

Päťhviezdičkové otvorené údaje v SR

5StarOpenData = LinkedData



Aplikovaný Sémantický Web

Miroslav Liška

Obsah

A. Problém + Riešenie

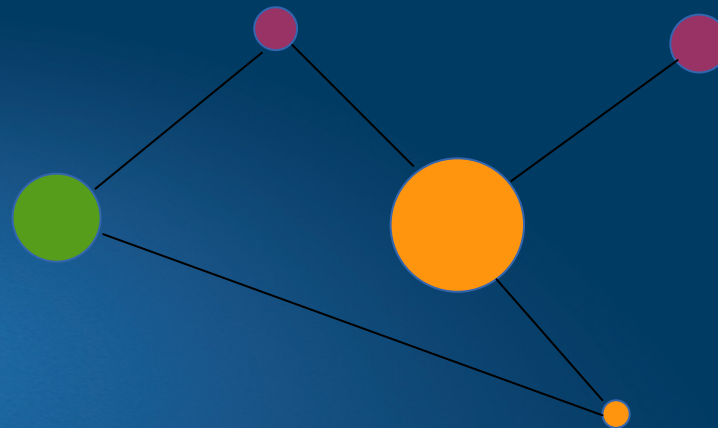
1. Zásadný problém GOV dát
2. Sémantický web
3. 5StarOpenData

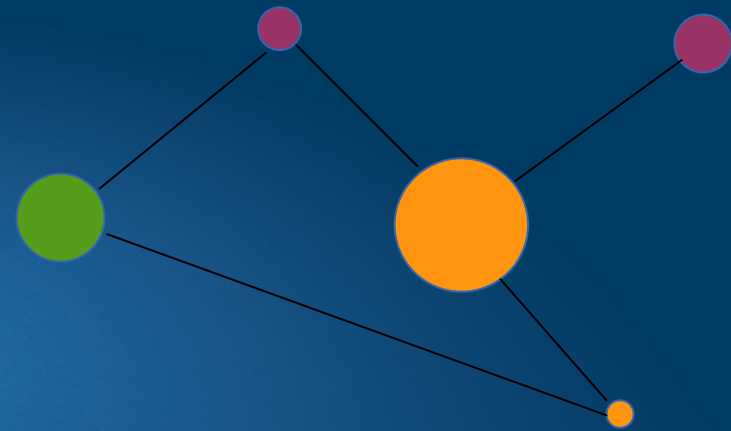
B. 5StarOpenData a Slovensko

4. Návrh: Sémantické datové štandardy GOV dát SR
5. Slovpedia, 6.Tripleskop, 7.PharmAX

C. Prínosy 5hviezdičkových dát

8. Štátna správa/samospráva a sémantický fairplay
9. Pozícia biznisu





A. Problem + Riesenie



1) Zásadný problém GOV dát

The screenshot shows the DATA.GOV.SK portal interface. At the top, there is a navigation bar with the logo and text 'PORTÁL OTVORENÝCH DÁT'. Below the logo, there are links for 'Hľadať', 'Skupiny', and 'O projekte'. The main content area is titled 'Hľadať - data.gov.sk' and features a search input field with the placeholder 'Hľadať...'. Below the search field, there are several data records, each with a title, description, and a 'CSV' download button. A lock icon and the text 'Nemá otvorenú licenciu' are present next to each record, indicating that the data is not open. The records include: 'Adresár poberateľov povinných výťažkov', 'Agrofilm – Register víťazných filmov medzinárodného festivalu', 'Bonitované pôdno-ekologické jednotky - BPEJ', 'Centrálna evidencia hospodárskych zvierat', 'Centrálny register geodetov', and 'Centrálny register poskytovateľov sociálnych služieb'. On the right side, there is a 'Tags' section with a list of tags and their counts, and a 'Resource formats' section with a list of formats and their counts. A network diagram is overlaid on the right side of the screenshot, consisting of several colored circles (green, orange, pink) connected by lines, representing a network structure.

