

VPP „SOPHIS” 2.projekts
**„Uz ontoloģijām balstītas tīmekļa videi
pielāgotas zināšanu inženierijas
tehnoloģijas”**

1.posms

vadītājs:

Dr.habil.sc.comp., profesors J.Bārzdiņš

VPP „SOPHIS” 2.projekts

„Uz ontoloģijām balstītas tīmekļa videi pielāgotas zināšanu inženierijas tehnoloģijas”

Apakšprojekti, organizācijas, galvenie izpildītāji:

- Uz ontoloģijām balstītas tīmekļa videi pielāgotas modelēšanas tehnoloģijas un rīki zināšanu analīzei, t.sk. Uz ontoloģijām balstītas dabīgās valodas semantikas izgūšanas metodes (**LU MII** : J.Bārzdiņš, A.Šostaks, E.Rencis, M.Grasmanis, A.Sproģis, G.Bārzdiņš, D.Goško, P.Paikens)
- Semantiskā tīmekļa tehnoloģijas zināšanu formalizācijai, vairākkārtīgai izmantošanai un koplietošanai (**RTU DITF** : J.Grundspenķis, V.Graudiņa, M.Kirikova, A.Novickis)
- Jaunas tīmekļa videi piemērotas datu atlasīšanas un vizualizācijas metodes un ar tām saistītas problēmas (**LU DF** : J.Bičevskis, G.Arnicaņš, Ģ.Karnītis)

Formālā puse par projektu:

Izpildīti



Projekta Nr.	Projekta vadītājs	Izpildītājs	Veicamie uzdevumi	Rezultatīvie rādītāji
2.	Jānis Bārzdīņš	LUMII	1. Uz ontoloģijām un tīmekļa tehnoloģijām balstītas grafiskas ātro vaicājumu valodas teorētisko pamatu izstrāde.	Sagatavota zinātniskā publikācija
			2. Uz ontoloģijām, tīmekļa tehnoloģijām un kontrolētas dabīgās valodas balstītas ātro vaicājumu valodas, kas „pa tiešo” (bez programmētāja starpniecības) būs lietojama nozaru speciālistiem, izstrāde.	Valodas apraksts
			3. Grūti formalizējamu sistēmu tīmeklī balstītu modelēšanas rīku būves metožu un tehnoloģiju izstrāde.	Metodikas apraksts
			4. Uz ontoloģijām balstītu saistīto datu tehnoloģiju izpēti e-pārvaldes un e-medicīnas lietojumiem.	Esošās situācijas apraksts
	Jānis Grundspenķis	RTU DITF	5. Uz lieliem semantiskiem grafiem (piemēram, BabelNet) un notikumu n-āru relāciju grafiem (piemēram, AMR, FrameNet) balstītu dabīgās valodas saprašanas (language understanding) metožu teorētisko pamatu izstrāde.	Sagatavota zinātniskā publikācija
			1. Zināšanu struktūras modeļu analīzes metodes un algoritmu, kā arī to atbalsta programmatūras prototipa izstrāde	Metodes un algoritma apraksts Sagatavota zinātniskā publikācija
			2. Saistīto darbu izpēti procesu, uzņēmumu arhitektūru un zināšanu strukturālās savietojamības jomā un ideālā sasaistes modeļa sākotnējās skices izstrāde	Zinātniskās atskaite
	Jānis Bičevskis	LU DF	3. Semantiskā tīmekļa servisu integrēšanas e-loģistikas portālā demonstrācijas prototipa izstrāde	Sagatavota zinātniskā publikācija
			1. Attīstīt metodes liela apjoma datu pieejamībai, kas balstītas uz modeļiem un nozares ontoloģijām, piedāvājot jaunas tīmekļa videi piemērotas datu atlasīšanas un vizualizācijas metodes.	Sagatavota zinātniskā publikācija
			2. Attīstīt biznesa procesu modeļu pielietošanas metodes programmas izpildes laika notikumu analīzē, lai paaugstinātu informācijas sistēmu drošības līmeni.	Sagatavota zinātniskā publikācija

Publikācijas:

- [1] Barzdins G., Paikens P., Gosko D.. Riga: from FrameNet to Semantic Frames with C6.0 Rules. SemEval 2015 Task 18: Semantic Dependency Parsing. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), Association for Computational Linguistics, pp. 960–964. (<http://www.aclweb.org/anthology/S15-2160>)
- [2] J. Barzdins, E. Rencis and A. Sostaks, **Fast Ad Hoc Queries Based on Data Ontologies**. In: H.M. Haav, A. Kalja, T. Robal (Eds.), *Frontiers of AI and Applications*, Vol. 270, Databases and Information Systems VIII, IOS Press, pp. 43-56, 2014. (**SCOPUS**)
- [3] *Rudolfs Bundulis, and Guntis Arnicans, **Concept of virtual machine based high resolution display wall**. Information, Electronic and Electrical Engineering (AIEEE), 2014 IEEE 2nd Workshop on Advances in. pp. 1-6, IEEE, 2014. DOI: 10.1109/AIEEE.2014.7020317, <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7020317>(**IEEEExplore**)*
- [4] *I.Oditis, J.Bicevskis Asynchronous Runtime Verification of Business Processes 7th International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN'2015) pp. 103-108, 2015 (būs SCOPUS)*
- [5] Kirikova, M. Enterprise Architecture and Knowledge Perspectives on Continuous Requirements Engineering. Proceedings of REFSQ-2015 Workshops, Research Method Track, and Poster Track co-located with the 21st International Conference on Requirements Engineering: Foundation for Software Quality. Essen, Germany, March 23, 2015. CEUR-WS.org, Vol. 1342, ISSN 1613-0073, pp. 44-51. (**SCOPUS**)
- [6] Bartusevičs, A., Novickis, L., Lesovskis, A. Model-Driven Software Configuration Management and Semantic Web in Applied Software Development. No: Recent Advances in Telecommunications, Informatics and Educational Technologies: Proceedings of the 13th International Conference on Telecommunications and Informatics (TELE-INFO '14) , Turcija, Istanbul, 15.-17. decembris, 2014. Istanbul: WSEAS Press, 2014, 108.-116.lpp. ISBN 978-1-61804-262-0. (būs SCOPUS un ISI)
- [7] Bartusevičs, A., Lesovskis, A., Novickis, L. Semantic Web Technologies and Model-Driven Approach for the Development and Configuration Management of Intelligent Web-Based Systems. No: Proceedings of the 2015 International Conference on Circuits, Systems, Signal Processing, Communications and Computers, Austrija, Vienna, 15.-17. marts, 2015. Vienna: 2015, 32.-39.lpp. ISBN 978-1-61804-285-9. ISSN 1790-5117. (būs SCOPUS un ISI)
- [8] *Rudolfs Bundulis, and Guntis Arnicans. **Virtual Machine Based High Resolution Display Wall: Experiments on Proof of Concept**. International Conference on Systems, Computing Sciences and Software Engineering (SCSS 14) , Electronic CISSE 2014 Conference Proceedings, 2014. (tiks publicēts Innovations and Advances in Computer, Information, Systems Sciences, and Engineering. LNEE, Springer)*
- [9] *Posters: Rudolfs Bundulis, Guntis Arnicans, and Rihards Gailums. **NVENC Based H.264 Encoding for Virtual Machine Based Monitor Wall Architecture**, GPU Technology Conference, San Jose, Mart 17-20, 2015. http://ondemand.gputechconf.com/gtc/2015/posters/GTC_2015_Visualization_Large_Scale_Multi_Display_01_P5174_WEB.pdf*

Promocijas darbi:

[1p] Armands Šlihte. Integrētā priekšmetiskās vides modelēšana: pieeja un rīku kopa topoloģiskā funkcionēšanas modeļa iegūšanai. (Aizstāvēts RTU promocijas padomē P-07 8.06.2015)

Maģistra darbi:

[1m] Svetlana Otto. Daudzaģentu sistēmu un ontoloģiju projektēšanas metožu integrēšanas metodoloģiju analīze (Zinātniskais vadītājs prof. J.Grundspenķis. Aizstāvēts RTU DITF 15.06.2015)

[2m] Emil Bikjanov. Zināšanu bāzu un deduktīvo datu bāzu salīdzināšanas analīze (Zinātniskais vadītājs prof. J.Grundspenķis. Aizstāvēts RTU DITF 15.06.2015)

[3m] Andrejs Gaidukovs. Laika dimensija biznesa procesa un uzņēmumu arhitektūras modelēšanā (Zinātniskā vadītāja prof. M.Kirikova. Aizstāvēts RTU DITF 16.06.2015)

[4m] Edgars Kirkiļevičs. Sadarbojošos procesu testēšana (Zinātniskais vadītājs prof. J.Bičevskis. Aizstāvēts LU DF 02.06.2015)

[5m] Dace Damberga. Uzraudzītas mašīnmācīšanās klasifikatoru izpēte un empīriskā salīdzināšana (Zinātniskais vadītājs prof. Guntis Bārzdiņš. Aizstāvēts LU DF 05.06.2015)

[6m] Zane Siliņa. Latviešu valodas sintaktiskā analizatora 'Čankeris' modernizācija (Zinātniskais vadītājs prof. Guntis Bārzdiņš. Aizstāvēts LU DF 01.06.2015)

Apakšprojekts 3.1

Uz ontoloģijām balstītas tīmekļa videi pielāgotas modelēšanas tehnoloģijas un rīki zināšanu analīzei, t.sk. uz ontoloģijām balstītas dabīgās valodas semantikas izgūšanas metodes (uzdevumi 1, 2, 3, 4, 5)

Pakavēšos tikai pie trim galvenajiem rezultātiem

1. Ātro vaicājumu valoda VV, tās eksperimentāla realizācija, pārbaude uz reāliem datiem, efektivitātes apstiprinājums (uzdevumi 1, 2)

Pētīta „3-How” problēma:

1. Kā attēlot datu ontoloģiju, lai tā būtu viegli saprotama nozares ekspertam,
2. Kā uz šīs ontoloģijas pamata uzbūvēt nozares speciālistam ‘pa tiešo’ (bez programmētāja starpniecības) viegli lietojamu vaicājumu valodu,
3. Kā efektīvi realizēt šādu vaicājumu valodu, lai atbildi iegūtu aptuveni sekundes laikā uz vairāku Gigabaitu apjoma datiem (piemēram, BKUS viena gada datiem).

Ideāls risinājums acīmredzot neeksistē.

