležencev, informacija o denarnem nadomestilu udeležencem raziskave, varnost udeležencev raziskave, zavarovanje v primeru morebitne škode za zdravje).
- Naročnik in stroški raziskave.
- Izjave vodje raziskave, nadrejenih in nadzornikov raziskave.
- Obrazec za obveščeno soglasje udeleženca.
- Če gre za otroke in druge ranljive skupine, obrazec s privolitvijo staršev oz. skrbnikov.
- Informacije glede zbiranja, varnosti in hranjenja osebnih podatkov.
- Podpis vlagatelja.

# Obdelava in analiza podatkov

- Vnos podatkov, digitalizacija, prepis, prevod.
- Preverjanje, potrjevanje, čiščenje, anonimizacija podatkov, kjer je potrebno.
- Opis, strukturiranje in dokumentiranje podatkov.
- Analiza podatkov.
- Interpretacija rezultatov.
- Upravljanje s podatki in hramba ( varnostne kopije: [3-2-1 pravilo arhiviranja](#)).

# Objava in deljenje podatkov

- Razrešitev avtorskih pravic.
- Izbira ustrezne licence. Če je možno, uporabite strojno berljive licence, zapisane v [ODRL](#). Dober primer je storitev [Licentia](#), ki omogoča iskanje strojno berljivih licenc.
- Izbira repozitorija.
- Nadzor nad dostopom do podatkov.
- Promocija podatkov (raziskovalec, knjižnica, repozitorij, raziskovalna organizacija, financer).



# Primer ustrezno definirane licence

Title: **Odprti metapodatki slovenske nacionalne infrastrukture odprtega dostopa**





Authors: [Bezget, Jan](#) (Author)  
[Brezovnik, Janez](#) (Author)  
[Borovič, Mladen](#) (Author)  
[Ferme, Marko](#) (Author)  
[Ojsteršek, Milan](#) (Author)

Files:  [RAZ\\_Bezget\\_Jan\\_0.zip](#) (40,54 MB)  
Note: Metapodatki v metapodatkovnem formatu Dublin Core v obliki n-triples. Metadata in Dublin Core in n-triples format.  
 [RAZ\\_Bezget\\_Jan\\_0.zip](#) (40,54 MB)  
Note: Metapodatki v metapodatkovnem formatu Dublin Core v formatu RDF/XML. Metadata in Dublin Core in RDF format.  
 [RAZ\\_Bezget\\_Jan\\_0.zip](#) (3,72 MB)  
Note: Avtorji v formatu CSV. Authors in CSV format.  
 [RAZ\\_Bezget\\_Jan\\_0.zip](#) (36,76 MB)  
Note: Metapodatki v formatu CSV. Metadata in CSV format.  
 [RAZ\\_Bezget\\_Jan\\_0.zip](#) (321,10 KB)  
Note: Metapodatki o naborih raziskovalnih podatkov v formatu RDF/XML. Metadata about research data in RDF/XML format.  
 [RAZ\\_Bezget\\_Jan\\_0.zip](#) (1,72 MB)  
Note: Metapodatki o doktorskih disertacijah v formatu RDF/XML. Metadata about doctoral dissertations in RDF/XML format.  
 [RAZ\\_Bezget\\_Jan\\_0.txt](#) (2,95 KB)  
Note: Opis metapodatkov o naborih raziskovalnih podatkov. Description of metadata about research datasets.  
 [RAZ\\_Bezget\\_Jan\\_0.txt](#) (2,99 KB)  
Note: Opis metapodatkov o doktorskih disertacijah. Description of metadata about doctoral dissertations.  
 <http://openscience.si/OpenData.aspx>