DATA.GOV.SK
PORTÁL OTVORENÝCH DÁT

Hľadať Skupiny O projekte

Hľadať - data.gov.sk

Hľadať... Hľadať

Adresár poberateľov povinných výťažkov CSV CSV CSV
evidencia podľa zákona č.2012/1997 Z.z. Nemá otvorenú licenciu

Agrofilm – Register víťazných filmov medzinárodného festivalu CSV
Informácie pre vyhľadávanie filmov premietnutých na 28 ročníkoch filmového festivalu Agrofilm. Nemá otvorenú licenciu

Bonitované pôdno-ekologické jednotky - BPEJ CSV
Výskumný ústav pôdozvedectva a ochrany pôdy - Bonitované pôdnoekologické jednotky (BPEJ) sú ohraničené v poľnohospodárskej krajine na základe kvality pôdy (pôdno-substrátová jednotka,...) Nemá otvorenú licenciu

Centrálna evidencia hospodárskych zvierat CSV
Plemenárske služby Slovenskej republiky, š.p. Centrálna evidencia hospodárskych zvierat Aplikačné rozhranie: https://www.cehz.sk/ Nemá otvorenú licenciu

Centrálny register geodetov CSV CSV CSV CSV CSV CSV CSV CSV CSV
Zoznam geodetov oprávnených na vstup do štátnej dokumentácie a na cudzie nehnuteľnosti Nemá otvorenú licenciu

Centrálny register poskytovateľov sociálnych služieb CSV
Nemá otvorenú licenciu

Tags

- školsstvo (5)
- register (3)
- zmluvy (1)
- zdravotníctvo (1)
- zamestnanosť (1)
- upvsr (1)
- sudne (1)
- shmu (1)
- rozhodnutia (1)
- projekty (1)

Resource formats

- csv (131)
- xml (57)
- xsd (45)
- html (3)
- HTML (3)
- XLS (1)
- RTF (1)
- GZIP (1)
- CSV (1)

Groups

1) Zásadný problém GOV dát

The screenshot shows the DATA.GOV.SK portal with several search results. An orange callout box is overlaid on the page, containing the following text:

Dáta nie sú prepojené!!!
Neviem sa spýtať napr:
**Je niektorý geodet
poberateľom
povinného výtlačku?**

The search results visible on the page include:

- Adresár poberateľov povinných výtlačkov**
evidencia podľa zákona č.2012/1997 Z.z.
- Agrofilm – Register víťazných filmov medzinárodného**
Informácie pre vyberanie filmov premietnutých na 28 ročníkoch Agrofilm.
- Bonitované pôdno-ekologické jednotky - BPEJ**
Výskumný ústav pôdozvedectva a ochrany pôdy - Bonitované pôdnoekologické jednotky (BPEJ) sú ohraničené v poľnohospodárskej krajine na základe kvality pôdy (pôdno-substrátová jednotka,...
- Centrálna evidencia hospodárskych zvierat**
Plemenárske služby Slovenskej republiky, š.p. Centrálna evidencia hospodárskych zvierat. Aplikáčné rozhranie: <https://www.cehz.sk/>
- Centrálny register geodetov**
Zoznam geodetov oprávnených na vstup do štátnej dokumentácie a na cudzie nehnuteľnosti
- Centrálny register poskytovateľov sociálnych služieb**

Each result includes a 'CSV' button and a lock icon with the text 'Nemá otvorenú licenciu'. On the right side, there is a 'Resource formats' section with a list of file types and their counts:

- rozhodnutia (1)
- projekty (1)
- csv (131)
- xml (57)
- xsd (45)
- html (3)
- HTML (3)
- XLS (1)
- RTF (1)
- GZIP (1)
- CSV (1)

Below the formats is a 'Groups' section.