Datu ontoloģijas jāierobežo, varbūt tad. Bet:

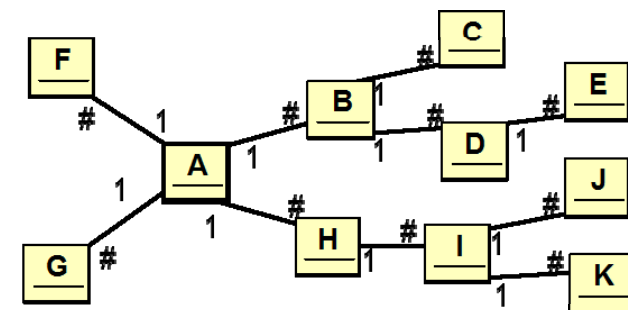
1. Ierobežojumiem jābūt «nelieliem» no praktiskā viedokļa,
2. Ontoloģijām jābūt «viegli» saprotamām nozares speciālistiem,
3. Vaicājumu valodai jābūt «viegli» saprotamai nozares speciālistiem, kuri «never took Com Sci 101», Berner-Lee, 2001,
4. Vaicājumu valodai jābūt pietiekoši «bagātai» (t.i. neder 90%, jo tad 10% gadījumu tāpat būs vajadzīgs programmētājs),
5. Realizācijai – tādai, ka atbilde uz reāliem datiem «1 sekundes laikā»

Mūsu pirmā ideja: Granuālas ontoloģijas -> Grafiska vaicājumu valoda

Nākamā ideja: Granuālas ontoloģijas -> Zvaigņveida ontoloģijas

Publikācija:

J. Barzdins, E. Rencis and A. Sostaks, **Fast Ad Hoc Queries Based on Data Ontologies**.
In: Frontiers of AI and Applications, Vol. 270, Databases and Information Systems VIII,
IOS Press, pp. 43-56, 2014. <http://ebooks.iospress.nl/volumearticle/38187> (SCOPUS)



In: The Semantic Web – ISWC 2014 →

How Semantic Technologies can Enhance Data Access at Siemens Energy

Evgeny Kharlamov^{1,2}, Nina Solomakhina², Oleg Zgur'ev³, Dmitriy Zheleznyakov¹, Thomas Hubauer², Steffen Lamparter², Mikhail Roshchin², and Ahmet Soylu⁴

¹ University of Oxford, UK, ² Siemens Corporate Technology, Germany,

³ Hamburg University of Technology, Germany, ⁴ University of Oslo, Norway

Abstract. We present a description and analysis of the data access challenge in the Siemens Energy. We advocate for Ontology Based Data Access (OBDA) as a suitable Semantic Web driven technology to address the challenge. We derive requirements for applying OBDA in Siemens, review existing OBDA systems and discuss their limitations with respect to the Siemens requirements. We then introduce the Optique platform as a suitable OBDA solution for Siemens. Finally, we describe our preliminary installation and evaluation of the platform in Siemens.

1 Introduction

The growth of available information in enterprises requires new efficient methods for data access by domain experts whose ability to analyse data is at the core of making business decisions. Current centralised approaches, where an IT expert translates the requirements of domain experts into Extract-Transform-Load (ETL) processes to integrate the data and to apply predefined analytical reporting tools, are too heavy-weight and inflexible [1]. In order to support interactive data exploration, domain experts therefore want to access and analyse available data sources directly, without IT experts being involved.

This direct data access is particularly important for Siemens Energy¹ that runs several service centres for power plants. The main task of a service centre is remote monitoring and diagnostics of many thousands appliances, such as gas and steam turbines, generators, and compressors installed in plants. Monitoring and diagnostics are performed by service engineers and are typically conducted in four steps: (i) engineers receive a notification about a potential or detected issue with an appliance, (ii) they gather data relevant to the case, (iii) analyse the data, and finally (iv) report about ways to address the issue to the appliance owner. Currently, Step (ii) of the process is the bottleneck consuming up to 80% of the overall time needed by the engineer to accomplish the task. The main reason for this time consumption is the indirect data access, i.e., in many cases the engineers have to access data via IT experts. Involvement of IT experts in the process slows it down dramatically due to reasons including the overload of IT experts and miscommunication between them and the engineers.

Enabling direct data access for engineers in Siemens is a challenging task primarily due to the Big Data dimensions as well as the conceptual mismatch between the language and structures that the engineers use to describe the data, and the way the data is actually

Ko saka citi →

! → ^y The research was supported by the FP7 grant Optique (n. 318338).

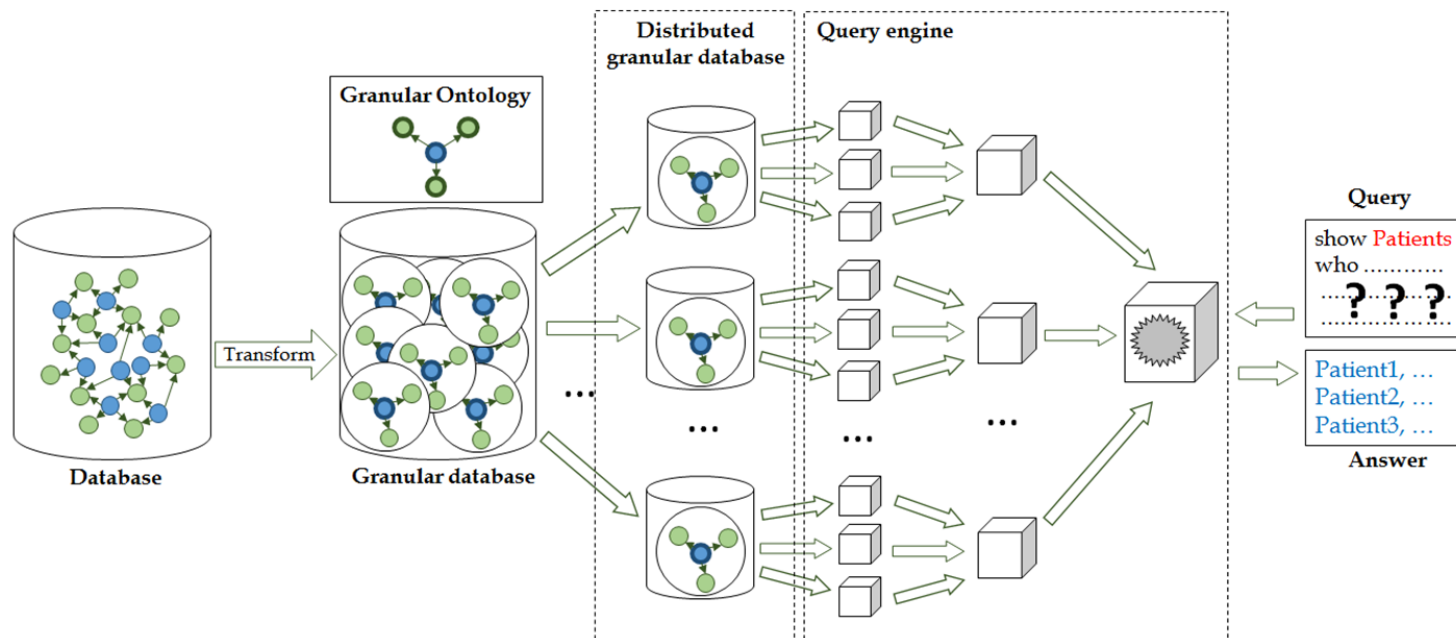


Figure1

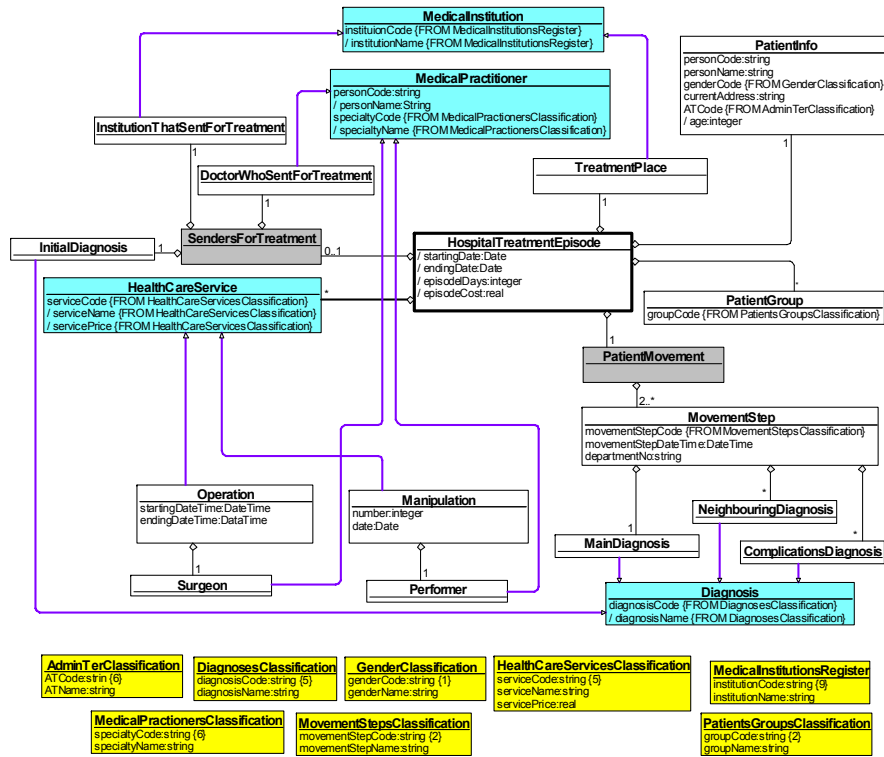


Fig. 4. Star ontology capturing data from the hospital patient card.

