

Ogrodje TextProc in njegova uporaba pri procesiranju besedil

Milan Ojsteršek

Laboratorij za heterogene računalniške sisteme
Univerza v Mariboru, Fakulteta za elektrotehniko,
računalništvo in informatiko

milan.ojstersek@um.si

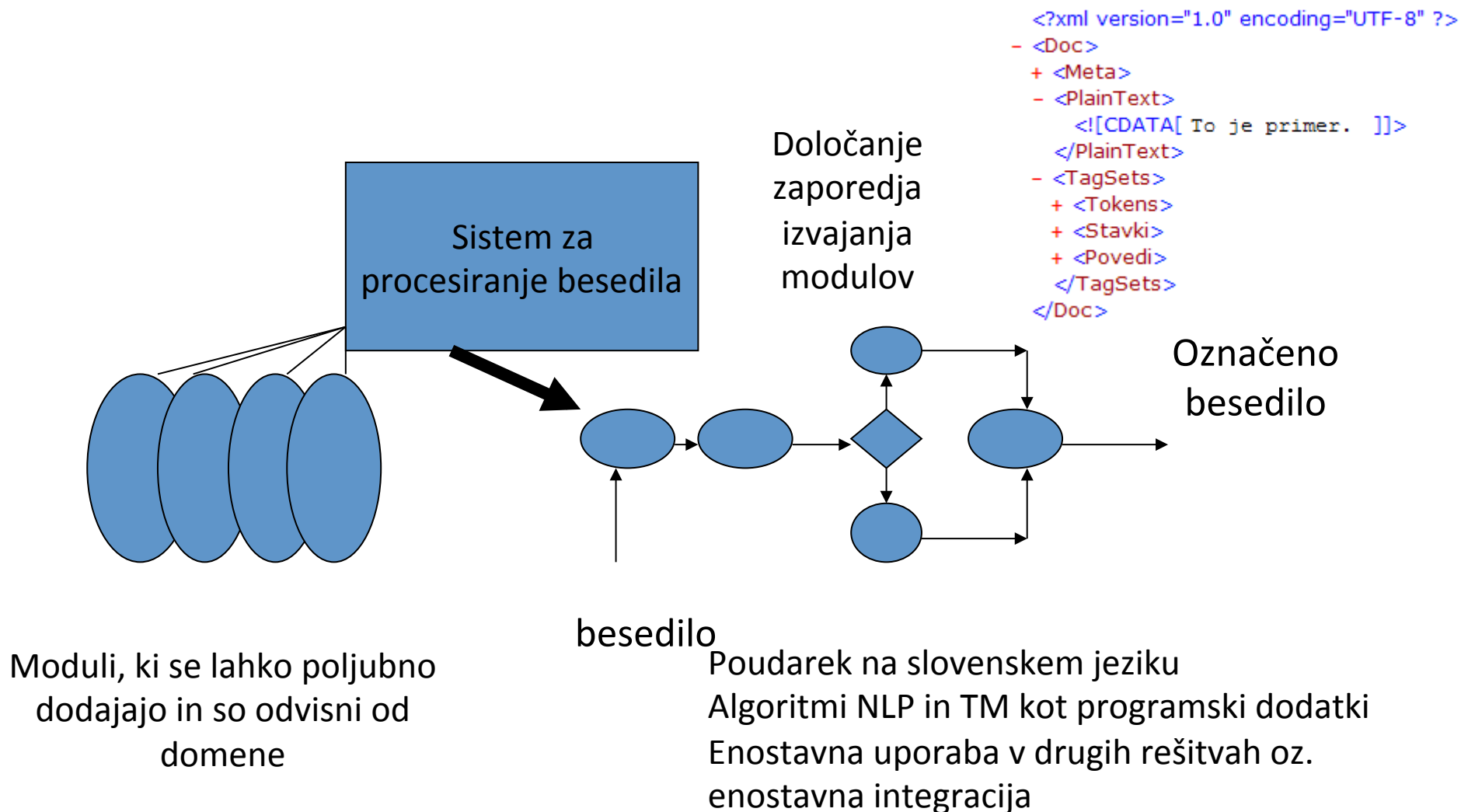
Vsebina predavanja

- O LHRS
- Ogrodje TextProc
- Jezikovna infrastruktura
- Nacionalna infrastruktura odprtega dostopa
 - Priporočilni sistem
 - Detekcija podobnih vsebin