1) Zásadný problém GOV dát

The screenshot shows the DATA.GOV.SK portal interface. At the top, there is a search bar with the text "Hľadať - data.gov.sk" and a search input field containing "Hľadať...". Below the search bar, there are several search results, each with a title and a brief description. The results include:

- Adresár poberateľov povinných výťažkov**: evidencia podľa zákona č.2012/1997 Z.z.
- Agrofilm – Register víťazných filmov medzinárodného**: Informácie pre vyhľadávanie filmov premietnutých na 28 ročných Agrofilm.
- Bonitované pôdno-ekologické jednotky - BPEJ**: Výskumný ústav pôdozvedectva a ochrany pôdy - Bonitované pôdnoekologické jednotky (BPEJ) sú ohraničené v poľnohospodárskej krajine na základe kvality pôdy (pôdno-substrátová jednotka,...
- Centrálna evidencia hospodárskych zvierat**: Plemenárske služby Slovenskej republiky, š.p. Centrálna evidencia hospodárskych zvierat Aplikačné rozhranie: <https://www.cehz.sk/>
- Centrálny register geodetov**: Zoznam geodetov oprávnených na vstup do štátnej dokumentácie a na cudzie pohybnosti
- Centrálny register poskytovateľov sociálnych služieb**

Each result has a "CSV" button and a lock icon with the text "Nemá otvorenú licenciu". On the right side, there is a "Resource formats" section with a list of file formats and their counts:

- rozhodnutia (1)
- projekty (1)
- csv (131)
- xml (57)
- xsd (45)
- html (3)
- HTML (3)
- XLS (1)
- RTF (1)
- GZIP (1)
- CSV (1)

Below the formats, there is a "Groups" section.

An orange callout box is overlaid on the search results, containing the text:

Dáta nenesú význam!!!

Neviem si vyžiadať napr:

Chcem vidieť zoznam všetkých organizácií zo sídlom na Dunajskej ulici.

1) Zásadný problém GOV dát

The screenshot shows the DATA.GOV.SK portal interface. At the top, there is a search bar and navigation links. Below the search bar, several data records are listed, each with a title, description, and a 'CSV' download button. An orange callout box is overlaid on the right side of the page, containing the text: 'Dáta nie sú štandardizovane skatalogizované!!!' and 'Neexistuje jednotný menný priestor entít SR a ich identifikátory'. Arrows point from this box to the 'CSV' buttons of the data records. On the right side of the page, there is a 'Resource formats' section with a list of file formats and their counts, and a 'Groups' section.

Dáta nie sú štandardizovane skatalogizované!!!

Neexistuje jednotný menný priestor entít SR a ich identifikátory

• rozhodnutia (1)
• projekty (1)

Resource formats

- csv (131)
- xml (57)
- xsd (45)
- html (3)
- HTML (3)
- XLS (1)
- RTF (1)
- GZIP (1)
- CSV (1)

Groups

2) Sémantický web (Web 3.0)