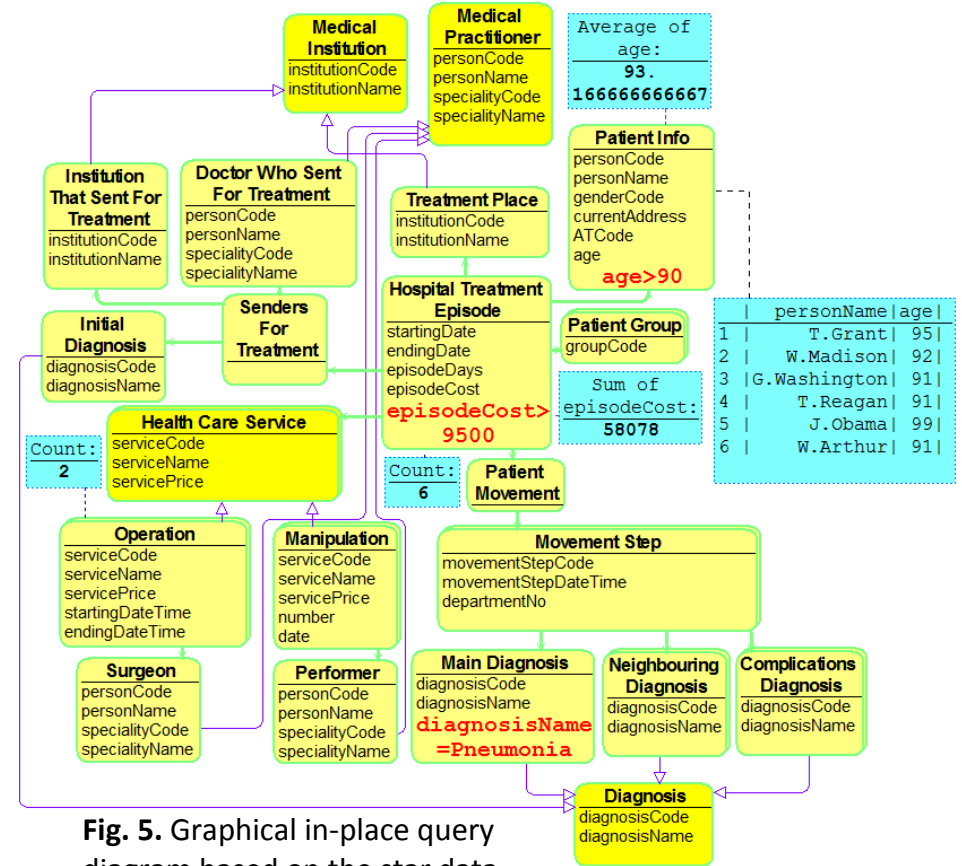
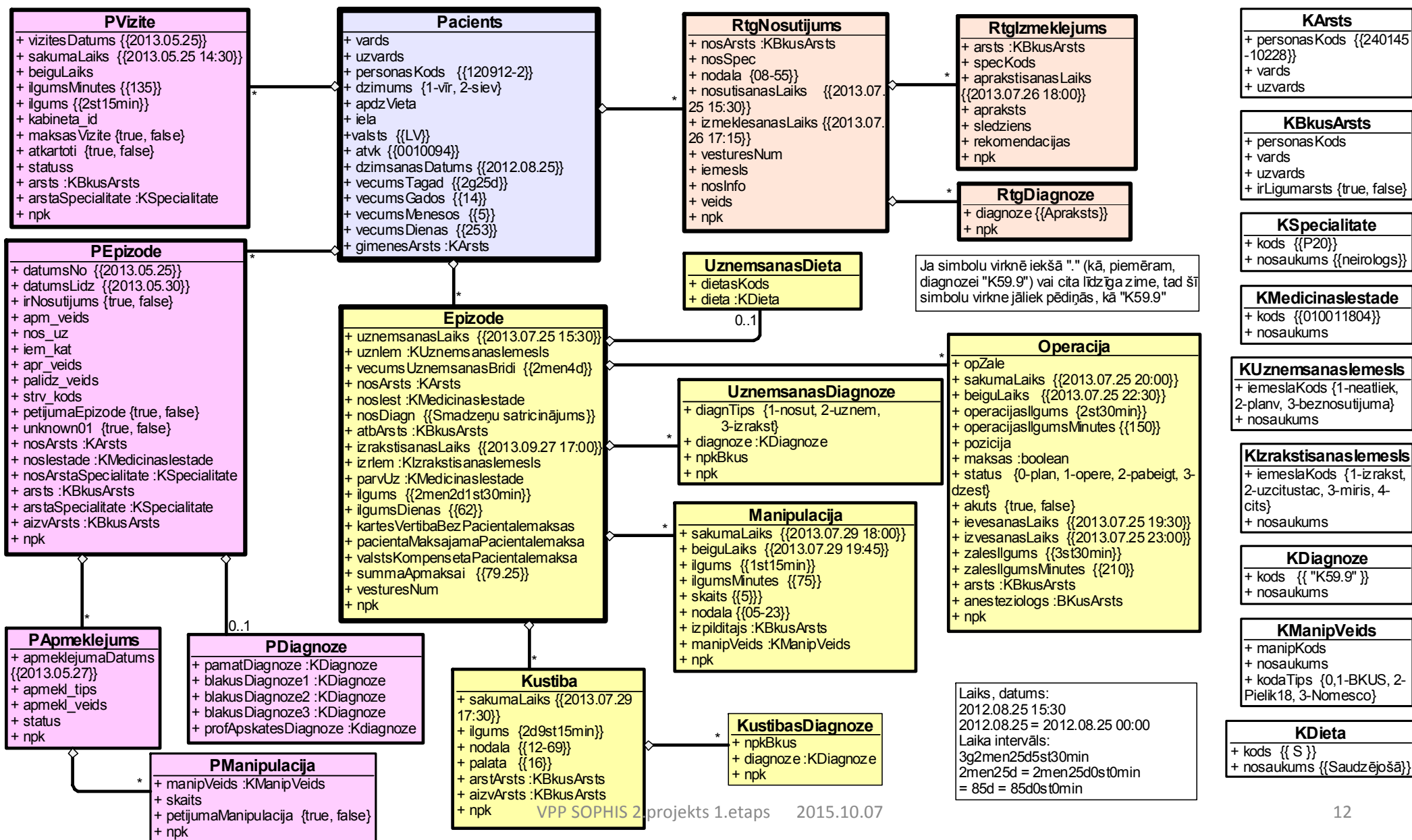


Fig. 5. Graphical in-place query diagram based on the star data ontology.

[illegible]



Pirmais svarīgākais rezultāts: Esam atraduši 6 kontrolētās dabīgās valodas vaicājumu šablonus (papildinātus ar skalārās izteiksmes formālo jēdzienu), kas pārklāj praktiski visus iedomājamus ad-hoc vaicājumus slimnīcas vadības vajadzībām (pie nosacījuma, ka vadītāji prot lietot Excel)). Šo hipotēzi esam pārbaudījuši uz BKUS reāliem datiem (par 2014.gadu) un reāliem vaicājumiem, kas bija nepieciešami vienas BKUS klīnikas (Intensīvās terapijas klīnikas) 2014.gada darbības pārskata un analīzes veikšanai. Minētais eksperiments ir apstiprinājis šo hipotēzi – tika sasniegts 100% atskaitei nepieciešamo vaicājumu-atbilžu pārklājums. Tiesa, eksperimenta laikā valoda vēl tika pilnveidota, acīmredzot, šādi reāli eksperimenti vēl ir jāturpina projekta 2.etapa laikā, kad tiks veikta vaicājumu valodas praktiska realizācija tīmekļa vidē.

Otrais svarīgākais rezultāts: Eksperimentālajā realizācijā uz BKUS reālajiem datiem (par 2014.gadu) praktiski uz visiem vaicājumiem tika sasniegta kritiskā ātrdarbība – iegūta atbilde aptuveni vienas sekundes laikā, kas gala lietotājām neradīja ‘kaitinošu’ gaidīšanas sajūtu.

Secinājums: Ir visas sagatves, lai projekta 2.etapā novestu šo vaicājumu valodu un ‘vebisko’ realizāciju līdz «Programmas ietvaros izstrādāto jauno tehnoloģiju, metožu, prototipu vai pakalpojumu skaits, kas aprobētas uzņēmumos» stāvoklim (aprobācija paredzēta BKUS).

Būs demo!

6 vaicājumu šabloni, ar kuriem praktiski pietiek

(šie šabloni būtiski izmanto datu ontoloģijas zvaiznes formu)

[parādīt / pilnparādīt] [10 / visas] Epizodes [e] kurām <no e -atkar log izt>

skaits Epizodēm [e] kurām <no e atkar log izt>

[max/min/biež/vid/sum] /<e-atr vai e-skalara izt>/ Epizodes [e] kurām <no e -atkar log izt>

atrast [10 / visas] Epizodes [x] kurām <no x -atkar log izt> ,
izveidot tabulu <x-atkar izt1> (kolonna A), <x-atkar izt2> (kolonna B), <x-atkar izt3> (kolonna C),
Var mainīt secību
→ atlasīt rindiņas kurām C<5,
→ sakārtot [augoši/dilstoši] pēc kolonnas B,
→ atstāt [pirmās/pēdējās] 5 rindiņas

atrast Epizodēm [e] , kurām <no e atkar log izt> , atribūta <e-atr vai izt> [10 / visas] atšķirīgās vērtības x ,
izveidot tabulu <x-atkar izt1> (kolonna A), <x-atkar izt2> (kolonna B), <x-atkar izt3> (kolonna C),
.....

atrast intervāla (8–23) [10 / visas] vērtības x ,
izveidot tabulu <x-atkar izt1> (kolonna A), <x-atkar izt2> (kolonna B), <x-atkar izt3> (kolonna C),
.....

*Ideja no
daudzsortīgās
predikātu valodas
(relāc. algebr. vietā)*

Vaicājuma piemērs

- Mirušo pacientu skaits nodaļā 12-69

- *dabīgā valodā*

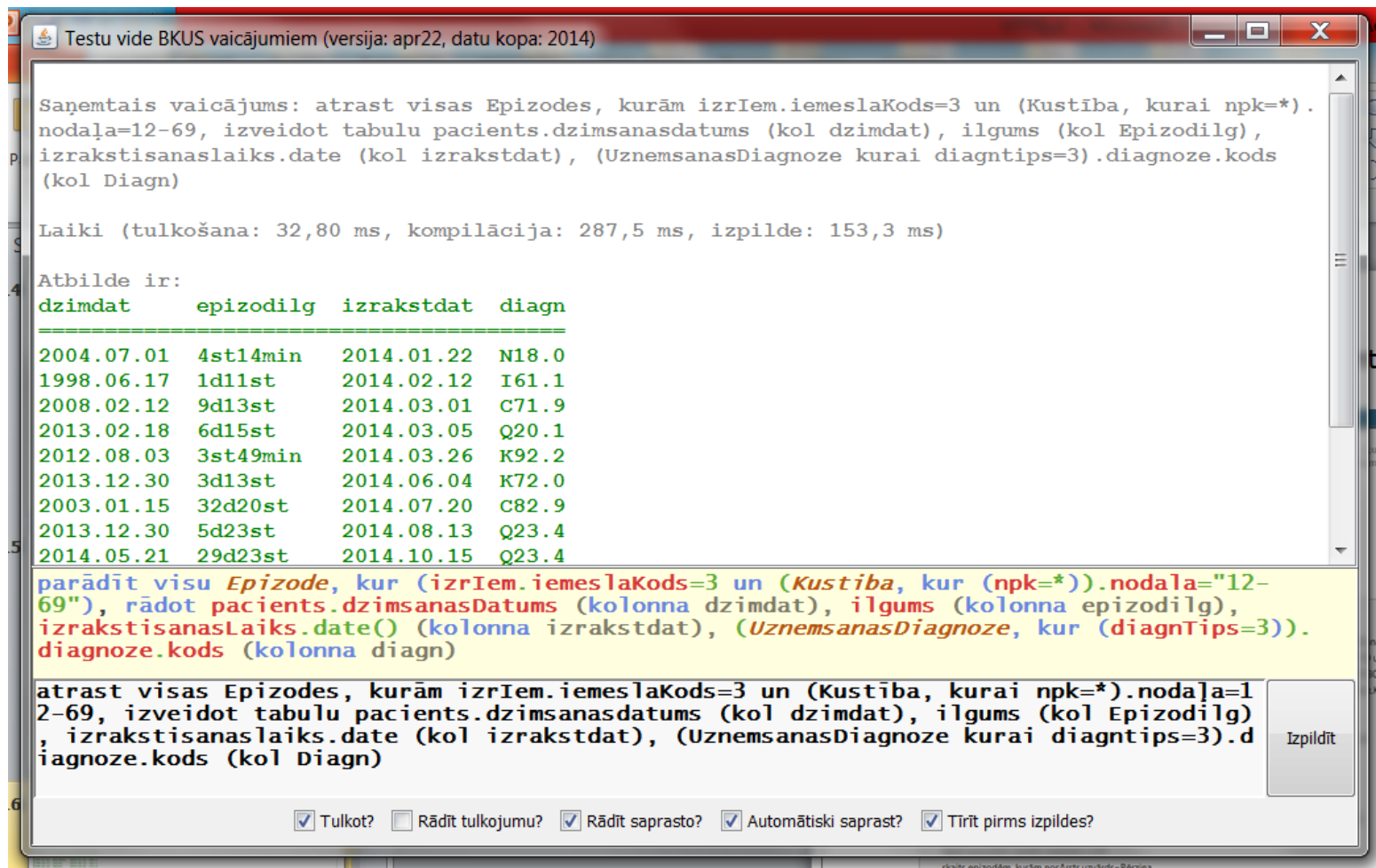
- `SELECT COUNT(DISTINCT pid) FROM uznemsanas_informacija, uznemsanas_kustiba WHERE izriem=3 AND nodaļa = '12-69' AND uznemsanas_informacija.bsid = uznemsanas_kustiba.bsid AND parvdat = ( SELECT MAX (uk2.parvdat) FROM uznemsanas_kustiba AS uk2 WHERE uk2.bsid = uznemsanas_informacija.bsid ))`