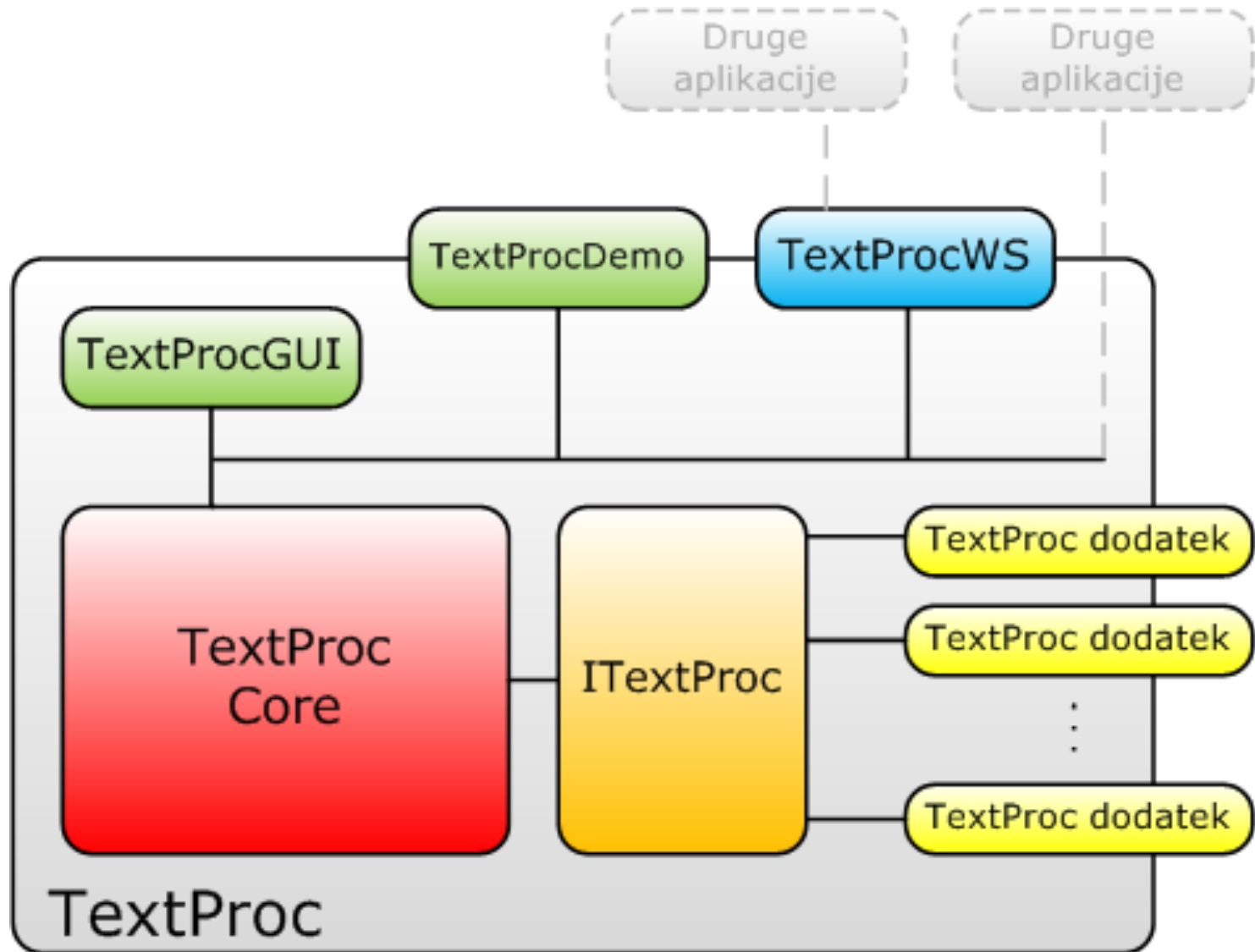
Opravila pri obdelavi naravnega jezika, uporabljena v naših aplikacijah