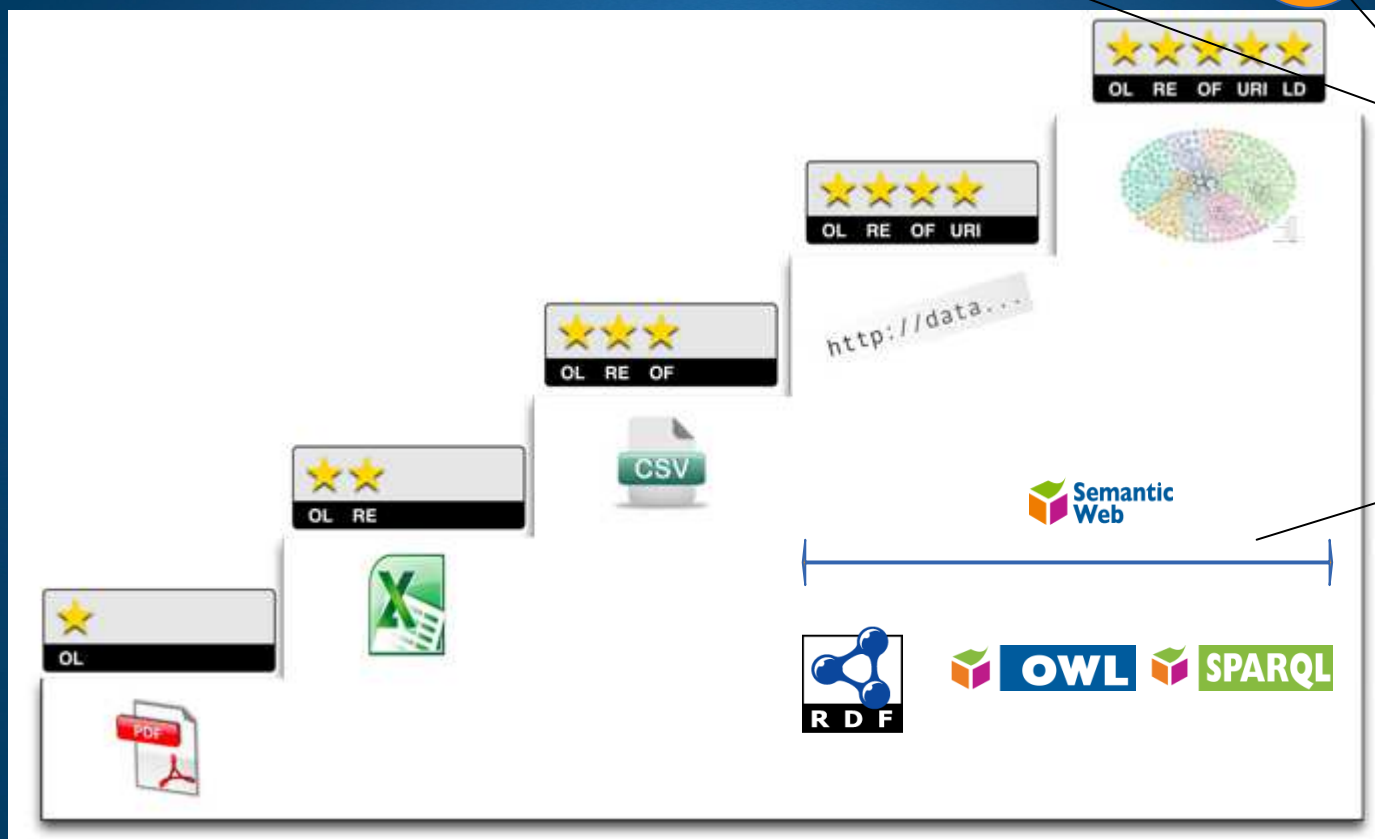


- nový spôsob spracovania, uchovávaní údajov s ohľadom na **význam (sémantiku)** dát (db. je **GRAF/triplet**)
- **dat' dátam význam** znamená popísať dáta štandardizovanými schémami/znalosťami = **ontológie**
- používanie unikátnych identifikátorov zdrojov (**URI**) na referencovanie ontológií a objektov sveta
 - s dátami sa pracuje ako s jedným celkom (mashup obsahu pre tvorbu aplikácií)
- **podpora strojového odvodzovania:**
čo z dát ešte vyplýva?



3) Trendy v štandardizácii GOV dát

<http://5stardata.info/>



2) Trendy štandardizácie GOV dát

An Official Web Site of the United States Government

Change Text Size SHARE

DATA.GOV
EMPOWERING PEOPLE

Search catalog Search

Open ID Login

HOME ABOUT DATA METRICS OPEN GOVERNMENT BLOGS COMMUNITIES

SANDY DAMAGE ESTIMATES BY BLOCK GROUP

U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT

Latest Datasets

- Library holdings for EX1304: Northeast...
- Fort Sumner 1:100000 Quad Hydrography...
- Santa Rosa 1:100000 Quad Hydrography ...
- Shiprock, NM 1:250,000 Quad East Half...
- Roswell 1:100000 Quad Hydrography DLGs
- Las Cruces 1:100000 Quad Transportat...
- Roswell, NM 1:250,000 Quad West Half ...
- Silver City, NM AZ 1:250,000 Quad USG...
- NM traffic count data for 2011
- Ruidoso 1:100000 Quad Hydrography DLGs