- *SQL*

- skaits *pacientiem*, kuriem eksistē *kustība*, kurai *nodaļa=12-69* un *npk=** un *epizode.izrlem.iemeslaKods=3*

- *mūsu vaicājumu valodā*

Vaicājums ar VV rīka prototipu uz BKUS reāliem datiem



Apmācība ar piemēriem, izmantojot VV rīka prototipu

Testu vide BKUS vaicājumiem (versija: apr22, datu kopa: 2014)

Saņemtais vaicājums: atrast visas Epizodes, kurām izIem.imeslaKods=3 un (Kustiba, kurai npk=*) .nodala=12-69, izveidot tabulu patients.dzimsanasdatums (kol dzimdat), ilgums (kol Epizodilg), izrakstisanaslaiks.date (kol izrakstdat), (UznemsanasDiagnoze kurai diagntips=3).diagnoze.kods (kol Diagn)

Laiki (tulkošana: 32,80 ms, kompilācija: 287,5 ms, izpilde: 153,3 ms)

Athilde ir:

dzimdat	epizodilg	izrakstdat	diagn
2004.07.01	4st14min	2014.01.22	N18.0
1998.06.17	1d11st	2014.02.12	I61.1
2008.02.12	9d13st	2014.03.01	C71.9
2013.02.18	6d15st	2014.03.05	Q20.1
2012.08.03	3st		
2013.12.30	3d1		
2003.01.15	32d		
2013.12.30	5d2		
2014.05.21	29d		
1800.01.01	10s		
2014.03.22	5st		
2011.01.13	1d1		

parādīt 2 Epizodes
pilnparādīt 2 Epizodes
parādīt 2 epizodes, kurām uzņemšanaslaiks.date>=2014.06.25 un uzņemšanaslaiks.date<2014.12.30 un nosārsts.uzvārds=Bērziņ...
atrast visas epizodes, kurām uzņemšanaslaiks.date>=2014.06.25 un uzņemšanaslaiks.date<2014.07.30 un nosārsts.uzvārds=Bēr...
atrast visas epizodes x, kurām izIem.imeslaKods<>3, izveidot tabulu x.vesturesNum (kol A), (skaits x.Kustibas) (kol B)...
atrast visas Epizodes x, kurām izIem.imeslaKods<>3, izveidot tabulu x.vesturesNum (kol A), (skaits x.Kustibas) (kol B)...
atrast visus KBkusārstus, kuriem personasKods.substring(1,4)=3001
atrast visus KBkusārstus, kuriem personasKods.substring(1,4)=3001, izveidot tabulu vārds, uzvārds
atrast visus KBkusārstus, kuriem vārds=Arta, izveidot tabulu vārds, uzvārds, personaskods.substring(1,4)
skaits kustības, kurām nodala=12-69
skaits Epizodes, kurām eksistē Kustiba, kurai nodala=12-69
skaits Pacienti, kuriem eksistē Kustiba, kurai nodala=12-69 un npk=1
skaits Pacienti, kuriem eksistē Kustiba, kurai nodala=12-69 un npk=*
skaits Pacienti kuriem eksistē Kustiba kurai (nodala=12-69 vai nodala=02-25) un (npk=1 vai npk=2)
skaits epizodēm, kurām nosArsts=nil
skaits epizodēm, kurām nosArsts<>nil
skaits epizodēm, kurām nosArsts.uzvārds=Bērziņa
skaits epizodēm, kurām (nosArsts.uzvārds=Bērziņa vai nosArsts.uzvārds=Kalniņa) un uznemsanasLaiks.date>2014.06.25
- mirušo un izdzīvojušo pacientu skaits par 2014.g.:
skaits Pacienti, kuriem eksistē Epizode, kurai izIem.imeslaKods=3
skaits Pacienti, kuriem neeksistē Epizode, kurai izIem.imeslaKods=3
- mirušo pacientu skaits nodaļā 12-69 (dažādi pieraksti):
skaits Pacienti, kuriem eksistē Epizode, kurai izIem.imeslaKods=3 un eksistē Kustiba, kurai nodala=12-69 un npk=*
skaits Pacienti, kuriem eksistē Kustiba, kurai epizode.izIem.imeslaKods=3 un nodala=12-69 un npk=*

parādīt visu / izrakstisanaslaiks

Piemēri
Palīgs
Atribūts
Notīrīt
Izgriezt
Kopēt
Iespraust
Iespraust vietā
Atzīmēt visu

patients.dzimsanasDatums (kolonna dzimdat), ilgums (kolonna epizodilg), kods (kolonna diagn)
t tabulu patients.dzimsanasdatums (kol dzimdat), ilgums (kol Epizodilg (kol Diagn)

Izpildīt

☒ Tulkot? ☐ Rādīt tulkojumu? ☒ Rādīt saprasto? ☒ Automātiski saprast? ☒ Tīrīt pirms izpildes?

LV 98% 16:33 2015.07.07.

II. Grūti formalizējamu sistēmu tīmeklī balstītu modelēšanas rīku būves platforma (strādājošs prototips) (uzdevums 3)

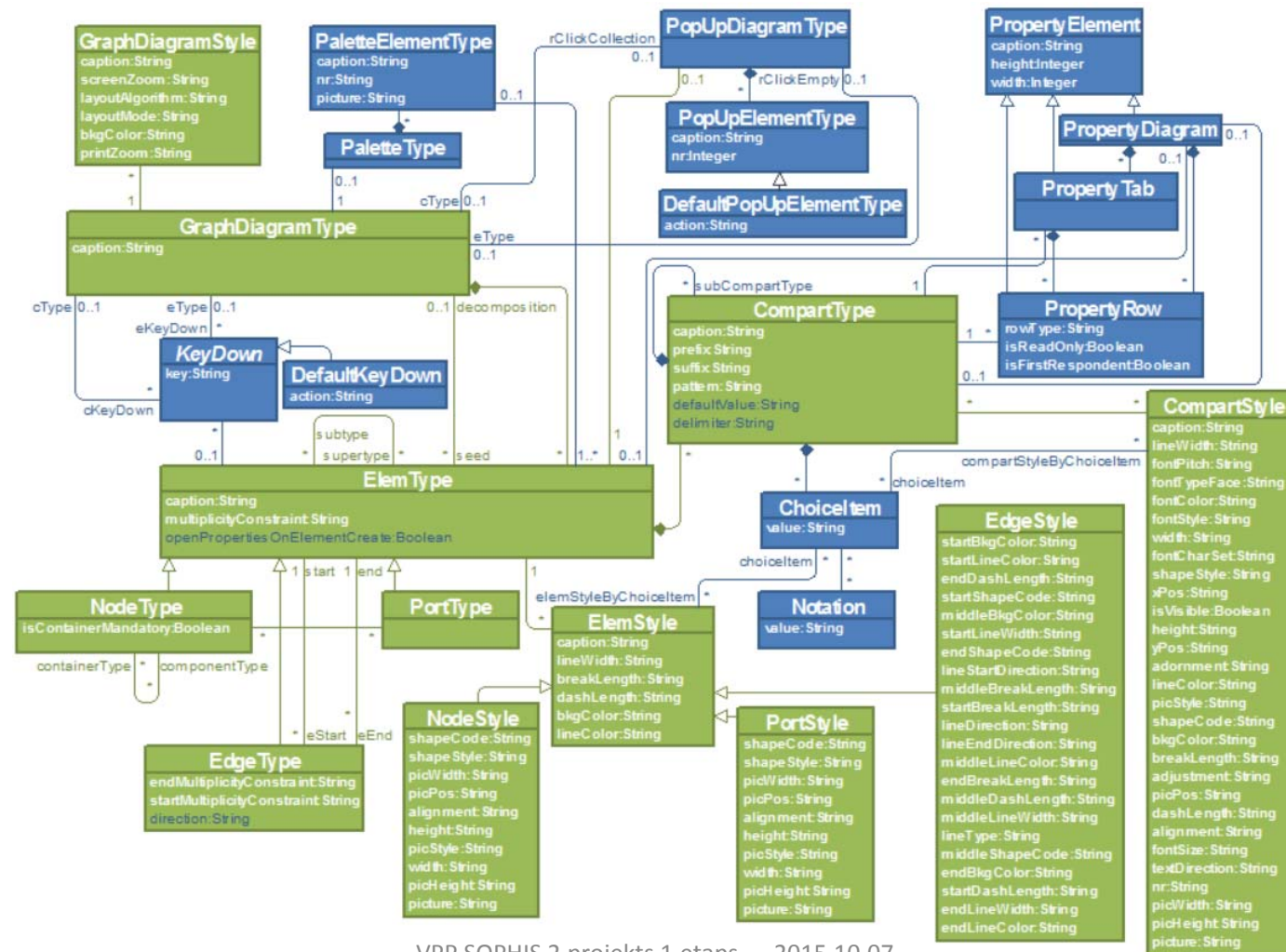
Galvenā piedāvātās tehnoloģijas ideja – Domēnspecifiskās Modelēšanas Valodas (DSL) un to atbalsta rīki.

