

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

3. 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:

2.1. Uz ontoloģijām balstītas tīmekļa videi pielāgotas modelēšanas tehnoloģijas un rīki zināšanu analīzei un Uz ontoloģijām balstītas dabīgās valodas semantikas izgūšanas metodes (**LU MII** : J.Bārzdiņš, A.Šostaks, E.Rencis, M.Grasmanis, A.Kalniņš, A.Sproģis, G.Bārzdiņš, D.Goško, P.Paikens)

2.2. Semantiskā tīmekļa tehnoloģijas zināšanu formalizācijai, vairākkārtīgai izmantošanai un koplietošanai (**RTU DITF** : J.Grundspenķis, M.Kirikova, A.Novickis, E.Lavendelis)

2.3. Modeļu bāzēta datu vizualizācija un biznesa procesu izpildes laika verifikācija (**LU DF** : J.Bičevskis, G.Arnicaņš, Ģ.Karnītis)

VPP „SOPHIS” 2.projekts

„Uz ontoloģijām balstītas tīmekļa videi pielāgotas zināšanu inženierijas tehnoloģijas”

Apakšprojekti, organizācijas, galvenie izpildītāji:

2.1. Uz ontoloģijām balstītas tīmekļa videi pielāgotas modelēšanas tehnoloģijas un rīki zināšanu analīzei un Uz ontoloģijām balstītas dabīgās valodas semantikas izgūšanas metodes (**LU MII** : J.Bārzdiņš, A.Šostaks, E.Rencis, M.Grasmanis, A.Kalniņš, A.Sproģis, G.Bārzdiņš, D.Goško, P.Paikens)

2.2. Semantiskā tīmekļa tehnoloģijas zināšanu formalizācijai, vairākkārtīgai izmantošanai un koplietošanai (**RTU DITF** : J.Grundspenķis, M.Kirikova, A.Novickis, E.Lavendelis)

2.3. Modeļu bāzēta datu vizualizācija un biznesa procesu izpildes laika verifikācija (**LU DF** : J.Bičevskis, G.Arnicaņš, Ģ.Karnītis)

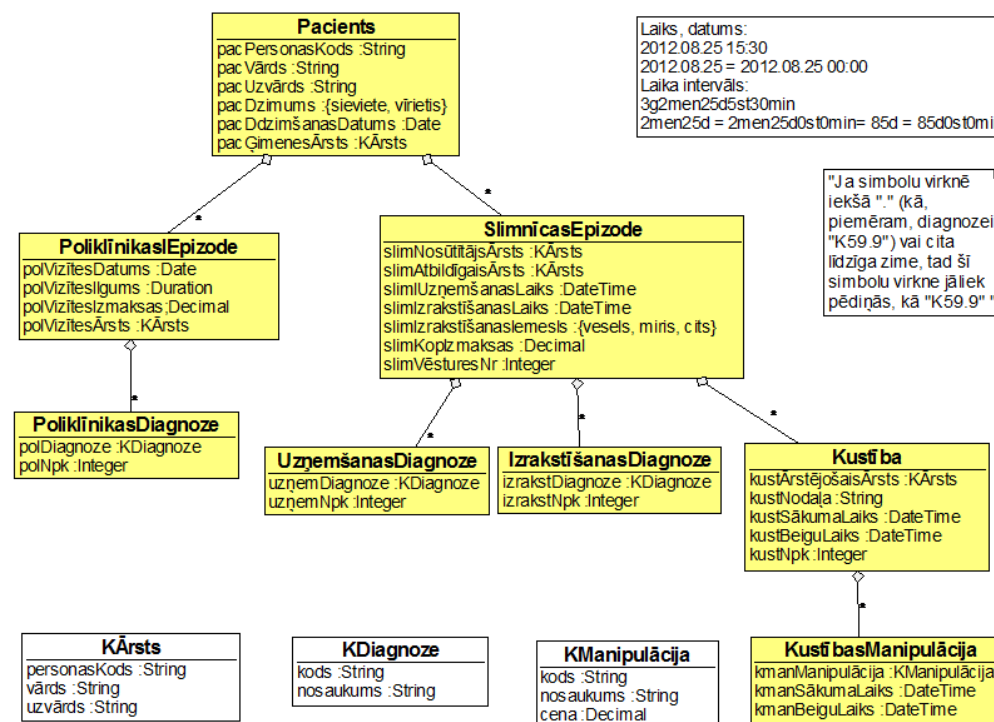
Uz ontoloģijām balstītas tīmekļa videi pielāgotas modelēšanas metodes un rīki zināšanu analīzei:

1. Tālāk attīstīt uz ontoloģijām un tīmekļa tehnoloģijām balstīto ātro vaicājumu valodu un tās efektīvu realizāciju, izstrādājot uz skatiem balstītu paplašināšanas mehānismu.
2. Izstrādāt uz datu ontoloģijām un tīmekļa tehnoloģijām balstītas zināšanu pieejas tiesību modelēšanas metodes sensitīvu datu pieejas tiesību mehānisma realizēšanai ātro vaicājumu valodas realizācijas vajadzībām.
3. Izstrādāt modeļu specializācijas metodes un to lietojumus tīmekļa vidē balstītu domēna specifisko valodu rīku būvei.

Vispirms īss atkārtojums par Ātro vaicājumu valodu

PROBLĒMA: Kā ātri, bez programmētāja starpniecības, piemēram, slimnīcas vadība var iegūt iepriekš neparedzētos griezumus ziņas par slimnīcas darbību.

Ieviešam **puszvaigžņu (semistar)** datu ontoloģijas. Piemērs:



"2016.g. Epizodei 3 jauni atribūti; kliniskā Diagnoze, pamat Diagnoze, galvenā Blakus Diagnoze; Uzņemšanas Diagnozei: 4-apmaksas Diagn"



7 vaicājumu šabloni kontrolētajā dabīgajā valodā

1. SKAITS <klases vārds> KUR <klases selekcijas izteiksme>

Semantika: Tiek saskaitītas Aclass instances, kurām izpildās <Aclass cond expr>.