- Priprava besedila
- Razčlenjevanje
- Lematizacija
- Krnjenje
- Označevanje besednih vrst
- Pregibanje besed in besednih zvez
- Normalizacija sopomenk (sinonimov)
- Kategorizacija besedil
- Ekstrakcija entitet ali konceptov
- Določevanje pomena besed
- Razreševanje sklicev

Sistem za obdelavo besedila TextProc



TextProc



Primer rezultata obdelave besedila iz infrastrukture TextProc



<http://predator.uni-mb.si:8026/TextProc/>

Pomenski slovar

hose
gumijasta cev
Schlauch
flex sleeve of the cable
gumijasta uvodnica kabla
Kabelmantel

back
hrbet
Rücken
neon screwdriver
iskalnik faz
Spannungsprüfer
selector switch
izbirno stikalo
Bereichsumschalter

insulated blade
izolirani ploščati nastavek
isolierte Klinge
insulated handle
izolirani ročaj
isolierter Griff

screwdriver
izvijač
Schraubenzieher

core
jedro
Kern
acetylene cylinder
jeklenka za acetilen

Opis pojma

Predlogi relacij

Sorodni Pojmi

Povezave



Jezik: angleški

Naziv: screwdriver

Jezik: slovenski

Naziv: izvijač

Jezik: nemški

Naziv: Schraubenzieher





REPUBLIKA SLOVENIJA
MINISTRSTVO ZA GOSPODARSKI
RAZVOJ IN TEHNOLOGIJO