DSL ideja nav jauna. Taču, salīdzinot ar tradicionālajām DSL valodām un rīkiem, grūti formalizējamu sistēmu gadījumā jānodrošina lietotājam ērtā veidā vairākas lietas:

- DSL valodas dināmisko (t.i. modelēšanas laika) izmaiņu atbalsts , pārkonfigurācijas iespējas modelēšanas laikā (valodas iterācija kā DSL rīka sastāvdaļa),
- DSL valodas formālismu un dabīgās valodas aprakstu ērti sasaistes līdzekļi (jo grūti formalizējamās sistēmas tieši ar to raksturojas, ka formāli var aprakstīt tikai kādu daļu no šīm sistēmām, bet pārējais jāatstāj dabīgās valodasa līmenī, bet abiem šiem aprakstu līmeņiem attiecīgajā DSL rīkā ir jābūt lietotājam ērtā veidā sasaistītiem),
- Jānodrošina ērts hipersaišu mehānisms ar aprakstu resursiem internetā.

Atbilstoši šīm prasībām, izstrādāts DSL valodu un rīku definēšanas metamodelis

DSML rīku definēšana metamodelis



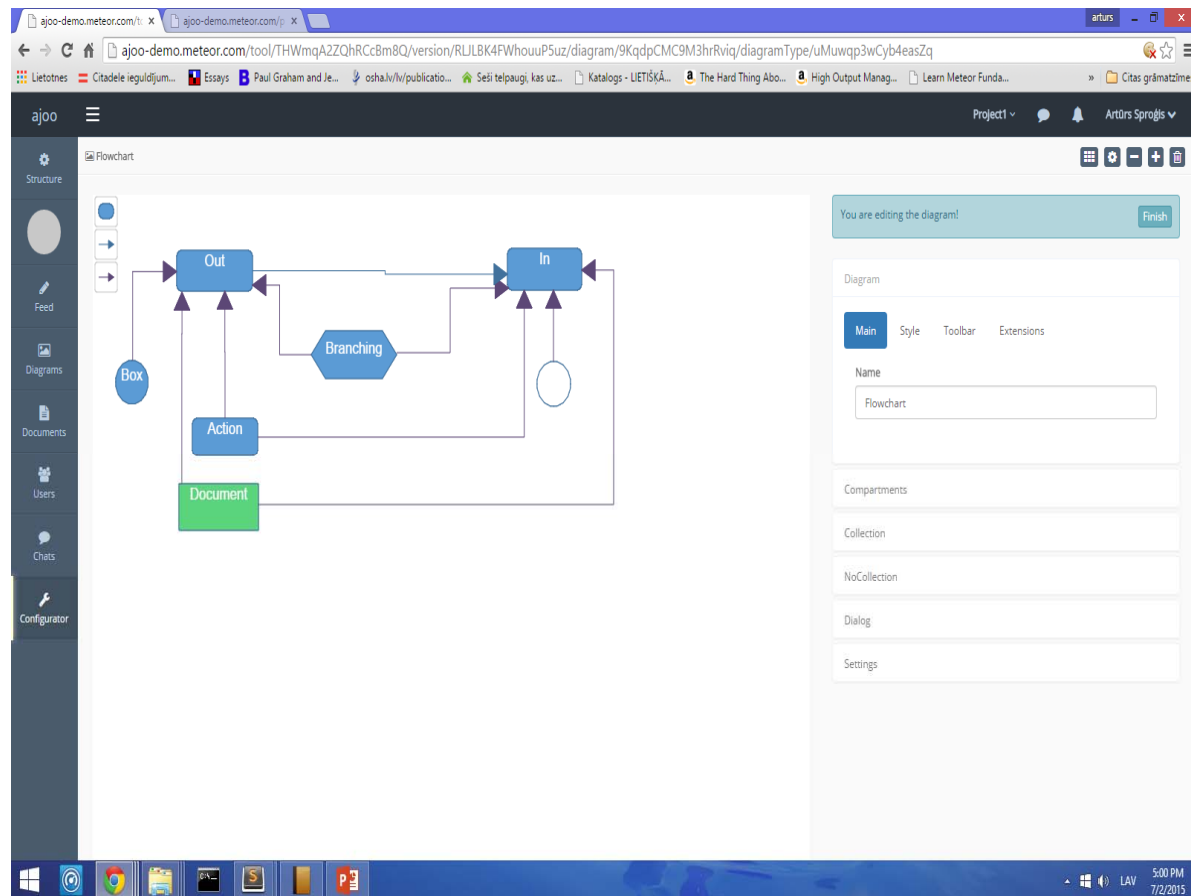
Platformas būvē Izmantotās tehnoloģijas

- Meteor (web aplikāciju izstrādes platforma) un MongoDB (datubāze)
- KonvaJS (diagrammu zīmēšana)
- Bootstrap (HTML, CSS, JavaScript)

Grafisku modelēšanas rīku būves platforma tīmeklī

- Interaktivitāte
- Daudzlietotāju režīms (collaboration)
- Dažādas iekārtas – datori, planšetes, telefoni...
- Reālā laikā (reactivity and live HTML)
- Viegli pieejams un ieviešams

Konfigurācija



Rīka piemērs I

ajoo-demo.meteor.com/project/NT94suKd82aMj8xe4/diagram/TK9g3i4Xwvdk5HZmu/type/Rk2imgnw82DeDwfHq/version/3LfjHaNmXck32Qh7N

Lietotnes Citadele ieguldījums... Essays Paul Graham and Je... osha.lv/publicatio... Seši telpaugi, kas uz... Katalogs - LIETIŠKĀ... The Hard Thing Abo... High Output Manag... Learn Meteor Funda... Citas grāmatzīmes

ajoo Project1 Artūrs Sprūģis

Diagram1

```
graph TD; Start(( )) --> A["(SP direktors)  
Sagatavo priekšlikumu par noslēguma  
pārbaudījumu komisijas sastāvu"]; A --> B["Priekšlikums"]; B --> C["(Fakultātes dome)  
Apstiprina priekšlikumu"]; C --> D["Domes lēmums par apstiprināšanu"]; D --> E["(Domes sekretārs)  
Sagatavo rīkojuma projektu par  
Komisijas sastāvu"]; E --> F["Rīkojuma projekts"]; F --> G["(Domes sekretārs)  
Rīkojuma projektu iesniedz LU  
akadēmiskajam sekretāram"]; G --> H["(Mācību prorektors)  
Apstiprina rīkojuma projektu"]; H --> I["Apstiprināts komisijas sastāvs"]; I --> End((( )));
```

You can edit the diagram! Edit

Documents

3.3. Priekšlikumu par noslēguma pārbaudījumu komisijas (turpmāk – Komisija) sastāvu sagatavo studiju programmas direktors vai akadēmiskās struktūrvienības (katedras, nodaļas) vadītājs un iesniedz apstiprināšanai fakultātes domei. Pamatojoties uz domes lēmumu, domes sekretārs sagatavo rīkojuma projektu par Komisijas sastāvu un ne vēlāk kā četras nedēļas pirms Komisijas pirmās sēdes iesniedz LU akadēmiskajam sekretāram. Komisijas sastāvu katram akadēmiskajam gadam ar LU rīkojumu apstiprina mācību prorektors.

Nolikums1

3.5. Komisiju izveido katrā studiju programmā. Nepieciešamības gadījumā vienā studiju programmā var veidot vairākas komisijas.

Nolikums1

3.6. Gala pārbaudījumu komisijas sastāvu veido studiju programmas akadēmiskā personāla locekļi. Komisijas sastāvā var iekļaut ar studiju programmas izpildi iesaistītus ekspertus.

Nolikums1

3.7. Gala pārbaudījumu komisijas priekšsēdētājs un vietnieks ir personas ar doktora grādu. Bakalaura gala pārbaudījumu komisijas locekļiem jābūt akadēmiskā personāla pārstāvjiem, kuriem ir vismaz maģistra grāds. Maģistra gala pārbaudījumu komisijas locekļiem, izņemot sekretāru, jābūt doktora zinātniskajam grādam vai docenta akadēmiskajam nosaukumam.

Nolikums1

3.8. Gala pārbaudījumu komisijas sekretārs ir akadēmiskā personāla pārstāvis, kuram ir vismaz maģistra grāds. Komisijas sekretāram ir balss

VPP SOPHIS 2.projekts 1.etaps 2015.10.07

Windows Taskbar: 5:23 PM 7/2/2015

Rīka piemērs II

ajoo-demo.meteor.com/p x ajoo-demo.meteor.com/p x arturs

ajoo-demo.meteor.com/project/NT94suKd82aMjBxe4/diagram/TK9g3i4Xwvdk5HZmu/type/Rk2imgw82DeDwfHq/version/3LfjHaNmXck32Qh7N

Lietotnes Citadele ieguldījums Essays Paul Graham and Je... osha.lv/lv/publicatio... Sesī telpaugi, kas uz... Katalogs - LIETIŠKĀ... The Hard Thing Ab... High Output Manag... Learn Meteor Funda... Citas grāmatzīmes

ajoo Project1 Artūrs Sproģis

Diagram1

Structure

Feed

Diagrams

Documents

Users

Chats

Configurator

You can edit the diagram! Edit

Documents

prof Juris Borzovs Datorikas fakultate 2009.03.18

07:12 Zvanīt tālrunī 67200047

Video

Nolikums1

Main

VPP SOPHIS 2.projekts 1.etaps 2015.10.07

5:32 PM 7/2/2015

```
graph TD; Start(( )) --> A["(SP direktors)  
Sagatavo priekšlikumu par noslēguma  
pārbaudījumu komisijas sastāvu"]; A --> B["Priekšlikums"]; B --> C["(Fakultātes dome)  
Apstiprina priekšlikumu"]; C --> D["Domes lēmums par apstiprināšanu"]; D --> E["(Domes sekretārs)  
Sagatavo rīkojuma projektu par  
Komisijas sastāvu"]; E --> F["Rīkojuma projekts"]; F --> G["(Domes sekretārs)  
Rīkojuma projektu iesniedz LU  
akadēmiskajam sekretāram"]; G --> H["(Mācību prorektors)  
Apstiprina rīkojumu projektu"]; H --> I["Apstiprināts komisijas sastāvs"]; I --> End((( )))
```


III. Uz lieliem semantiskiem grafiem (piemēram, BabelNet) un notikumu n-āru relāciju grafiem (piemēram, AMR, FrameNet) balstītu dabīgās valodas saprašanas (language understanding) metožu teorētisko pamatu izstrāde (uzdevums 5)