Piemēri:

SKAITS Pacienti, KURIEM EKSISTĒ SlimnīcasEpizode, KURAI slimNosūtītājsĀrsts=pacĢimenesĀrsts
-- tiek saskaitīti pacienti, kurus uz slimnīcu nosūtījis to ģimenes ārsts,

SKAITS SlimnīcasEpisodes, KURĀM slimIzrakstīšanasLaiks-slimUzņemšanasLaiks > 15d

-- cik bija tādu ārstēšanās reižu, kuras bija ilgākas par 15 dienām,

SKAITS Pacienti, KURIEM EKSISTĒ PoliklīnikasEpizode, KURAI NEEKSISTĒ SlimnīcasEpizode, KURAI slimUzņemšanasLaiks.DATE>=polVizītesDatums

-- cik tādu pacientu, kas pēc poliklīnikas vizītes vairs nav nokļuvuši slimnīcā,

SKAITS SlimnīcasEpisodes e1, KURĀM EKSISTĒ SlimnīcasEpizode e2, KURAI e2<>e1 UN e2.slimUzņemšanasLaiks>e1.slimIzrakstīšanasLaiks UN (e2.slimUzņemšanasLaiks-e1.slimIzrakstīšanasLaiks)<30d

-- cik ir tādu slimnīcas epizožu, pēc kurām pacients ir atgriezies slimnīcā 30 dienu laikā.

7 vaicājumu šabloni kontrolētajā dabīgajā valodā

2. [SUM / MAX / MIN / VID / BIEŽ] /<klases atribūtizteiksme>/ <klases vārds> KUR <klases selekcijas izteiksme>

Semantika: Vispirms tiek atlasītas Aclass instances, kurām izpildās <Aclass select cond>”, pēc tam atlasītajām Aclass instancēm tiek izrēķināta attiecīgās atribūtizteiksmes vērtība (jeb, ja atribūtizteiksme ir vienkārši Aclass atribūts, tad tiek paņemta šī atribūta vērtība), rezultātā tiek iegūts vērtību saraksts, kuram tālāk pielieto attiecīgi operācijas SUM, MAX, MIN, VID, BIEŽ (operācijas SUM, MAX, MIN, AVG ir definētas tikai Integer, Decimal un Duration tipiem, operācija MOST – tikai String tipam).

Piemēri:

SUM /slimKopizmaksas/ SlimnīcasEpisodes, KURĀM slimIzrakstīšanaslemesls=vesels UN pacDzimšanasDatums.YEAR=2012

-- cik izmaksājušas veiksmīgas ārstēšanās slimnīcā pacientiem, kas dzimuši 2012.gadā,

BIEŽ /izrakstDiagnoze.kods/ IzrakstīšanasDiagnozes KURĀM izrakstNpk=1 UN slimIzrakstīšanaslemesls=miris
-- biežākās galvenās (izrakstNpk=1) miršanas diagnozes,

“BIEŽ” vaicājumam ir paredzēta vēl viena opcija – “BIEŽ m”:

BIEŽ 3 /izrakstDiagnoze.kods/ IzrakstīšanasDiagnozes KURĀM izrakstNpk=1 UN slimIzrakstīšanaslemesls=miris

-- atrast 3 biežākās miršanas diagnozes.

7 vaicājumu šabloni kontrolētajā dabīgajā valodā

3. **ATLASĪT NO <klases vārds> KUR <klases selekcijas izteiksme> ATRIBŪTA <klases atribūtizteiksme> VISAS VĒRTĪBAS**

Semantika: Acīmredzama. Piemēri:

ATLASĪT NO SlimnīcasEpizodēm, KURĀM slimIzrakstīšanasIemesls=miris, ATRIBŪTA slimAtbildīgaisĀrsts.uzvārds
VISAS VĒRTĪBAS

ATLASĪT NO SlimnīcasEpizodēm KURĀM slimIzrakstīšanasIemesls=miris, ATRIBŪTA (IzrakstīšanasDiagnoze,
KURAI izrakstNpk=1).izrakstDiagnoze.kods VISAS VĒRTĪBAS

ATLASĪT NO IzrakstīšanasDiagnozēm, KURĀM izrakstNpk=1 UN slimIzrakstīšanasIemesls= miris, ATRIBŪTA
izrakstDiagnoze.kods VISAS VĒRTĪBAS

ATLASĪT NO IzrakstīšanasDiagnozēm, KURĀM izrakstNpk=1 UN slimIzrakstīšanasIemesls= miris, ATRIBŪTA
izrakstDiagnoze.kods.SUBSTRING(1,3) VISAS VĒRTĪBAS

4. **PARĀDĪT [n / VISAS] <klases selekcijas izteiksme>**

Semantika: acīmredzama.

5. **PILNPRĀDĪT [n / VISAS] <klases selekcijas izteiksme>**

Semantika: ...

7 vaicājumu šabloni kontrolētajā dabīgajā valodā

6. **ATLASĪT Aclass x KURĀM <Aclass select cond>, IZVEIDOT TABULU <x-expr'1> [(KOL a1)], ... ,<x-expr'n> [(KOL an)], [ATLASĪT RINDAS KURĀM ai=<konst>] (* = vietā var būt <>, <, <=, >=*) [SAKĀRTOT [AUGOŠI / DILSTOŠI] PĒC aj] [ATSTĀT [PIRMĀS / PĒDĒJĀS] s RINDAS]**

Piemēri:

ATLASĪT SlimnīcasEpisodes x KURĀM slimIzrakstīšanasIemesls=miris, IZVEIDOT TABULU x.pacUzvārds (KOL uzvards), x.slimIzrakstīšanasLaiks.DATE (KOL miršanas_datums), (SKAITS x.KustībasManipulācijas KURĀM kmanManipulacija.kods=02078) (KOL skaits_02078), (SUM /kmanManipulacija.cena/ x.KustībasManipulācijas KURĀM kmanManipulacija.kods=02078) (KOL izmaksas_02078), (x.IzrakstīšanasDiagnoze KURAI izrakstNpk=1).izrakstDiagnoze.kods (KOL pamatdiagnoze), (x.Kustība KURAI kustNpk=).kustNodaļa (KOL pēdējā_nodaļa)*

ATLASĪT Kārstus x KURIEM vārds=Gatis UN EKSISTĒ SlimnīcasEpizode KURAI slimAtbildīgaisĀrsts=x, IZVEIDOT TABULU x.uzvārds (KOL ārsta_uzvārds), (SKAITS SlimnīcasEpisodes KURĀM slimAtbildīgaisĀrsts=x) (KOL epizožu_skaits), (BIEŽ /uznemDiagnoze.kods/ UzņemšanasDiagnozes KURĀM uzņemNPK=1 UN slimAtbildīgaisĀrsts=x) (KOL biežākā_galv_diagnoze)

7 vaicājumu šabloni kontrolētajā dabīgajā valodā

- 7a. ATLASĪT [NO] INTERVĀLA (i-j) VISAS VĒRTĪBAS **x**
[IZVEIDOT TABULU/ RĀDĪT] <izteiksme'1> [(KOL a1)], ... , <izteiksme'n> [(KOL an)]
- 7b. ATLASĪT [NO] <klases vārds> KUR <klases selekcijas izteiksme>
ATRIBŪTA <klases atribūtizteiksme> VISAS VĒRTĪBAS **x**
IZVEIDOT TABULU <izteiksme'1> [(KOL a1)], ... , <izteiksme'n> [(KOL an)]