Language: Slovenian  
Work type: Other (9)  
Typology: 2.20 - Complete Scientific Database or Corpus  
Organization: FERI - Faculty of Electrical Engineering and Computer Science  
Abstract: Podatkovna zbirka nacionalne infrastrukture odprtega dostopa je sestavljena iz metapodatkov in podatkov o avtorjih zaključnih del (diplome, magistrski in doktorati) in raziskovalnih publikacij slovenskih univerz in drugih slovenskih zbirk (dLib, Videolectures.NET, DKMORS, repozitorij SciVie...). Nacionalni portal (najdete ga na <http://www.openscience.si/>) agregira vsebine iz univerzitetnih repozitorijev in drugih slovenskih zbirk za potrebe skupnega iskalnika, priporočilnega sistema in detektorja podobnih vsebin. Podatki so dosegljivi v formatih CSV, RDF/XML in n-triples. Dodana je tudi ontologija, ki jo uporablja nacionalna infrastruktura odprtega dostopa. Po podatkovni zbirki OpenScience lahko povprašujete tudi preko spletnega vmesnika. Vsaka poizvedba vrne objekt JSON s spremenljivo količino lastnosti (odvisno od podatkov, ki so na voljo in tipa zahteve).

NUK URN: URN:SI:UM:DK:YO8DQRIX  
License: 

This work is available under this license: [Creative Commons Attribution 4.0 International](#)

Average score: ★★★★★ (0 votes)  
Your score: Voting is allowed only for [logged in users](#).  
Share:  SHARE   

#### Similar works from our repository:

1. [THE DEVELOPMENT OF NATIVE MOBILE APPLICATIONS FOR SEARCHING IN THE REPOSITORIES OF THE NATIONAL OPEN ACCESS INFRASTRUCTURE](#)
2. [Designing responsive web pages for repositories of national open access infrastructure](#)
3. [Collection of materials related to the national open access infrastructure](#)
4. [Organizacijski, tehnični in pravni vidiki vzpostavitve nacionalne infrastrukture odprtega dostopa](#)
5. [Establishing of a Slovenian open access infrastructure](#)

#### Similar works from other repositories:

1. [Bottlenecks in the Open-Access System](#)
2. [History of Open Access](#)
3. [Iniciativa odprtega dostopa v znanstvenem komuniciranju na področju gradbeništv in institucionalni repozitorij kot odziv nanjo - primer repozitorija DRUGG](#)
4. [Impact of open access on citation of scientific publications in civil engineering](#)
5. [Research on the impact of open access on the citation of scientific publications](#)

Hover the mouse pointer over a document title to show the abstract or click on the title to get all document metadata.

# Trajno hranjenje rezultatov raziskav

- Kateri podatki imajo dolgotrajno rednost in naj se hranijo?
- Prenos podatkov v format, ki je neodvisen od trenutne tehnologije.
- Zagotavljanje podatkovnega arhiva po [3-2-1 pravilu arhiviranja](#).
- Katere podatke je potrebno shraniti ali uničiti zaradi pogodbenih razmerij z dajalci podatkov?
- Kako dolgo naj se podatki hranijo ali ohranjajo?
- Kako bo poskrbljeno za dostop in varnost?

# Citiranje podatkov

**Navedba podatkovnega vira v publikaciji naj v splošnem vključuje naslednje informacije:**

- avtorja oz. avtorje raziskovalnih podatkov,
- ime nabora podatkov,
- producent (ustanova, pod okriljem katere so podatki nastali),
- kraj in leto izdelave podatkovnega nabora,
- verzija podatkovnega nabora,
- distribucija,
- serija,
- kraj in leto objave raziskovalnih podatkov v podatkovnem repozitoriju,
- podatkovni repozitorij, kjer do dostopni podatki,
- stalni identifikator.

# Povezave v zvezi z ravnanjem z raziskovalnimi podatki

- [CESSDA Data Management Expert Guide](#)
- [Data Management Skillbuilding Hub](#)
- [UK data service youtube channel](#)
- [Funders' data plan requirements](#)
- [DMPOnline](#)
- [Checklist for a Data Management Plan](#)
- [Consent for data sharing](#)
- [DMP Assistant](#)

?