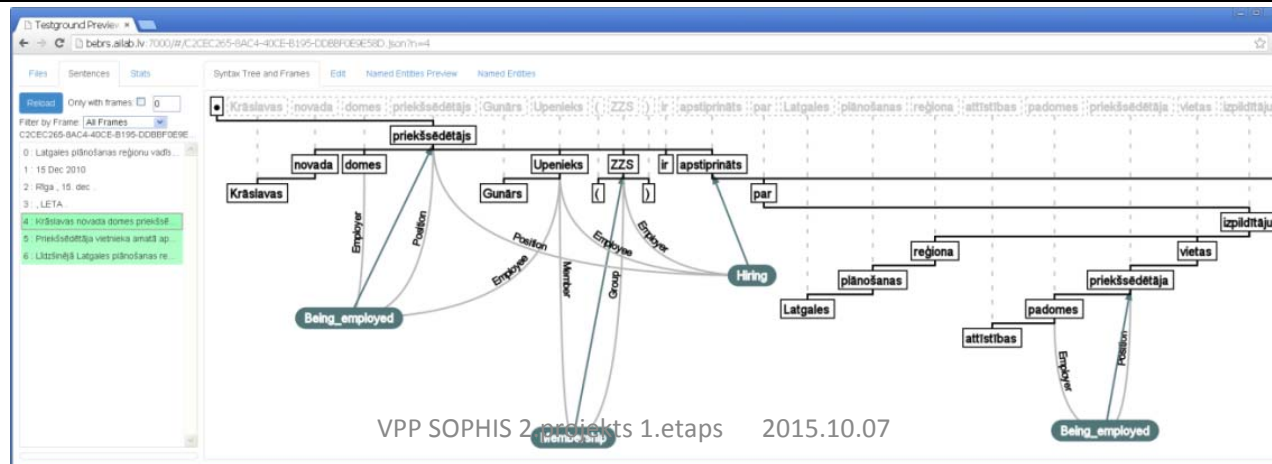
Apvienojot FrameNet n-āro relāciju un Named Entity Linking pieeju vienotā lingvistiskā ontoloģijā izstrādāta tehnoloģija Curriculum Vitae tipa datu (semantisko grafu) izgūšanai no dabīgās valodas tekstiem. Šī pētījuma teorētiskie rezultāti atspoguļoti publikācijā [1] un praktiski aprobēti ERAF Nr.2DP/2.1.1.1.0/13/APIA/VIAA/014 projektā "Pētījums par publicistikā pieminēto entītiju savstarpējo saišu identificēšanu, tām atbilstošo grafu strukturēšanu un datu bāzu vaicājumu attēlošanu grafu veidā", pasūtītājs SIA „LETA” (kopējās projekta izmaksas 395KEUR). Tāpat uz šī pētījuma pamata kopā ar SIA „LETA” sagatavots un konkursā vinnēts Eiropas Horizon-2020 ICT-16-2015 „Big data – research” projekts SUMMA (Nr. 688139) par kopējo summu 9.85MEUR, no kuriem Latvija saņems 1.16MEUR. Pētniecībā bija iesaistīti LU Datorikas fakultātes doktoranti Didzis Goško un Pēteris Paikens, kā arī par darba tēmu sekmīgi aizstāvēti magistra darbi [5m,6m].

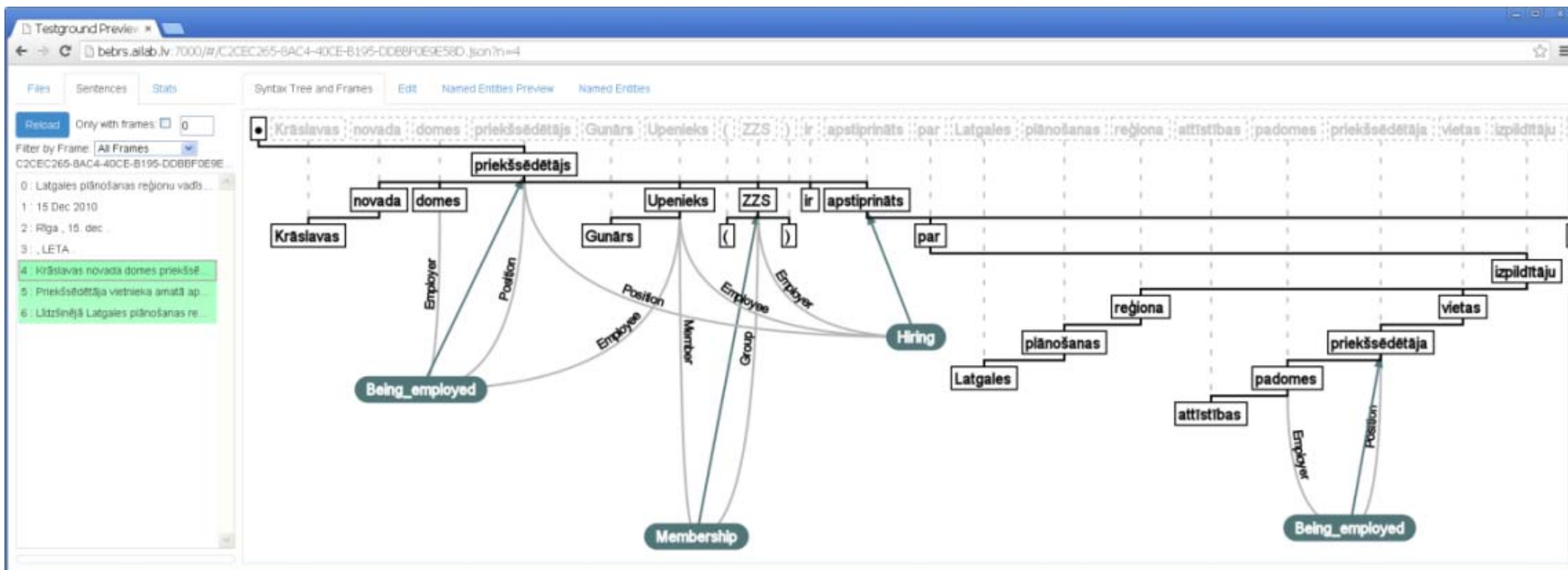
Language understanding: information extraction

Notikumi

Relation	Attributes						
Being_born	Child	Time	Place	Relatives			
People_by_origin	Origin	Person	Ethnicity				
Education_teaching	Student	Institution	Subject	Qualification	Time	Place	
People_by_vocation	Person	Vocation	Time	Descriptor			
Being_employed	Employee	Employer	Position	Place	Time	Employment_start	Employment_end
Personal_relationship	Partner_1	Partner_2	Partners	Relationship	Time		
Participation	Participant_1	Event	Time	Place	Institution	Manner	
Membership	Member	Group	Time	Standing			
Statement	Speaker	Message	Medium	Time			
Intentionally_create	Created_entity	Creator	Manner	Industry	Time	Place	
Earnings_and_losses	Earners	Earnings	Goods	Profit	Unit	Time	Growth
Change_of_leadership	Candidate	Body	Role	New_leader	Result	Time	Place
Attack	Victim	Assailant	Result	Circumstances	Weapon	Place	Time
Death	Protagonist	Time	Place	Cause	Manner		
...							

No FrameNet
(tur ~1000,
CV -26)

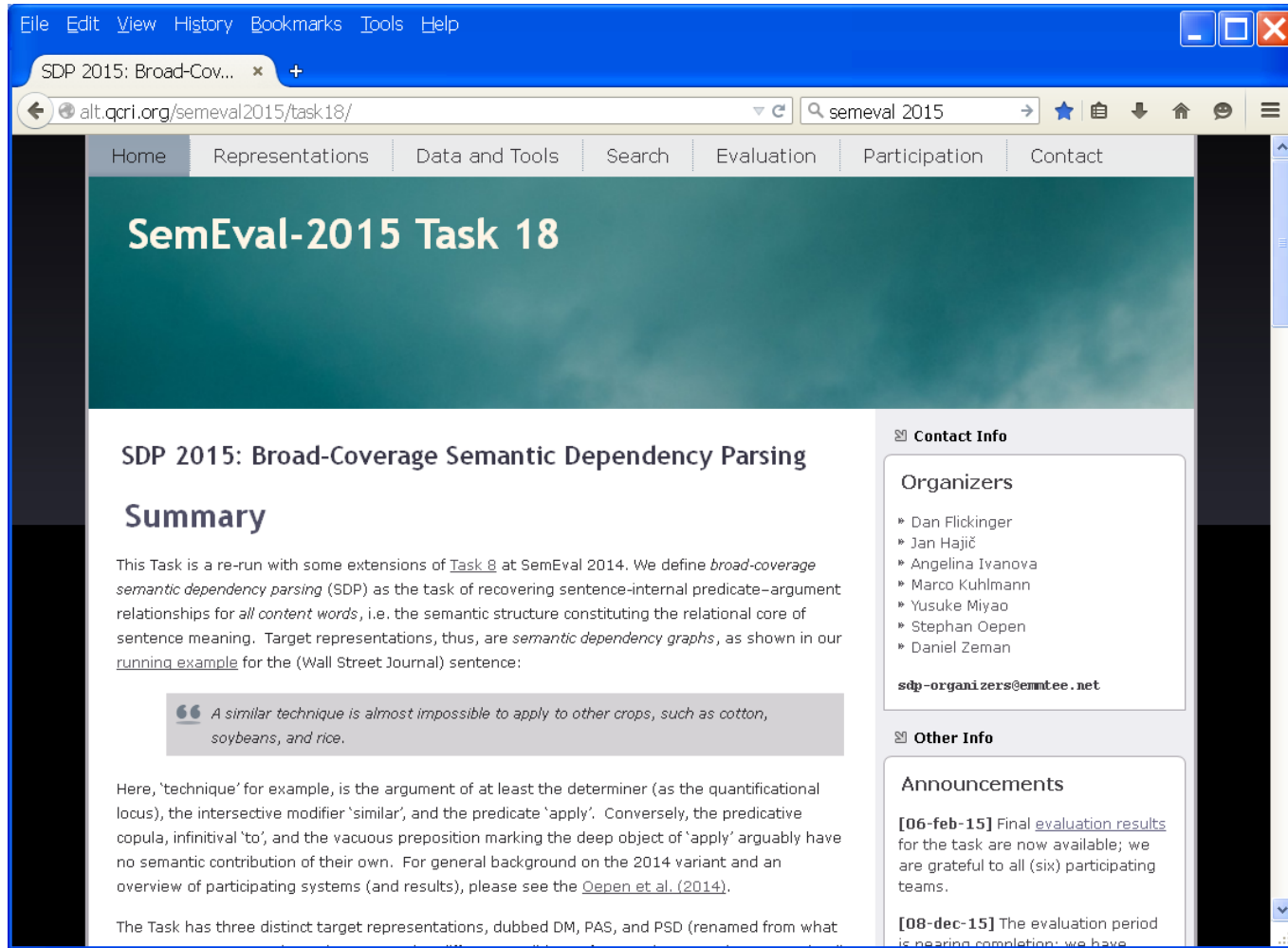




Ar šo bildi dabīgās valodas konkrētu tekstu ieliek FrameNet
= 'valodas saprašana'

Šis uzdevums jau sen tiek risināts, tikai kā to labāk izdarīt? Īsti labi nevienam neizdodas.
Šī projekta ietvaros izstrādāts t.s. C6.0 algoritms. Ar ko labāks par esošajiem?