Piemēri:

ATLASĪT NO SlimnīcasEpizodēm, KURĀM slimIzrakstīšanasIemesls=vesels, ATRIBŪTA slimAtbildīgaisĀrsts.personasKods VISAS VĒRTĪBAS x, IZVEIDOT TABULU x (KOL a1), (KĀrstis KURAM personasKods=x).vārds (KOL a2), (KĀrstis KURAM personasKods=x).uzvārds (KOL a3)

ATLASĪT NO INTERVĀLA (0-23) VISAS VĒRTĪBAS x, IZVEIDOT TABULU x (KOL stunda), (SKAITS SlimnīcasEpisodes KURĀM slimUzņemšanasLaiks.HOUR=x) (KOL skaits)

Piemērs vaicājumu rīkā

Testu vide BKUS vaicājumiem (versija: feb08, datu kopa: 2015)

Saņemtais vaicājums: atlasīt visas Epizodes kuram izriem.iemeslakods=3, izveidot tabulu patients.dzimums (kol Dzimum), patients.dzimsanasdatums (kol Dzim_dat), izrakstisanaslaiks.Date (kol Mir_atums), pamatdiagnoze (kol Pam_diagn), (Kustiba kurai npk=*).nodala (kol Mir_nod)

Tulkošana: 4 ms, kompilācija: 518 ms, izpilde: 8 ms

Atbilde ir:

dzimum	dzim_dat	mir_atums	pam_diagn	mir_nod
1	██████████	2015.01.09		01_1-69
2	██████████	2015.01.23		12-69
1	██████████	2015.02.23		12-69

parādīt visu *Epizode*, kur (izrIem.iemeslaKods=3), rādot patients.dzimums (kolonna dzimum), patients.dzimsanasDatums (kolonna dzim_dat), izrakstisanasLaiks.date() (kolonna mir_atums), pamatdiagnoze (kolonna pam_diagn), (Kustiba, kur (npk=*)).nodala (kolonna mir_nod)

atlasīt visas Epizodes kuram izriem.iemeslakods=3, izveidot tabulu patients.dzimums (kol Dzimum), patients.dzimsanasdatums (kol Dzim_dat), izrakstisanaslaiks.Date (kol Mir_atums), pamatdiagnoze (kol Pam_diagn), (Kustiba kurai npk=*).nodala (kol Mir_nod)

Izpildīt

☒ Rādīt saprasto? ☒ Automātiski saprast? ☒ Tīrīt pirms izpildes?

1. Tālāk attīstīt uz ontoloģijām un tīmekļa tehnoloģijām balstīto ātro vaicājumu valodu un tās efektīvu realizāciju, izstrādājot uz skatiem balstītu paplašināšanas mehānismu

Uz skatiem balstīts Ātrās vaicājumu valodas paplašinājums:

DEFINĒT Wclass = Aclass [x] KUR <selection condition>

Piemērs:

definēt VeiksmīgsPacients=Pacients kuram visām SlimnīcasEpizodēm izpildās
slimIzrakstīšanasIemesls=vesels

Ātrdarbība: ~0.3 sek. uz tipiskiem BKUS ātrajiem vaicājumiem

Publikācijas:

1. J.Barzdins, M.Grasmanis, E.Rencis, A.Sostaks, J.Barzdins, Ad-Hoc Querying of Semistar Data Ontologies Using Controlled Natural Language. // In: G.Arnicans, V.Arnican, J.Borzovs, L.Niedrite (Eds.), Frontiers of AI and Applications, Vol. 291, Databases and Information Systems IX, IOS Press, pp. 3-16, 2016. (to be indexed **SCOPUS**), <http://ebooks.iospress.com/volumearticle/45695>
2. J.Barzdins, M.Grasmanis, E.Rencis, A.Sostaks, A.Steinsbekk, Towards a More Effective Hospital: Helping Health Professionals to Learn from their Own Practice by Developing an Easy to use Clinical Processes Querying Language. // J.E.Q. Varajão et. al. (Eds.), Procedia Computer Science, Vol. 100, Elsevier, pp. 498-506, 2016.

Adrese, kur ar Ātro vaicājumu rīku var “paspēlēties” lietojot latviešu valodas versiju: <http://85.254.199.40/> (lietot Google Chrome!)

2. Izstrādāt uz datu ontoloģijām un tīmekļa tehnoloģijām balstītas zināšanu pieejas tiesību modelēšanas metodes sensitīvu datu pieejas tiesību mehānisma realizēšanai ātro vaicājumu valodas realizācijas vajadzībām

- ✓ Role-based Access Control, Attribute-based Access Control,
- ✓ XACML (eXtensible Access Control Markup Language) - ietver iepriekšējās u.c.
- ✓ Metodes universālas, bet komplicētas.
- ✓ Puszvaigžņu (semistar) datu ontoloģijas un ātrā vaicājumu valoda – vai nevar vienkāršāk?

Galvenais rezultāts:

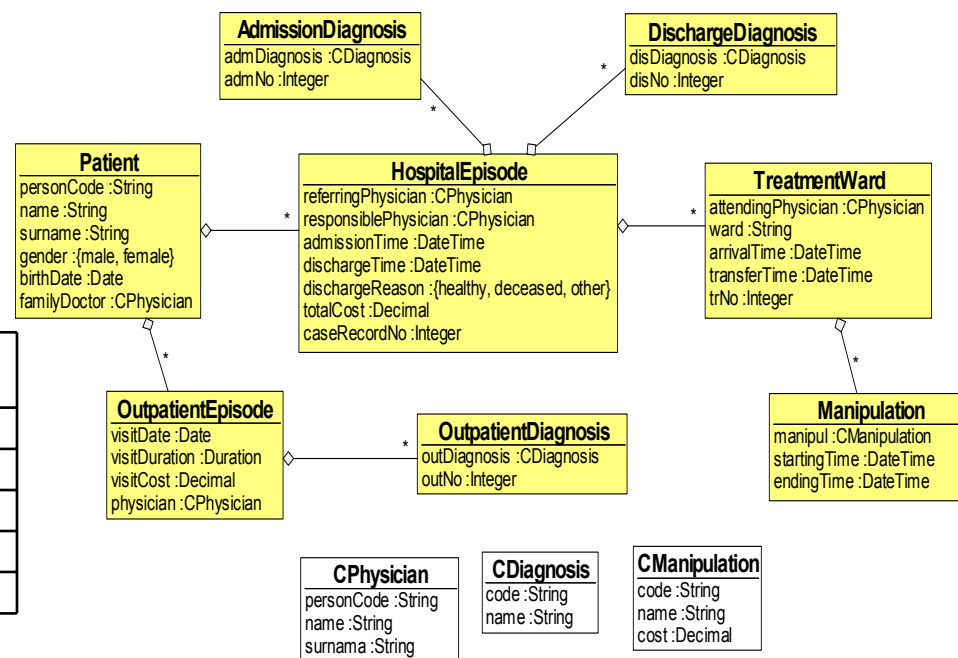
- Izstrādāts pietiekoši vispārīgs pieejas tiesību definēšanas mehānisms priekš minētās vaicājumu valodas, kas izmanto jauna veida lomas jēdzienu, kas tiek definēts ar šīs pašas vaicājumu valodas līdzekļiem, tikai nedaudz tos papildinot.
- Un proti, lomas definēšana balstās uz datu atlasī priekš attiecīgā lietotāja ar formulas tipa “FULLSELECT <class name> WHERE <select condition> WITHOUT <class name>[.<attribute name>], ...” palīdzību, kur FULLSELECT nozīmē, ka reizē ar norādītas klases instanču atlasī notiek arī tās “bērnu” un “vecāku” atlase.
- Korekta šādas atlases definīcija ir iespējam pateicoties aplūkojamo ontoloģiju zvaigžņuveida struktūrai
- Pētīta arī minēto pieejas tiesību ietekme uz vaicājumu izpildes ātrdarbību, sagaidāmais palēninājums – ne vairāk kā 2 reizes, bet tas prasa vēl tālākus eksperimentālus pētījumus, kas paredzēti projekta nākamajā posmā.

Pieejas tiesību definēšanas piemērs:

Authentication	Person code C	Role	Role parameter	Access rule
Password	C	Responsible Physician	–	FULLSELECT HospitalEpisode WHERE responsiblePhysician.personCode = C WITHOUT Patient.personCode
Password	[C]	Ward manager	n	FULLSELECT HospitalEpisode WHERE EXISTS TreatmentWard WHERE ward = n WITHOUT Patient.personCode
Password	[C]	Hospital CEO	–	FULLSELECT Patient WITHOUT Patient.personCode
External authentication (via Swedbank)	C	Patient	–	FULLSELECT Patient WHERE personCode = C
Open access	–	Journalist	–	FULLSELECT Patient WITHONLY HospitalEpisode.totalCost WITHOUT Patient.ALL, TreatmentWard.ALL, Manipulation.ALL, OutPatientEpisode.ALL, AdmissionDiagnosis, DischargeDiagnosis

An example of a filled access table:

Authentication	Person code C	Role	Role parameter
ArpCrV01	011040-11111	Responsible Physician	–
BrpHrtD02	301191-99999	Responsible Physician	–
CmnEmT01	–	Ward manager	12
DmnPlm03	–	Ward manager	07
EceDrBKS0	–	Hospital CEO	–



3. Izstrādāt modeļu specializācijas metodes un to lietojumus tīmekļa vidē balstītu domēna specifisko valodu rīku būvei

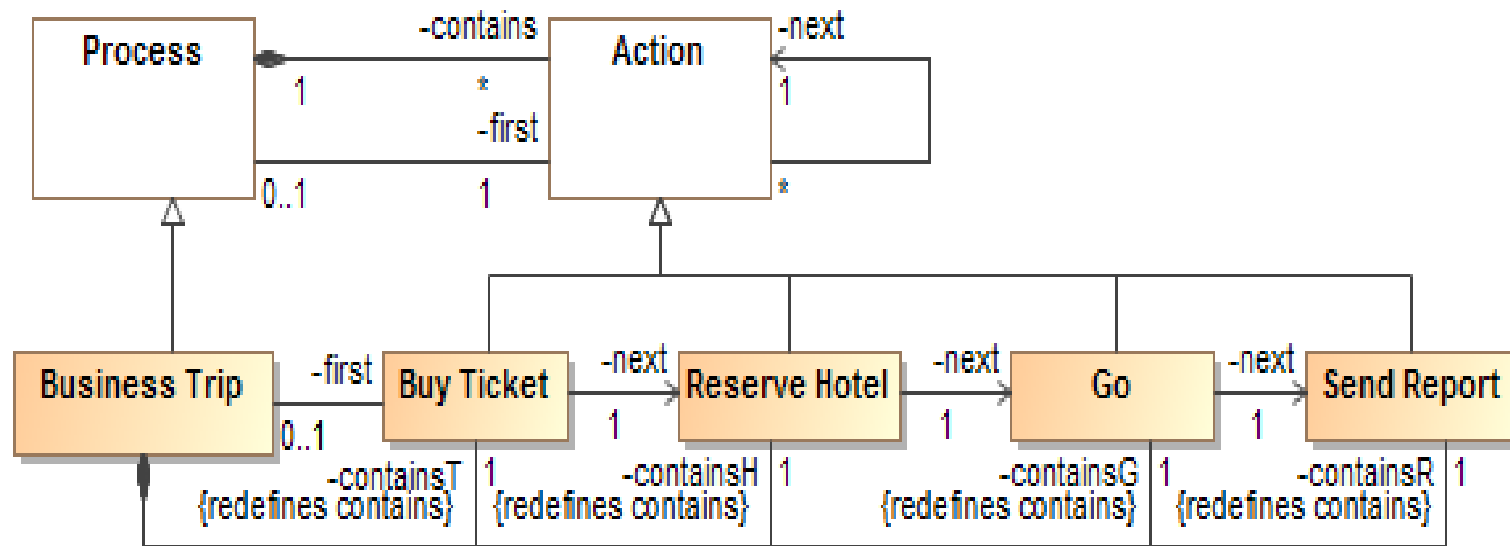
Publikācijas:

1. A.Kalnins, J.Barzdins, Metamodel Specialization for DSL Tool Building. // In: G.Arnicans, V.Arnicanē, J.Borzovs, L.Niedrite (Eds.), Databases and Information Systems, 12th International Baltic Conference, DB&IS 2016, Riga, Latvia, July 4-6, 2016, Proceedings, Communications in Computer and Information Science Vol. 615, Springer, pp.68-82, 2016. (SCOPUS)
2. A.Sprogis, DSML Tool Building Platform in WEB. // In: G.Arnicans, V.Arnicanē, J.Borzovs, L.Niedrite (Eds.), Databases and Information Systems, 12th International Baltic Conference, DB&IS 2016, Riga, Latvia, July 4-6, 2016, Proceedings, Communications in Computer and Information Science Vol. 615, Springer, pp.99-109, 2016. (SCOPUS)
3. **A.Kalnins, J.Barzdins, Metamodel specialization for graphical modeling language support. // In: Proceedings of the ACM/IEEE 19th International Conference on Model Driven Engineering Languages and Systems. ACM, pp.103-112, 2016. (to be indexed SCOPUS)**
4. A.Sprogis, ajoo: WEB Based Framework for Domain Specific Modeling Tools. // In: G.Arnicans, V.Arnicanē, J.Borzovs, L.Niedrite (Eds.), Frontiers of AI and Applications, Vol. 291, Databases and Information Systems IX, IOS Press, pp. 115-126, 2016. (to be indexed SCOPUS), <http://ebooks.iospress.com/volumearticle/45704>

Dear Audris and Janis,

Thank you for your submission to one of the MODELS 2016 tracks. It is our pleasure to inform you that your paper: "Metamodel Specialization for Graphical Modeling Language Support" with id: 70 has been accepted for publication in the Foundations Track of the conference proceedings. Each paper was reviewed by at least three members of the Program Committee (PC) and the reviews were monitored by the Program Board (PB). Each paper was also extensively discussed during the online PC meeting, and due consideration was given to author responses that were provided. On July 1st-2nd, a PB meeting was held in Vienna, Austria, where all PB members attended. During that meeting, each paper was again discussed, and the paper selection was finalized. **This year, out of 118 papers submitted to the Foundations Track, the PC and the PB accepted 28 (acceptance rate 23.7%).**

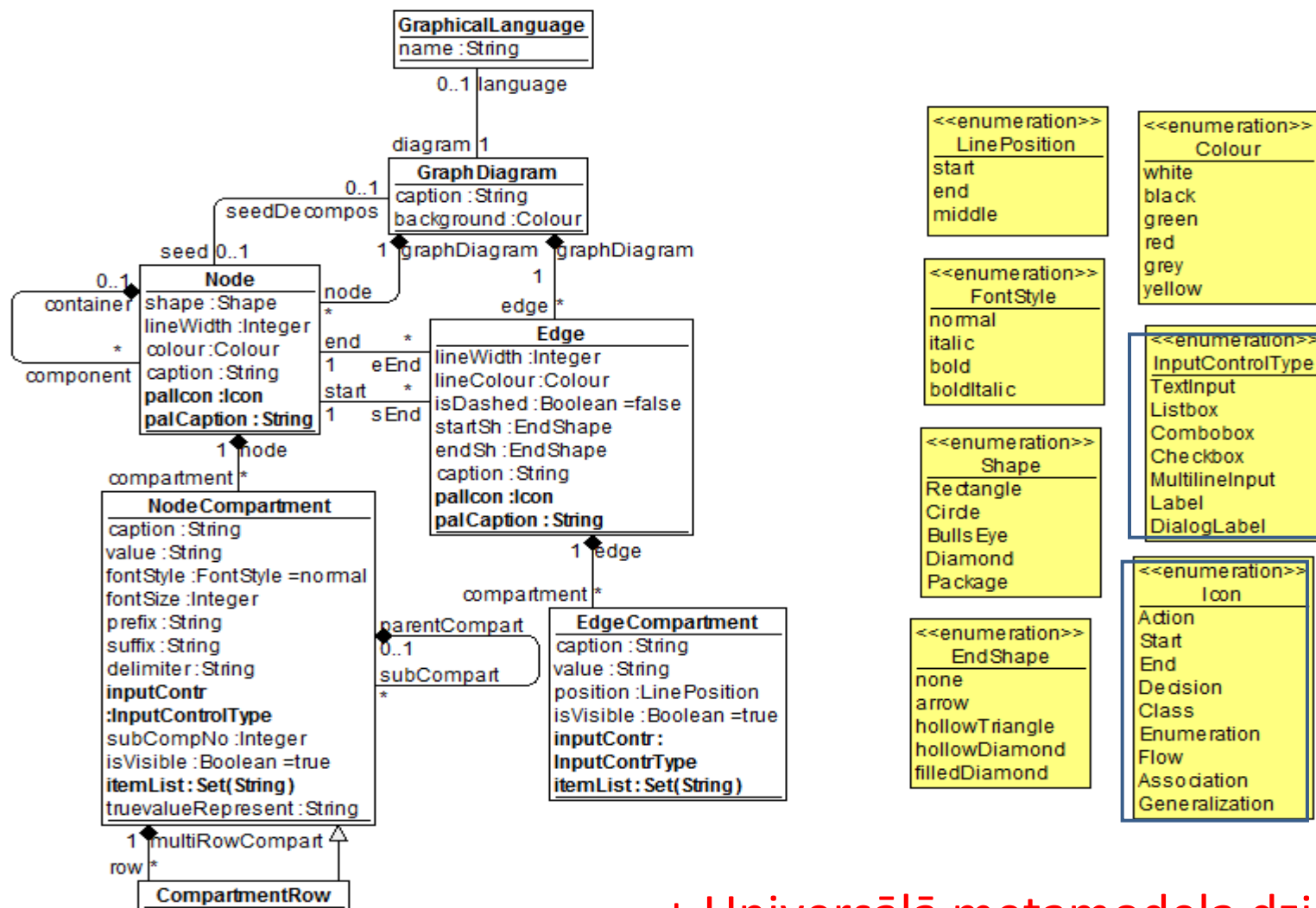
Kas ir modeļu (metamodeļu) specializācija – piemērs:



- Apakšklases
- Atribūtu «redefines»
- Asociāciju galu «redefines»
- OCL ierobežotāji

← Likumīgi UML līdzekļi

Universālais metamodelis priekš grafiskās valodas sintakses + redaktoru definīcijas ar **specializācijas** palīdzību



+ Universālā metamodeļa dzinis

Izstrādāta arī grafisko rīku būves platforma tīmekļa videi, kas lielā mērā balstīta uz minētajām idejām

Publikācija:

A.Sprogis, ajoo: WEB Based Framework for Domain Specific Modeling Tools. // In: G.Arnicans, V.Arnican, J.Borzovs, L.Niedrite (Eds.), Frontiers of AI and Applications, Vol. 291, Databases and Information Systems IX, IOS Press, pp. 115-126, 2016. (to be indexed SCOPUS), <http://ebooks.iospress.com/volumearticle/45704>

Nodrošina:

- Daudzlietotāju režīmu (collaboration)
- Dažādas iekārtas – datori, planšetes, telefoni...
- Darbību reālā laikā (reactivity and live HTML)

Izmantotās tehnoloģijas:

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

Turpmākie pētījumi:

- MM specializācijas (un to dziņu) tālāka attīstība
- MM specializācijas jauni lietojumi

Uz ontoloģijām un dziļās mašīnāpmācības metodēm balstītas dabīgās valodas semantikas izgūšanas metodes:

Piedalīties SemEval-2016 starptautiskajās sacensībā ar uzlabotu C6.0 klasifikācijas algoritma versiju, kas pielāgota "Abstract Meaning Representation" (AMR) izgūšanai no dabiskās valodas teksta.