Naložba v vašo prihodnost

Operacijo delno financira Evropska unija
Evropski sklad za regionalni razvoj

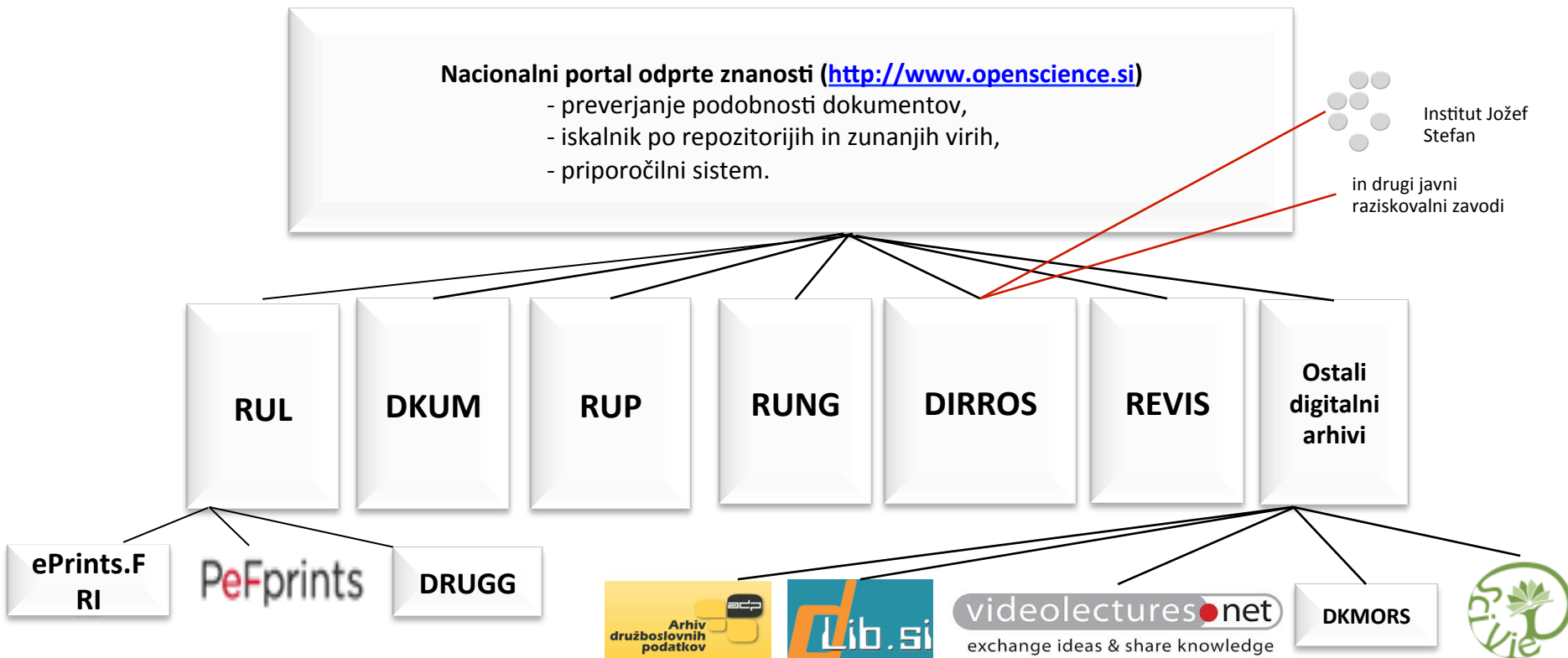


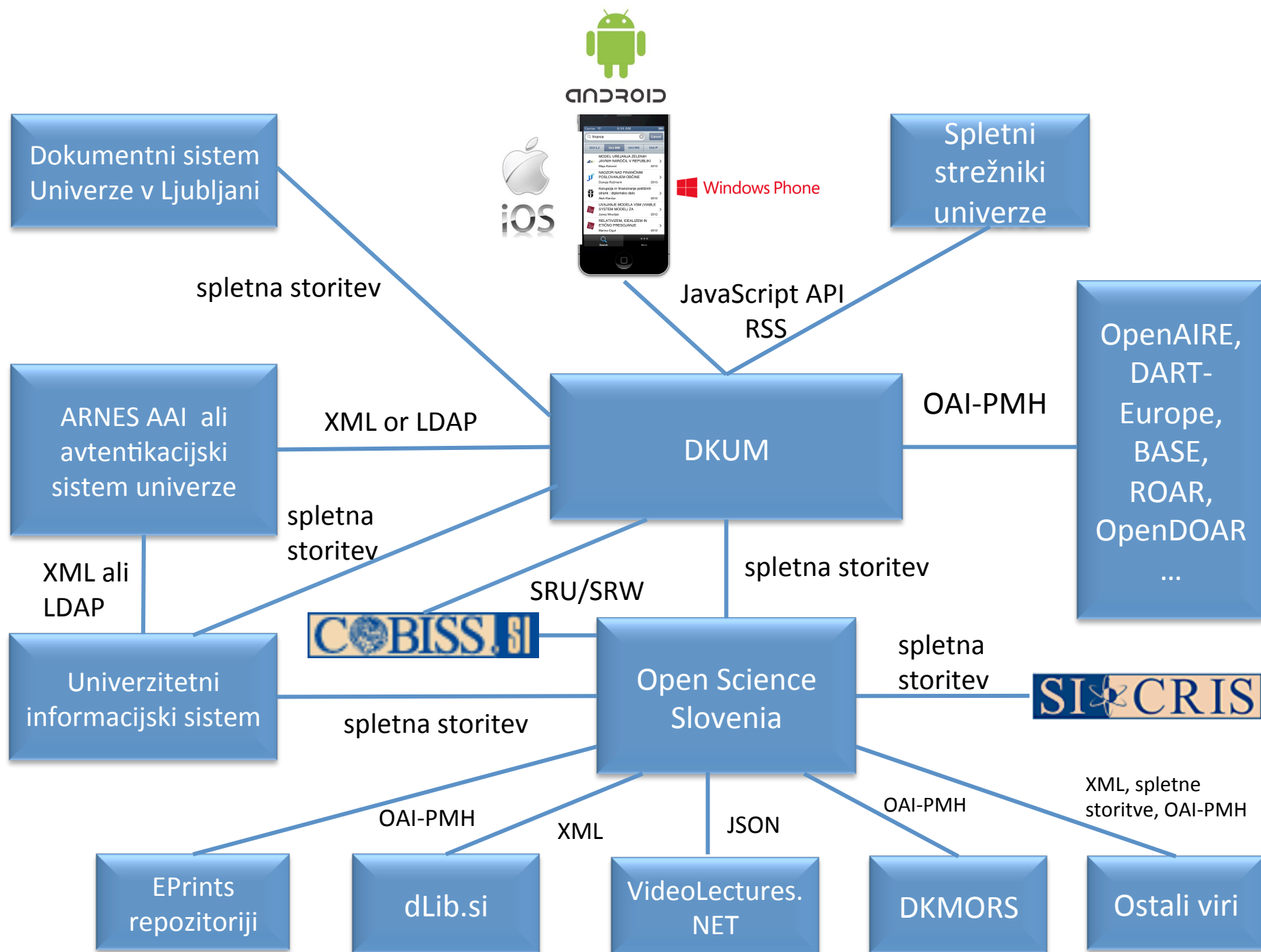
Investition in Ihre Zukunft

Operation teilweise finanziert von der Europäischen Union
Europäischer Fonds für regionale Entwicklung



Nacionalna infrastruktura odprtega dostopa





Priporočilni sistem

Izpis gradiva

Naslov: **Testiranje programske opreme**

Avtorji: [Moharič, Tadej](#) (Avtor)
[Kokol, Peter](#) (Mentor) [Več o mentorju...](#)

Datoteke:  [VS_Moharic_Tadej_2015.pdf](#) (1,83 MB)

Jezik: Slovenski jezik

Vrsta gradiva: Diplomsko delo/naloga (mb11)

Organizacija: FERI - Fakulteta za elektrotehniko, računalništvo in informatiko

Opis: V diplomskem delu predstavimo proces testiranja programske opreme. Na začetku podamo pregled zgodovine testiranja programske opreme ter osnov testiranja le-te, nato raziščemo različne stopnje in tipe testiranja ter kako najbolje dokumentiramo rezultate, da lahko razvijalci hitro najdejo in odpravijo najdene napake. Ko s tem končamo, predstavimo razvoj QT C++ aplikacije, ki jo lahko poganjamo na različnih operacijskih sistemih (Windows, Linux) in spremljamo porabo pomnilnika aplikacije, ki jo testiramo. Na koncu dela našo aplikacijo uporabimo za testiranje porabe pomnilnika same aplikacije na različnih operacijskih sistemih in predstavimo rezultate.

Ključne besede: [testiranje programske opreme](#), [razvoj programske opreme](#), [testiranje](#), [poraba pomnilnika](#), [ogrodje QT](#)

Leto izida: 2015

Izvor: Maribor

Število ogledov: 89

Število prenosov: 18

Metapodatki:        

Področja: Gradivo ni uvrščeno v področja.

Skupna ocena: ★★★★★ (0 glasov)

Vaša ocena: Ocenjevanje je dovoljeno samo [prijavljenim](#) uporabnikom.

Objavi na:  SHARE    ...

Podobna dela iz repozitorija:

1. [Uporaba orodja TestExpert za podporo testiranju](#)
2. [Model spletne aplikacije za e-učne inštrukcije](#)
3. [Podpora aktivnostim testiranja v okolju Microsoft Visual Studio 2013](#)
4. [Model razvoja programske opreme za notranjega naročnika](#)
5. [Ogrodje za napovedovanje napak programske opreme v agilnih okoljih](#)

Podobna dela iz ostalih repozitorijev:

1. [Primerjava ogrodij za testiranje enot v programskem jeziku C#](#)
2. [Avtomatsko funkcionalno testiranje programske opreme](#)
3. [Avtomatizacija testiranja kot ključ agilnosti razvoja programske opreme](#)
4. [Primerjava ogrodij za testiranje enot v programskem jeziku C](#)
5. [Ogrodja za avtomatsko izvajanje testnih aktivnosti](#)

Okapi BM25

$$tf(t, d) = ||n : t \in d||$$

$$n(t) = ||d \in D : t \in d||$$

$$idf(t) = \log \frac{||D|| - n(t) + 0.5}{n(t) + 0.5}$$

$$s(d, q) = \sum_{i=1}^{||q||} idf(q_i) \cdot \frac{tf(q_i, d) \cdot (k_1 + 1)}{tf(q_i, d) + k_1 \cdot \left(1 - b + b \cdot \frac{||d||}{A}\right)}, \quad q_i \in q$$

$$||d|| = \sum_{i=1}^n tf(t_i, d), \quad t_i \in d \quad \begin{matrix} k_1=2 \\ b=0,75 \end{matrix}$$

t beseda

d dokument

$tf(t, d)$ tf utež besede t v dokumentu d

$n(t)$ število dokumentov, v katerih se pojavlja beseda t

$idf(t)$ idf utež besede t v celotni zbirki dokumentov

D zbirka dokumentov

$||D||$ velikost zbirke D

q zahteva

q_i beseda v zahtevi

$||q||$ število besed v zahtevi

$s(d, q)$ ocena BM25 za dokument d in zahtevo q

$||d||$ število besed v dokumentu

A povprečno število besed v dokumentih iz zbirke D

$$itd(t_i, d_j) = \frac{tf(t_i, d_j)}{||d_j||}, \quad t_i \in d_j$$

itd (t_i, d_j) - index term

density – gostota

pojavljanja besedne zveze

ti znotraj dokumenta j

Vrste plagiatov

- Uporaba metode “Kopiraj – prilepi” brez citiranja originalnega vira.
- Prikrit prepis vira:
 - Uporaba istih besednih zvez ali strukture vira. Delno je vir citiran delno ni.
 - Sprememba besednih zvez iz vira. Obdrži se enak pomen vsebine.
- Agregiranje vsebine iz množice virov, nekateri so lahko celo neobstoječi. Nič v delu ni originalnega.
- Uporaba drugih besednih zvez v vsebini (parafraziranje) brez navedbe vira.
- Kraja ideje.
- Uporaba avtorjevih predhodnih del brez citiranja (samoplariat , recikliranje).
- Prevajanje besedila iz drugega jezika brez citiranja vira.

Povzeto po Plagiarism.org:

<http://plagiarism.org/plagiarism-101/types-of-plagiarism> , obiskano 7. 5. 2015.

Kako delujejo programi za preverjanja plagiatorstva?

- Izračun zgoščenih vrednosti na n besedah.
- Izračun zgoščenih vrednosti z uporabo tehnik procesiranja naravnega jezika.
- Določanje vreče besed iz dokumenta in uporaba statističnih metod primerjanja podobnosti dokumentov ali odstavkov.
- Ugotavljanje zaporedja citiranja.
- Ugotavljanje stila avtorja.
- Uporaba pomenskih podpisov na odstavkih besedila.
- Določanje najdaljših skupnih podnizov med dvema dokumentoma.
- Uporaba tehnik prevajanja programske kode.

Detektor podobnih vsebin

[odjava] [Iskanje po gradivih...](#)

Predstavitev UM 15. 9. 2015 (1) \ Testni dokument.doc \ Testni dokument.doc

znižanje ravnovesne količine.
Spremembe dejavnikov ponudbe in povpraševanja povzročijo tržna stanja, kjer lahko enkrat prevladuje presežno povpraševanje, ki sili cene navzgor ali pa presežna ponudba, ki si cene navzdol...