SemEval-2015 Task 18



The screenshot shows a web browser window with the address bar displaying 'alt.qcri.org/semEval2015/task18/'. The browser's menu bar includes 'File', 'Edit', 'View', 'History', 'Bookmarks', 'Tools', and 'Help'. The page features a navigation bar with links to 'Home', 'Representations', 'Data and Tools', 'Search', 'Evaluation', 'Participation', and 'Contact'. The main heading is 'SemEval-2015 Task 18'. Below this, the section is titled 'SDP 2015: Broad-Coverage Semantic Dependency Parsing' with a 'Summary' sub-header. The summary text describes the task as a re-run of Task 8 from SemEval 2014, defining broad-coverage semantic dependency parsing (SDP) as recovering sentence-internal predicate-argument relationships for all content words. A quote from a Wall Street Journal sentence is provided: 'A similar technique is almost impossible to apply to other crops, such as cotton, soybeans, and rice.' The text explains that 'technique' is the argument of 'apply', 'similar' is an intersective modifier, and 'apply' is the predicate. It also mentions the availability of general background on the 2014 variant and participating systems. A right sidebar contains 'Contact Info' with organizers' names and email, and 'Other Info' with announcements regarding final evaluation results and the evaluation period.

File Edit View History Bookmarks Tools Help

SDP 2015: Broad-Cov... x +

alt.qcri.org/semEval2015/task18/

semEval 2015

Home Representations Data and Tools Search Evaluation Participation Contact

SemEval-2015 Task 18

SDP 2015: Broad-Coverage Semantic Dependency Parsing

Summary

This Task is a re-run with some extensions of [Task 8](#) at SemEval 2014. We define *broad-coverage semantic dependency parsing* (SDP) as the task of recovering sentence-internal predicate-argument relationships for *all content words*, i.e. the semantic structure constituting the relational core of sentence meaning. Target representations, thus, are *semantic dependency graphs*, as shown in our [running example](#) for the (Wall Street Journal) sentence:

“ A similar technique is almost impossible to apply to other crops, such as cotton, soybeans, and rice.

Here, ‘technique’ for example, is the argument of at least the determiner (as the quantificational locus), the intersective modifier ‘similar’, and the predicate ‘apply’. Conversely, the predicative copula, infinitival ‘to’, and the vacuous preposition marking the deep object of ‘apply’ arguably have no semantic contribution of their own. For general background on the 2014 variant and an overview of participating systems (and results), please see the [Oepen et al. \(2014\)](#).

The Task has three distinct target representations, dubbed DM, PAS, and PSD (renamed from what

Contact Info

Organizers

- Dan Flickinger
- Jan Hajič
- Angelina Ivanova
- Marco Kuhlmann
- Yusuke Miyao
- Stephan Oepen
- Daniel Zeman

sdp-organizers@emntee.net

Other Info

Announcements

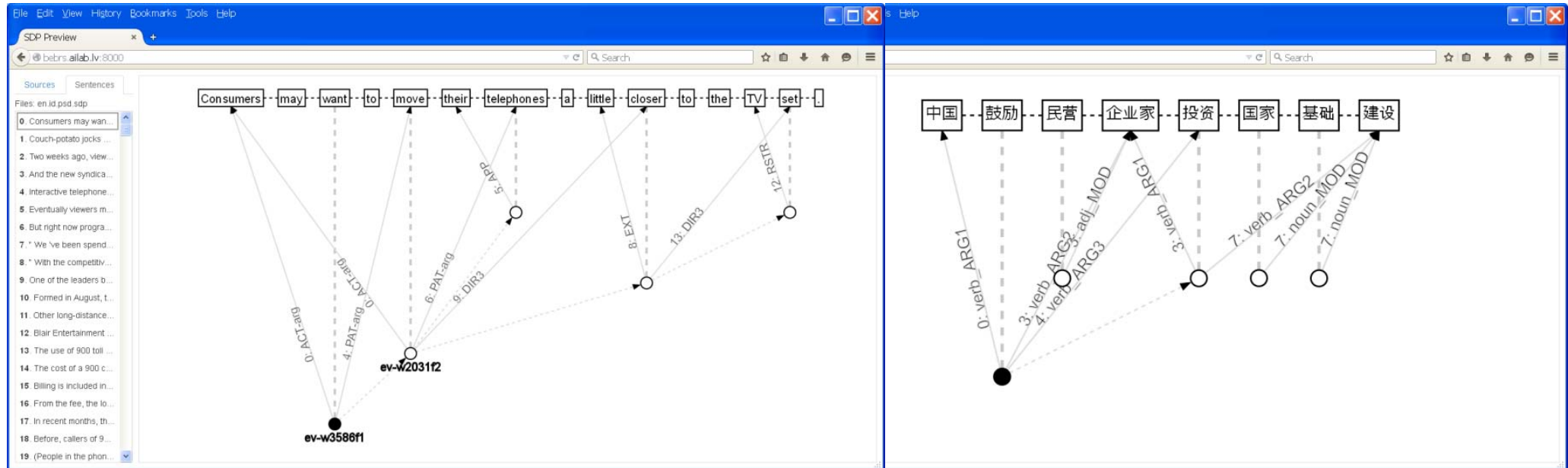
[06-feb-15] Final [evaluation results](#) for the task are now available; we are grateful to all (six) participating teams.

[08-dec-15] The evaluation period is nearing completion; we have

Treninkorpusi

EN.PSD

CZ.PAS



System	LF	LM	PF	SF	FF
Peking	80.51	21.14	62.64	69.45	48.70
Lisbon	80.42	20.05	63.59	--	--
Riga	78.68	21.84	61.29	73.76	48.33
Minsk	78.18	15.04	56.40	79.40	47.32

Ranking of scores averaged over all available datasets for the best runs of the systems in the closed track:

- labeled dependencies (LF),
- labeled exact match of the complete semantic dependency graphs (LM),
- complete predications (PF),
- sense identification (SF),
- semantic-frames (FF).

Riga: from FrameNet to Semantic Frames with C6.0 Rules

Guntis Barzdins, Peteris Paikens, Didzis Gosko

University of Latvia, IMCS

Rainis Blvd. 29, Riga, LV-1459, Latvia

{guntis.barzdins,peteris.paikens,didzis.gosko}@lumii.lv

Abstract

For the purposes of SemEval-2015 Task-18 on the semantic dependency parsing we combined the best-performing closed track approach from the SemEval-2014 competition with state-of-the-art techniques for FrameNet semantic parsing. In the closed track our system ranked third for the semantic graph accuracy and first for exact labeled match of complete semantic graphs. These results can be attributed to the high accuracy of the C6.0 rule-based sense labeler adapted from the FrameNet parser. To handle large SemEval training data the C6.0 algorithm was extended to provide multi-class classification and to use fast greedy search without significant accuracy loss compared to exhaustive search. A method for improved FrameNet parsing using semantic graphs is proposed.

1 Introduction

The trend of natural language processing in recent years is shifting towards multilingual natural language understanding based on full-text shallow semantic parsing (e.g., Banarescu et al., 2013). Despite various formalisms proposed, these approaches are characterized by direct extraction of a bi-lexical semantic graph rather than a bi-lexical dependency tree from the surface form of the sentence.

Following the best practice for semantic parsing established already by the SemEval-2014 Task 8 (Open et al., 2014) we modified the best-performing closed-track system there (Du et al., 2014) by removing some less essential components

while adding a new component of our own. The newly added component is the C6.0 rule-based classifier (Barzdins et al., 2014) used both for graph parsing and for sense labeling. Sense labeling is a novelty of SemEval-2015 Task 18 and was not present in the previous year competition. Semantic frame is comprised of a complete predication combined with the sense identifier of its predicate as shown in Figure 1. Semantic frames are similar to FrameNet (Fillmore et al., 2003) frames, except that FrameNet argument labels are sense-specific – this mismatch can be resolved by feeding the semantic graph (instead of dependency tree) through the regular FrameNet parser.

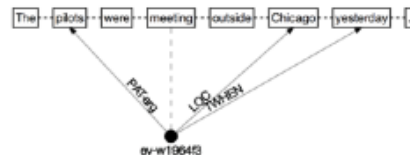


Figure 1. Semantic frame from the PSD corpus.

We participated only in the closed track. Despite ranking third for the semantic graph accuracy, our system ranked first for exact labeled match of complete semantic graphs, and close second for semantic frame accuracy.

2 Baseline Architecture

For semantic graph parsing we started by implementing a straight-forward baseline architecture described on the SemEval-2015 Task-18 evaluation page by the task organizers. The baseline architecture consists of two components: reduction



Guntis Barzdins

@guntisb

Mūsu daturs ķīniešu valodu (LM jomā)
saprof labāk kā Pekinas daturs! :) goo.gl/vUQ9ML goo.gl/trK4rF

View translation

Retweets 6 Favorites 4

Chinese: In-Domain (Closed Track)											
PAS											
#	Team	Run	LP	LR	LF	LM	UP	UR	UF	UM	PP
1	Peking	2	0.8475	0.8215	0.8341	0.2809	0.8638	0.8374	0.8504	0.2937	0.8703
	Peking	1	0.8452	0.8174	0.8310	0.2701	0.8618	0.8334	0.8473	0.2824	0.8596
2	Riga	1	0.8312	0.8184	0.8247	0.2884	0.8477	0.8346	0.8411	0.3017	0.8638
	Riga	2	0.8634	0.7837	0.8216	0.2438	0.8791	0.7979	0.8365	0.2532	0.8548
3	Lisbon	1	0.8381	0.8031	0.8202	0.2326	0.8556	0.8199	0.8374	0.2443	0.8688
4	Minsk	2	0.7927	0.7615	0.7768	0.1892	0.8170	0.7849	0.8007	0.2108	0.5917
	Minsk	1	0.7917	0.7612	0.7761	0.1894	0.8165	0.7850	0.8005	0.2118	0.5900

Chinese: In-Domain (Open Track)											
PAS											
#	Team	Run	LP	LR	LF	LM	UP	UR	UF	UM	PP
1	Turku	2	0.8081	0.7851	0.7964	0.1964	0.8255	0.8021	0.8136	0.1991	0.8264
	Turku	1	0.7990	0.7823	0.7906	0.1771	0.8192	0.8020	0.8105	0.1826	0.8191

RETWEETS 6

FAVORITES 4



3:40 AM - 9 Feb 2015

Pēteris Paikens un Didzis Goško (LU DF doktoranti)

Dace Dambergā, LU DF maģistra darbs «Uzraudzības mašīnmācīšanās klasifikatoru izpēte un empīriskā salīdzināšana», aizstāvēts 02/06/2015