DATA AND TOOLS

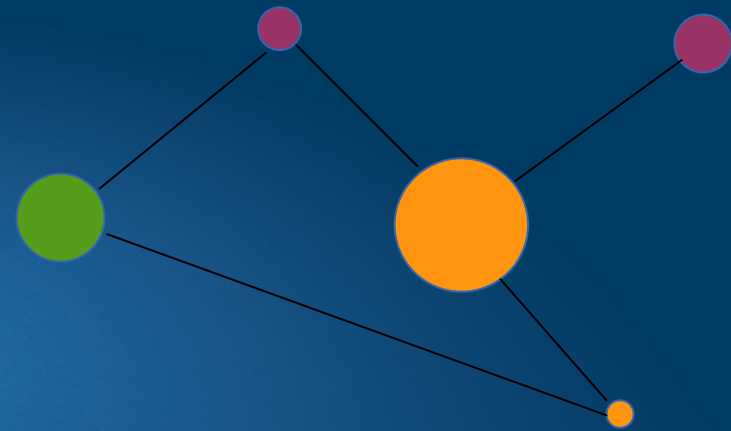
- 88,161 [datasets](#)
- 349 citizen-developed [apps](#)
- 137 [mobile apps](#)

SEMANTIC WEB

As the Web of linked documents evolves to include the Web of linked data, we're working to maximize the potential of Semantic Web technologies to

DEVELOPERS CORNER

Are you interested in sharing your mashups, apps, and ideas? Do you want to learn how to create app and mashups with some of the data hosted here on



B. 5StarOpenData a Slovensko



4) 5Hviezdičkové dáta a Slovensko

Návrh na dátové LinkedData štandardy SR

Proposed Data Semantic Annotation Using Standardized Web Ontologies and URI Identifiers to the Slovak Government



The document proposes standardized ontologies and URI identifier for the Slovak Government Data semantic annotation. The actual specification comprises semantic annotation of: **Person, Organization, Address, Administrative Unit, Drug** and **Geospatial Entity**. For the annotation purposes we use **RDF, OWL, SKOS, ADMS Vocabulary, Core Person & Business Vocabulary** together with **GeoNames Ontology, FOAF Vocabulary, WGS84 POS RDF Vocabulary, TMO Ontology, LAU/NUTS RAMON Ontology** and others. The document either proposes standardized URI identifier design pattern of a Slovak resource,

Príklady identifikátorov

BSK <http://data.gov.sk/id/region/SK01>

Betadine <http://data.gov.sk/id/drug/62321>



Nová verzia –
Január 2014

5. Slovpedia::Data

Slovpedia

★★★★★
Spojené otvorené dáta Slovenska

Úvod Data Znalosti Aplikácie API Prehliadanie O slovpédii Login

Aktuálne katalógy



Štátny úrad

Slovensko - NUTS úroveň 1	http://data.gov.sk/set/geo/00166197/statistics-sr-nuts1	16.10.2013
Slovensko - NUTS úroveň 0	http://data.gov.sk/set/geo/00166197/statistics-sr-nuts0	16.10.2013
Slovensko - NUTS úroveň 2	http://data.gov.sk/set/geo/00166197/statistics-sr-nuts2	16.10.2013
Slovensko - NUTS úroveň 3	http://data.gov.sk/set/geo/00166197/statistics-sr-nuts3	16.10.2013
Slovensko - LAU úroveň 1	http://data.gov.sk/set/geo/00166197/statistics-sr-lau1	16.10.2013
Slovensko - LAU úroveň 2	http://data.gov.sk/set/geo/00166197/statistics-sr-lau2	16.10.2013
Slovensko - ulice	http://data.gov.sk/set/geo/00166197/statistics-sr-street	16.10.2013



Ministerstvo zdravotníctva



Štátny ústav pre kontrolu liečiv

6. Tripleskop – analytický klient



File View Query Graph Resource

Sesame server: <http://data.talari.sk:8080/openrdf-sesame>
Repository:

Ibuprofén

6. Tripleskop – analytický klient



File View Query Graph Resource Settings Sesame server: http://data.gov.sk:8080/openrdf-sesame Repository: sl...