Uzvara SemEval-2016, Task 8: Meaning Representation Parsing (AMR)



1. *RIGA (University of Latvia, IMCS; LETA): 0.6196*
2. *CAMR (Brandeis University; Boulder Learning Inc.; Rensselaer Polytechnic Institute): 0.6195*
3. *ICL-HD (Ruprecht-Karls-Universität Heidelberg): 0.6005*
4. *UCL+Sheffield (University College London; University of Sheffield): 0.5983*
5. *M2L (Kyoto University): 0.5952*
6. *CMU (Carnegie Mellon University; University of Washington): 0.5636*
7. *CU-NLP (OK Robot Go, Ltd.; University of Colorado): 0.5566*
8. *UofR (University of Rochester): 0.4985*
9. *MeaningFactory (University of Groningen): 0.4702*
10. *CLIP@UMD (University of Maryland): 0.4370*
11. *DynamicPower (National Institute for Japanese Language and Linguistics): 0.3706*

<http://summa-project.eu/blog/leta-wins-amr-parsing-trophy-at-semeval-2016/>

Publikācija SemEval-2016, Task 8: Meaning Representation Parsing (AMR)

```
(f / fear-01
  :polarity "-"
  :ARG0 ( s / soldier )
  :ARG1 ( d / die-01
    :ARG1 s ))
```

The soldier was not afraid of dying.

The soldier was not afraid to die.

The soldier did not fear death.

```
(h / have-org-role-91
  :ARG0 (p / person
    :wiki "David_Cameron"
    :name (n / name :op1 "David"
      :op2 "Cameron"))
  :ARG1 (c / country
    :wiki "United_Kingdom"
    :name (n2 / name :op1 "United"
      :op2 "Kingdom"))
  :ARG2 (m / minister
    :mod (p2 / prime)))
```

“David Cameron is the Prime Minister of the United Kingdom.”

Guntis Barzdins, Didzis Gosko. RIGA at SemEval-2016 Task 8: **Impact of Smatch Extensions and Character-Level Neural Translation on AMR Parsing Accuracy**. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), SanDiego CA, Association for Computational Linguistics, pp. 1143-1147. (<http://aclweb.org/anthology/S16-1176>)

Mūsu open-source AMR parseris ātrakais un precīzākais pasaulē, tiek izmantots H2020 SUMMA projektā, NIST TAC-KBP sacensībā un citur

Novel AMR parsing methods

F1=66%

Sentences Other

Precision: 70.493 %
Recall: 61.496 %
Best F-Score: 65.688 %
Good: 1370
Skipped: 0
1st sort: Sentence #
2nd sort: no sorting

Filter text... (press Enter)
040. [0.0%] Late on a AU...
847. [82%] Tskhinvali is t...
848. [83%] Georgia is m...
849. [73%] Georgia is a r...
850. [0%] 2008-09-19
851. [100%] Iraq (IQ)
852. [78%] International; ...
853. [54%] Iranian News ...
854. [50%] The official w...
855. [75%] Group XP sta...
856. [81%] The official w...
857. [81%] It was not cle...
858. [91%] By the evenin...
859. [73%] Another web...
860. [74%] An alternative...
861. [62%] Iranian News ...
862. [83%] Neither the h...
863. [74%] Several Irania...
864. [56%] A statement ...
865. [25%] The statemen...

Georgia is mountainous.

Georgia is mountainous.

AMR

(m / mountainous
:domain (c / country :wiki "Georgia_(country)"
:name (n / name :op1 "Georgia")))

(x3 / mountain
:ARG1 (x1 / country
:wiki "Georgia_(country)" :name (n / name
:op1 "Georgia")))

Precision: 62.500 %
Recall: 62.500 %
Best F-Score: 62.500 %

Root Instance

m / mountainous x3 / mountain

Instances

c / country x1 / country
n / name n / name
no matching instance x3 / mountain
m / mountainous no matching instance

Attributes

c / country wiki Georgia_(country) x1 / country wiki Georgia_(country)
n / name op1 Georgia n / name op1 Georgia

Relations

c / country name n / name x1 / country name n / name
no matching relation x3 / mountain ARG1 x1 / country
m / mountainous domain c / country no matching relation

Rules

class: MATCHED_INSTANCE, score: 0.8421, yes/no: 31/5 data: GOLD_TYPE = president
class: GOLD_INSTANCE, score: 0.6667, yes/no: 1/0 data: GOLD_TYPE = mountainous
class: GOLD_ATTRIBUTE, score: 0.8750, yes/no: 13/1 data: GOLD_CONST = Foreign GOLD_TYPE = name GOLD_RELATION = op1
class: SILVER_RELATION, score: 0.9524, yes/no: 59/2 data: SILVER_TYPE1 = person, SILVER_RELATION = ARG1
class: GOLD_RELATION, score: 0.8871, yes/no: 109/13 data: GOLD_RELATION = ARG3
class: MATCHED_INSTANCE, score: 0.8421, yes/no: 31/5 data: GOLD_TYPE = president

Georgia is mountainous.

↓ F1=43%

(mountain-01
:ARG1 (country
:name (name :op1 "Georgia"))

↓ F1=97%

(x3 / mountain
:ARG1 (x1 / country
:wiki "Georgia_(country)" :name (n / name
:op1 "Georgia")))

Smatch Extended with C6.0
for systematic error spotting

Character-Level Neural Translation
English → simplified AMR
and deterministic extension to AMR

Ensemble: F1=67% (62% on the official scoring set)

Citas publikācijas:

- N. Gruzitis and G. Barzdins. The role of CNL and AMR in scalable abstractive summarization for multilingual media monitoring. Controlled Natural Language, Controlled Natural Language 5th International Workshop, CNL 2016, Davis, Brian, Pace, Gordon J., Wyner, Adam (Eds.), LNAI, Volume 9767, pp. 127-130, Springer 2016. doi = "10.1007/978-3-319-41498-0" (to be indexed SCOPUS)
- Peteris Paikens. Deep Neural Learning Approaches for Latvian Morphological Tagging. Frontiers in Artificial Intelligence and Applications, Volume 289: Human Language Technologies – The Baltic Perspective, I. Skadiņa and R. Rozis (Eds.). IOS Press, 2016, pp 160-166. DOI 10.3233/978-1-61499-701-6-160 <http://ebooks.iospress.nl/volumearticle/45531> (to be indexed SCOPUS)