.....
Drugi del vsebine

Mednarodni standardi ocenjevanja vrednosti (MSOV) Odbor za Mednarodne standarde ocenjevanja vrednosti (OMSOV) iz leta 2011 pojma tržne vrednosti ne opredeljujejo, natančno pa opredeljujejo pojem »tržne vrednosti«.
Že v poglavju, ki vsebuje opredelitve pojmov MSOV, je zaslediti naslednjo definicijo tržne vrednosti:
»Tržna vrednost - ocenjeni znesek, za katerega naj bi voljan kupec in voljan prodajalec zamenjala sredstvo ali obveznost na datum ocenjevanja vrednosti v poslu med nepovezanima in neodvisnima strankama po ustreznem trženju in pri čemer sta stranki delovali seznanjeno, preudarno in brez prisile«.

V naslednjem poglavju, ki določa t.i. okvir MSOV (le-ta vključuje splošno sprejete zasnove, načela in opredelitve pojmov ocenjevanja vrednosti na katerih temeljijo mednarodni standardi ocenjevanja vrednosti), se definicija v točki 30 ponovi, v točki 31 pa so izrazi iz osnovne definicije podrobneje obrazloženi.

Tako se izraz »ocenjeni znesek« nanaša na ceno, izraženo v denarju plačljivo za sredstvo v poslu med nepovezanima in neodvisnima strankama. Tržna vrednost je najverjetnejša cena, ki jo lahko razumno dosežemo na trgu na datum ocenjevanja vrednosti v skladu z opredelitvijo pojma tržna vrednost. To je najboljša cena, ki jo prodajalec lahko razumno doseže, in najugodnejša cena, ki jo kupec lahko razumno doseže. Ta ocena izrecne izključuje ocenjeno ceno, ki je povišana ali znižana zaradi posebnih pogojev ali okoliščin, kot so neobičajno financiranje, dogovori o prodaji in povratnem najemu, posebne obravnave ali ugodnosti, ki jih podeli nekdo, ki je povezan s prodajo, ali katera koli prvina posebne vrednosti.

Izraz »zamenjala sredstvo« se nanaša na dejstvo, da je vrednost sredstva ocenjeni znesek in ne vnaprej določeni znesek ali dejanska prodajna cena. To je cena v poslu, ki izpolnjuje vse sestavine opredelitve pojma tržna vrednost na datum ocenjevanja vrednosti.

Časovna opredelitev »na datum ocenjevanja vrednosti« zahteva, da je vrednost časovno vezana na določen dan, ker se trgi in tržne razmere lahko spreminjajo. Je ocenjena vrednost ob drugem času lahko napačna ali neprimerna. Znesek vrednosti bo izražal dejansko stanje na trgu in okoliščine na določen datum ocenjevanja vrednosti, ne pa na neki pretekli ali prihodnji dan. Opredelitev tudi predvideva hkratno zamenjavo in

Podobni dokumenti | Podobni nizi | Izseki | PDF poročilo | Prenos | Pomoč

Podobni dokumenti

dokumenta:

Testni dokument.doc

Predstavitev UM 15. 9. 2015 (vir)

Pokritost dokumenta: 74.46 %

Najdeno v 50 dokumentih.

☐ Izberi vse

- ☐ 1. VPLIV SOCIALNEGA KAPITALA SLOVENSKE DINAMIČNIH PODJETIJ - GAZEL NA USPEŠNOST
[Primerjava povedi] [Znakovna primerjava] 50,34 %
Marcel Rojko, Karin Širec (mentor), 2012, Ekonomsko-poslovna fakulteta
- ☐ 2. MSOV-2011-SLO.pdf
[Primerjava povedi] [Znakovna primerjava] 16,51 %
Matija Polajnar, 0, FGG Moodle (vir)
- ☐ 3. SESTO POGLAVJE DEFINICIJE VREDNOSTI 2015.pdf
[Primerjava povedi] [Znakovna primerjava] 16,40 %
Maruška Šubic-Kovač, 0, FGG Moodle (vir)
- ☐ 4. Mednarodni standardi ocenjevanja vrednosti

Hvala za pozornost!