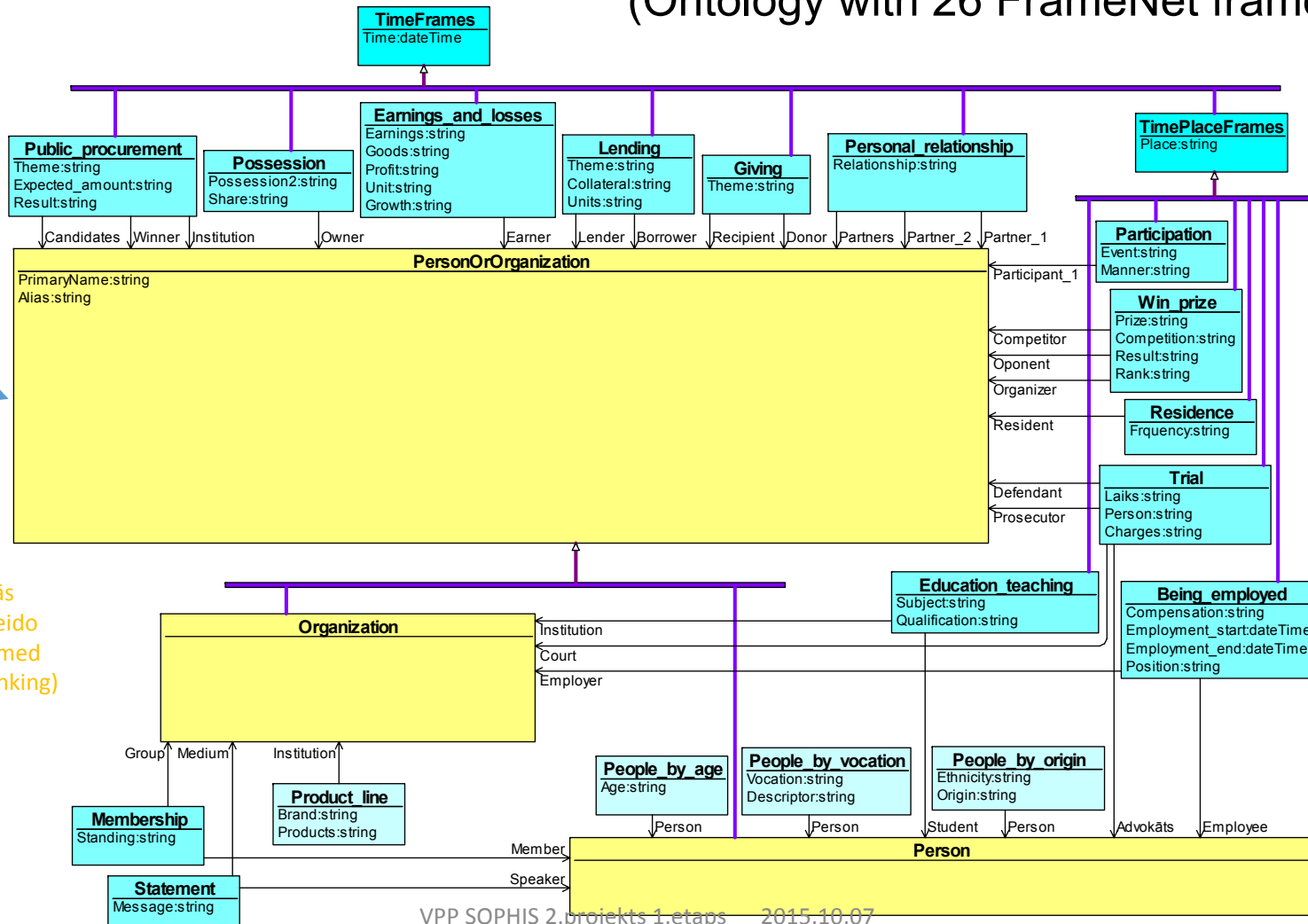
Abstract Knowledge Representation

(Ontology with 26 FrameNet frames)

2.sasniegums

Dzeltenās
klases veido
NEL (Named
Entity Linking)

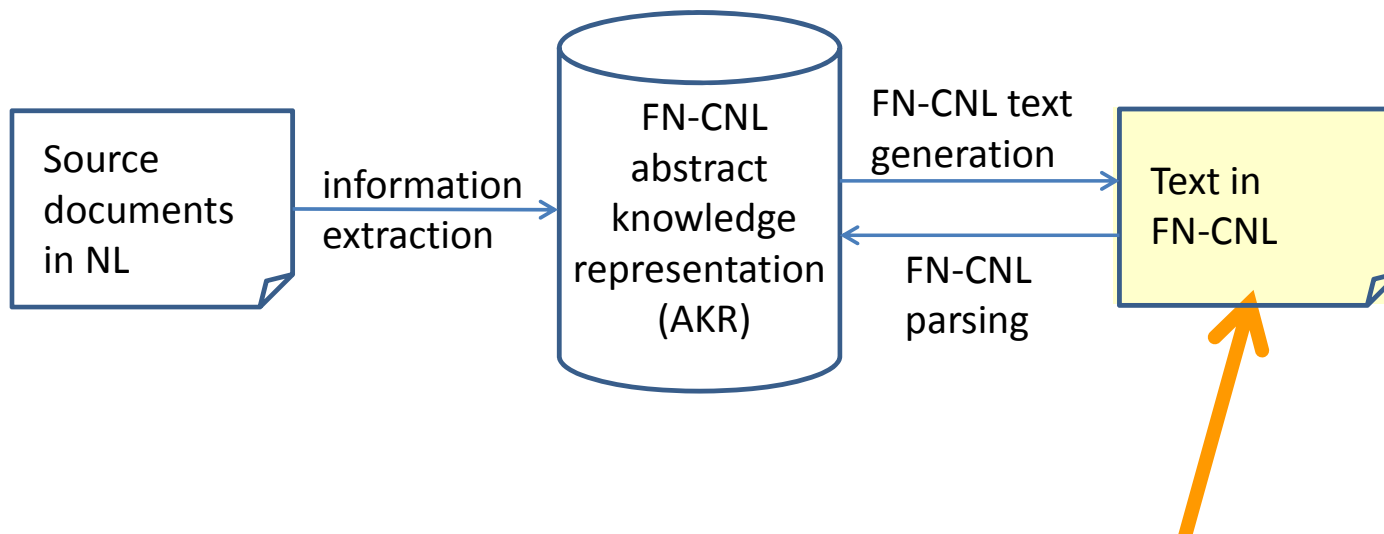
3.Sasniegums
«kurš Andris
Bērziņš tas ir»



VPP SOPHIS 2.projekts 1.etaps 2015.10.07

OWL ontology visualised with OWLGrEd: <http://owlgred.lumii.lv>

Valodas saprašanas sistēmas shēma



4.sasniegums

Ieva Akuratere bija solista amatā [23]
Ieva Akuratere bija Puķu burves amatā [8]
Ieva Akuratere bija mūziķes un aktrises amatā [5]
Ieva Akuratere bija deputātes amatā Rīgas domē [
Ieva Akuratere bija solista amatā Koncertuzvedumā [4]
Ieva Akuratere bija dziedātājas amatā [3]
Ieva Akuratere bija triju Zvaigžņu ordena virsnieka amatā Latvijā [3] (...held an Honorary position in Latvia)

Ieva Akuratere held a soloist position
 Ieva Akuratere held a Flower fairy position
 Ieva Akuratere held a musician and actress position
 Ieva Akuratere held a member position in Riga city council
 Ieva Akuratere held a soloist position in a Concert
 Ieva Akuratere held a singer position
 Ieva Akuratere held an Honorary position in Latvia



- 1.Sasniegums
- +
- 2.Sasniegums
- +
- 3.Sasniegums
- +
- 4.sasniegums

SUMMA

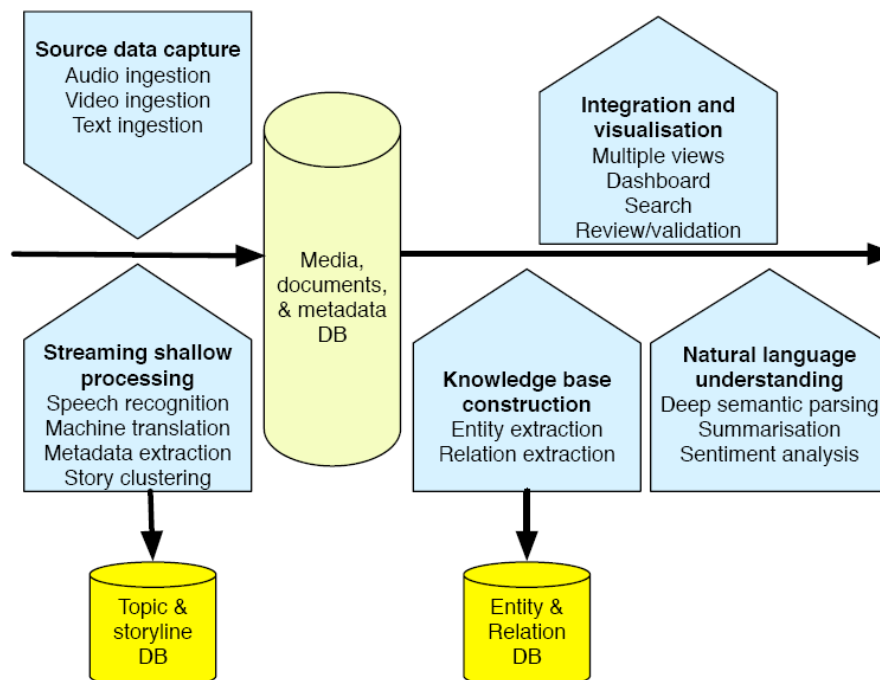
(Scalable Understanding of Multilingual Media)

9,85 MEUR; no tā Latvijai 1,16 MEUR

(lielākais Eiropas IKT pētniecības projekts Latvijā – pēc Horizon2020 Latvijas nacionālā kontaktpunkta datiem)

Participants

- THE UNIVERSITY OF EDINBURGH(999974941) - **COORDINATOR**
- BRITISH BROADCASTING CORPORATION(999589075) - **BENEFICIARY**
- DEUTSCHE WELLE(999586262) - **BENEFICIARY**
- UNIVERSITY COLLEGE LONDON(999975620) - **BENEFICIARY**
- LETA(949269258) - **BENEFICIARY**
- FONDATION DE L'INSTITUT DE RECHERCHE IDIAP(999644074) - **BENEFICIARY**
- Qatar Computing Research Institute(929080551) - **BENEFICIARY**
- PRIBERAM INFORMATICA S.A.(969256011) - **BENEFICIARY**



Project 688139 (SUMMA)

Responsible Unit:

CNECT/G/03

Call:

H2020-ICT-2015

Topic:

ICT-16-2015 - Big data - research

Type of Action:

RIA

Duration:

36 months (01/02/2016 – 31/01/2019)

Budget Information:

Total Costs in the Proposal:

9,858,326.25 €

<http://summa-project.eu>

Paldies par uzmanību!