Exploring properties of <http://data.gov.sk/id/drug/48225>

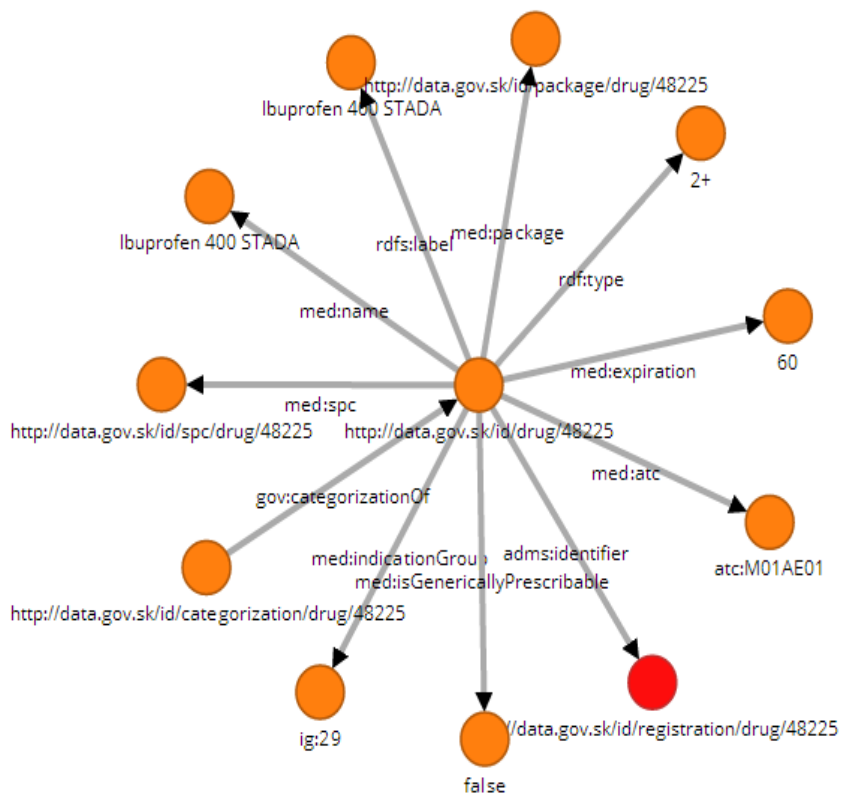
Predicate		Object	
rdf:type	🔍	drug	🔍
rdf:type	🔍	liek	🔍
rdf:type	🔍	medicínsky produkt	🔍
rdfs:label	🔍	Ibuprofen 400 STADA	🔍
ATC Classification	🔍	#M01AE01	🔍
ATC klasifikácia	🔍	#M01AE01	🔍
indication group	🔍	ANTIRHEUMATICA, ANTIPLHLOGISTICA, ANTIURATICA	🔍
indikačná skupina	🔍	ANTIRHEUMATICA, ANTIPLHLOGISTICA, ANTIURATICA	🔍
package	🔍	http://data.gov.sk/id/package/drug/48225	🔍
balenie	🔍	http://data.gov.sk/id/package/drug/48225	🔍
summary of product characteristcs	🔍	http://data.gov.sk/id/spc/drug/48225	🔍
súhrn charakteristických vlastností lieku	🔍	http://data.gov.sk/id/spc/drug/48225	🔍
identifier	🔍	http://data.gov.sk/id/registration/drug/48225	🔍
identifikátor	🔍	http://data.gov.sk/id/registration/drug/48225	🔍
expiration	🔍	60	🔍
expirácia	🔍	60	🔍
name	🔍	Ibuprofen 400 STADA	🔍
názov	🔍	Ibuprofen 400 STADA	🔍
med:isGenericallyPrescribable	🔍	false	🔍

6. Tripleskop – analytický klient

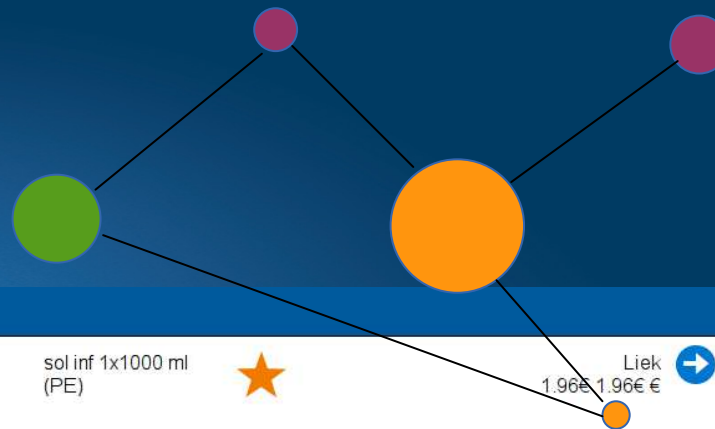


File View Query Graph Resource Settings

Sesame server: data.gov.sk:8080/openrdf-sesame
Repository:



7. PharmAX/Slovpedia



[-] Hľadaj

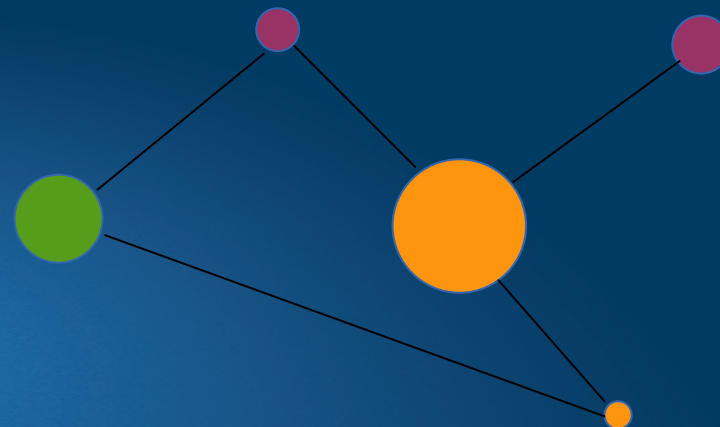
- [13602] 0,9% SODIUM CHL.IN W.F.I.FRESENIUS
atc: B05BB01
sol inf 1x500 ml (pp)
Liek
★ 1.49€ 1.15€ €
- [13584] 0,9% SODIUM CHL.IN.W.F.I.FRESENIUS
atc: B05BB01
sol inf 1x250 ml (sklo)
Liek
★ 1.08€ 0.95€ €
- [33596] 0,9% SODIUM CHL.IN W.F.I.FRESENIUS
atc: B05BB01

🔍 ★ 👤 ⚙️ ⓘ

[-] Hľadaj

[33596] 0,9% SODIUM CHL.IN W.F.I.FRESENIUS atc: B05BB01	sol inf 1x1000 ml (PE)	★	Liek 1.96€ 1.96€ €
[33593] 0,9% SODIUM CHL.IN W.F.I.FRESENIUS atc: B05BB01	sol inf 1x250 ml (PE)	★	Liek 1.29€ 0.95€ €
[33590] 0,9% SODIUM CHL.IN W.F.I.FRESENIUS atc: B05BB01	sol inf 1x500 ml (PE)	★	Liek 1.34€ 1.15€ €
[13599] 0,9% SODIUM CHL.IN W.F.I.FRESENIUS atc: B05BB01	sol inf 1x500 ml (sklo)	★	Liek 1.23€ 1.15€ €
[14455] 10% GLUKOSE IN WATER FOR INJECTION FRESENIUS atc: B05BA03	sol inf 1x250 ml (skl.fÅ%a.)	★	Liek 1.07€ 1.07€ €
[33635] 10% GLUKOSE IN WATER FOR INJECTION FRESENIUS atc: B05BA03	sol inf 1x500 ml (LDPE fÅ%a.)	★	Liek 1.88€ 1.31€ €
[14470] 10% GLUKOSE IN WATER FOR INJECTION FRESENIUS atc: B05BA03	sol inf 1x500 ml (PP fÅ%a.)	★	Liek 1.55€ 1.31€ €
[14473] 10% GLUKOSE IN WATER FOR INJECTION FRESENIUS atc: B05BA03	sol inf 1x500 ml (skl.fÅ%a.)	★	Liek 1.37€ 1.31€ €