Full Stack of Language Resources for Natural Language Understanding and Generation in Latvian

NACIONĀLAIS
ATTĪSTĪBAS
PLĀNS 2020



EIROPAS SAVIENĪBA
Eiropas Reģionālās
attīstības fonds

projekts

IEGULDĪJUMS TAVĀ NĀKOTNĒ



Latvijas Universitātes
Matemātikas un informātikas institūta
Mākslīgā intelekta laboratoriju (AI-Lab)

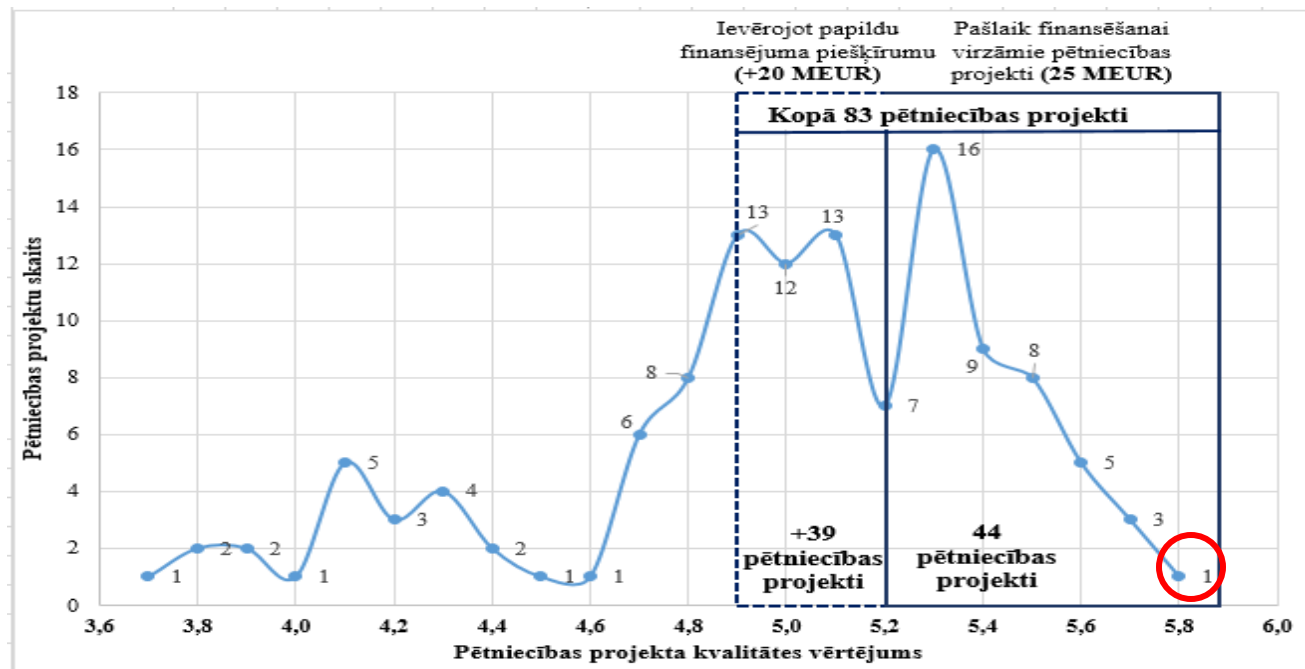


LETA

nacionālā informācijas aģentūra

full stack

- Total overall Budget: 600 K€
- Total timing of the project: 3 years
- Start date: December 1, 2016



Citi nākotnes plāni (1)

- Piedalīties SemEval-2017, Task 9: Abstract Meaning Representation (AMR) Parsing and Generation
 - Kopā ar IBM Brazīlija, Chalmers University, Tohoku university
 - Tekstrade izmantojot AMR→GF konvertāciju un neironu mašintulkošanu

Abstract Meaning Representation (AMR)	GF Abstract Syntax Tree (AST)
<pre>(x3 / take-01 :ARG0 (x2 / organization :wiki "Syrian_opposition" :name (n2 / name :op1 "Syrian" :op2 "rebels")) :ARG1 (x5 / city :wiki "Aleppo" :name (n / name :op1 "Aleppo")))</pre>	<pre>PredVP organization_Syrian_opposition_NP (ComplV2 take_over_V2 city_Aleppo_NP)</pre>

Fig. 2. AMR and AST representations of “*Syrian rebels took over Aleppo*”

Citi nākotnes plāni (2)

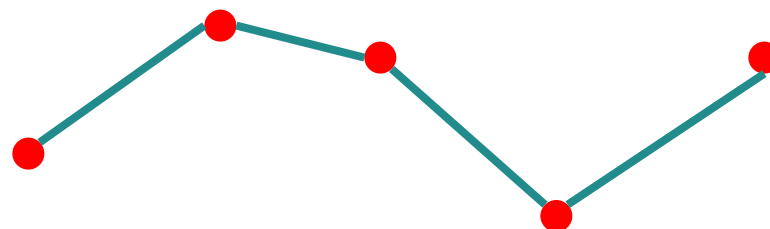
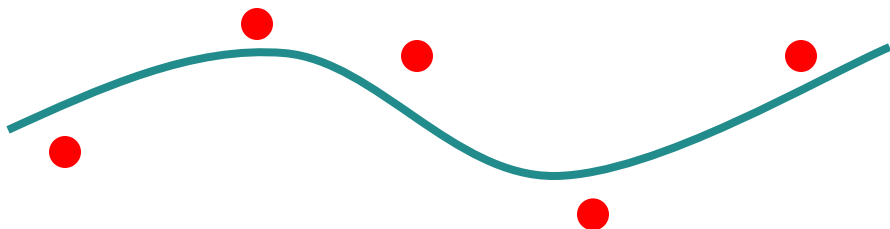
- Immitation Learning
 - Apvieno Deep Learning (Reinforcement Learning with SGD) un Episodic Memory (one-shot learning)

Deep Learning

- Mācās lēni ar SGD

Episodic Memory

- One-shot learning



Citi nākotnes plāni (3)

- Dalība NIST TAC-KBP Cold Start sacensībā
 - Iegūto rezultātu integrācija LETA ontoloģijas automatiskā profilu izgūšanas sistēmā
 - Data Science metožu izpēte un integrācija

The screenshot displays a web application interface for LETA visualization. The browser address bar shows the URL: `bebrs.aifab.lv:7600/#/databases/mkold/entities/2199813/summaryframes/1744312/frames/2`. The application header includes a navigation bar with tabs: **CV Personas**, **CV Organizācijas**, **Personas**, **Organizācijas**, **Entiņu meklēšana**, **Kartība: Gordons Brauns**, **Faktu redaktors**, and **Nepabeigto fakti**. Below the header, the main content area is titled **2199813 > Gordons Brauns** with a sub-header **Apraksts...**. The interface includes a filter section with **Tips: persona** and **Dzimte: bez dzimtes**. The main content area is divided into several sections: **Apbalvojumi, interesanti fakti un izteicieni**, **Nodarbošanās un publiskā darbība**, **Ipašumi un finanses**, and **Incidenti un tiesvedības**. Each section contains a list of items with details and action buttons. On the right side, there is a sidebar with a search bar and a list of related entities, including **ministers** and **Gordons Brauns**.

Režims: pilns Datubāze: mkold

CV Personas CV Organizācijas Personas Organizācijas Entiņu meklēšana Kartība: Gordons Brauns Faktu redaktors Nepabeigto fakti

2199813 > Gordons Brauns Apraksts... GRAFI Konsolidēt

Tips: persona Dzimte: bez dzimtes

Filtrēt: Kārtot: jaunākie Rādīt: ☒ Derīgs ☒ Jāpārbauda! ☐ Slēpts ☐ Noderīgs

Apbalvojumi, interesanti fakti un izteicieni

- [1] datums: Gordons Brauns ieguva formālu karalienes Elizabetes II piekrišanu ☒ Derīgs ☐ Noderīgs ☐ Cits
- [1] datums: Gordons Brauns informē: "Notiek pastāvīgas problēmas ar cilvēku izslēgšanu no troņa pārmantošanas tiesībām, un notiek lēmums par problēmas, kuras jāatrisina."

Nodarbošanās un publiskā darbība

- [1] datums: Gordons Brauns bija ekonomikas ministra amatā ☒ Derīgs ☐ Noderīgs ☐ Cits
- [2] datums: Gordons Brauns bija aizsardzības ministra amatā ☒ Derīgs ☐ Noderīgs ☐ Cits
- [1] datums: Gordons Brauns bija finanšu ministra amatā ☒ Derīgs ☐ Noderīgs ☐ Cits
- [10] datums: Gordons Brauns bija finanšu ministra amatā ☒ Derīgs ☐ Noderīgs ☐ Cits

Ipašumi un finanses

Incidenti un tiesvedības

- 4329222 pieminējums avota my_D8E4AE24-A8A4-4562-968D-6C88994BC72E teikumā nr. 7 [redīgēt >>](#)
- 4201206 pieminējums avota my_89B87409-B8CF-11D8-89DF-00C96B78513F teikumā nr. 29 [redīgēt >>](#)
- 3918246 pieminējums avota my_B2C53357-71D1-467D-8F34-D9CAC201B34A teikumā nr. 5 [redīgēt >>](#)
- 3846812 pieminējums avota my_518D5896-D5B7-44AB-A737-975B2B85EC9F teikumā nr. 3 [redīgēt >>](#)
- 3832206 pieminējums avota my_E34F3206-3DEA-4EAC-A14B-BC46521F0F30 teikumā nr. 2 [redīgēt >>](#)
- 3814010 pieminējums avota my_48F80F01-0D5A-46C8-88D0-9AC9B584F8F teikumā nr. 1A [redīgēt >>](#)

Nezināms Izrakstīties

atrisināt Amerikas problēmas

5. Lielbritānijas finanšu

ministers Gordons Brauns

Belong_employed Employee

, kas iestājas par parādu

atcelšanu, BBC televīzijai

atziņēja, ka sabiedriskā

doma, baznīcas un

cilvēktiesību grupējumi

pēdējā laikā jau ir

ietekmējuši lēmumus šai

jomā, tomēr vienlaikus

atzina, ka Āfrikas

nabadzības problēmas

risināšana ir "mūža darbs"

Rezultatīvais rādītājs	Rezultāti			
	plānots	sasniegts		
	2014.–2017. g.	2014.	2015.	2016.
Zinātniskie rezultatīvie rādītāji				
1. Zinātnisko publikāciju skaits:				
oriģinālo zinātnisko rakstu (SCOPUS) (SNIP > 1) skaits	4 (-3)			1
Oriģinālo zinātnisko rakstu IEEExplore, ACM DL, SCOPUS, Web of science datubāzēs iekļautajos izdevumos skaits	22 (+17)	6	12	21
Citu oriģinālo zinātnisko rakstu skaits		2	4	6
Sagatavotas un iesniegtas			1	3
2. Programmas ietvaros aizstāvēto darbu skaits:				
promocijas darbu skaits	9 (-4)	1	2	2
maģistra darbu skaits	22 (+1)	6	1	16
Programmas popularizēšanas rezultatīvie rādītāji				
3. Programmas gaitas un rezultātu popularizēšanas interaktīvie pasākumi, kuru mērķu grupās iekļauti arī izglītojamie, skaits:				
konferences	14 (-2)		3	9
semināri			1	1
rīkotie semināri	9 (-3)	1	3	2
rīkotās konferences	1 (+0)		1	1(DB&IS)
populārzinātniskas publikācijas	3 (-3)			
4. Ilgtermiņa tehnoloģiskā prognoze	1 (-1)	-		

Rezultatīvais rādītājs				
	plānots	sasniegts		
	2014.–2017. g.	2014.	2015.	2016.
Tautsaimnieciskie rezultatīvie rādītāji				
1. Zinātniskajai institūcijai programmas ietvaros piesaistītā privātā finansējuma apjoms, t. sk.:				
1.3. ieņēmumi no līgumdarbiem, kas balstās uz programmas ietvaros radītajiem rezultātiem un zinātības	320000 (+1037500)	197500	1160000**	
1.4 Zinātnisko institūciju līdzfinansējums no pašu līdzekļiem programmas izpildei (EUR):	10000 (-10000)			
3. Programmas ietvaros izstrādāto jauno tehnoloģiju, metožu, prototipu vai pakalpojumu skaits, kas aprobēti uzņēmumos	6 (-4)	-	2	

** Eiropas IKT pētniecības projekts SUMMA, kas ir iegūts, izmantojot VPP izstrādes, un kurš norisināsies no 2016. gada 1. februāra līdz 2019. gada 31. janvārim. Projektu realizē SIA „LETA” un VPP iesaistītie LUMII pētnieki.

Paldies par uzmanību!