🔍 ★ 👤 ⚙️ ⓘ

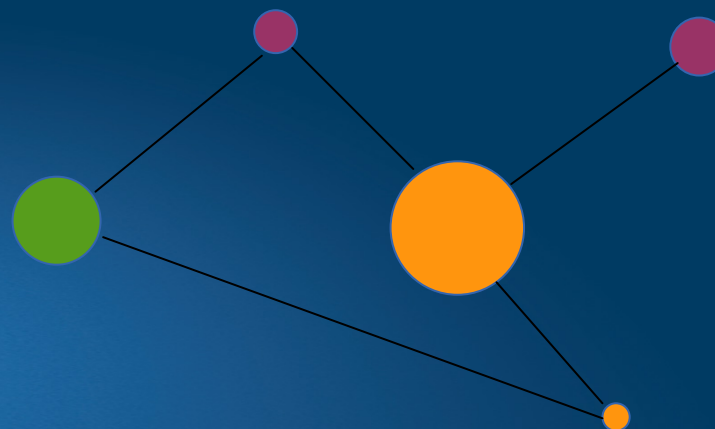


III. Prínosy OpenLinkedData pre štátnu správu / biznis?



Prínosy OpenLinkedData

8. Štátna správa & Samospráva



Fantastická dátová základňa
pre vyhľadávanie, prehliadanie dát,
tvorbu štatistík (analýz) - „Ako sa tu žije“

CHRBTICOVÉ Zefektívnenie fungovania dát
v štátnej správe

Lepšie predpoklady
na fair play



Prínosy OpenLinkedData

9. Pozícia biznisu

Občan

Poist'ovne

- Mapy záplavových území

Banky

- Klasifikácia BPEJ (December 2013)



Telco

- Meranie intenzity pokrytia 3G, júl 2013

Samospráva

- Prehľad školských zariadení BSK Marec 2013

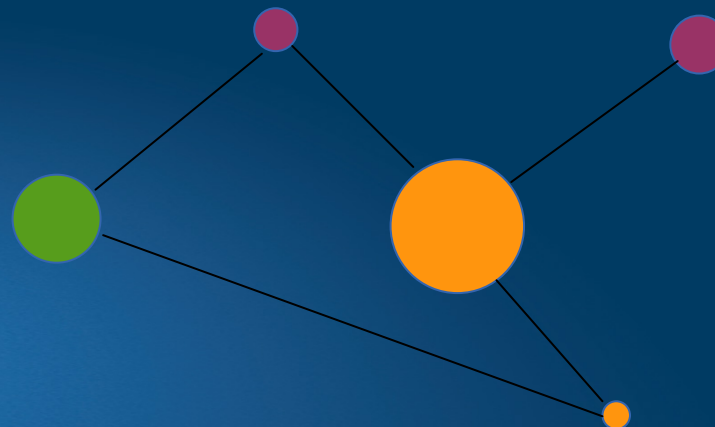


Sémantický fair play

Štátna správa

- Prehľad registrovaných liekov November 2012

VII. Referencie



<http://sk.linkedin.com/in/miroslavliska/>

<http://www.w3.org/standards/semanticweb/>

<https://joinup.ec.europa.eu/community/semic/description>

<http://www.openrdf.org/>

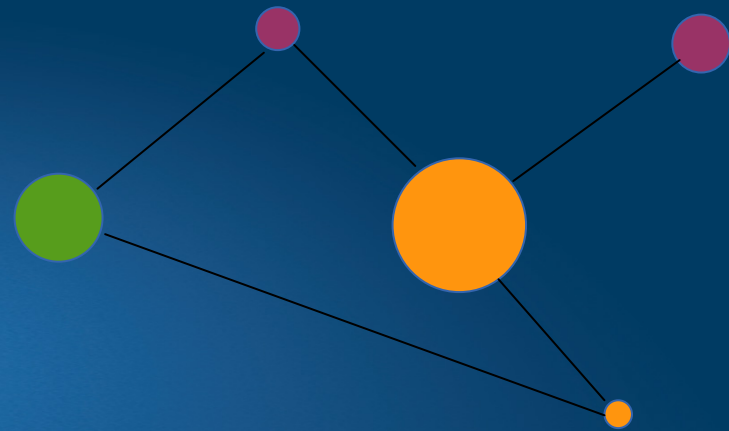
<http://www.w3.org/OWL/>

<http://www.w3.org/RDF/>

<https://joinup.ec.europa.eu/e-library/case/proposed-data-semantic-annotation-using-standardized-web-ontologies-and-uri-identifier>

Q1/2014

<http://www.slovpedia.sk/>



Ďakujem za pozornosť