

Pētījumi datu noliktavās

Datu noliktavu evolūcija

Problēma:

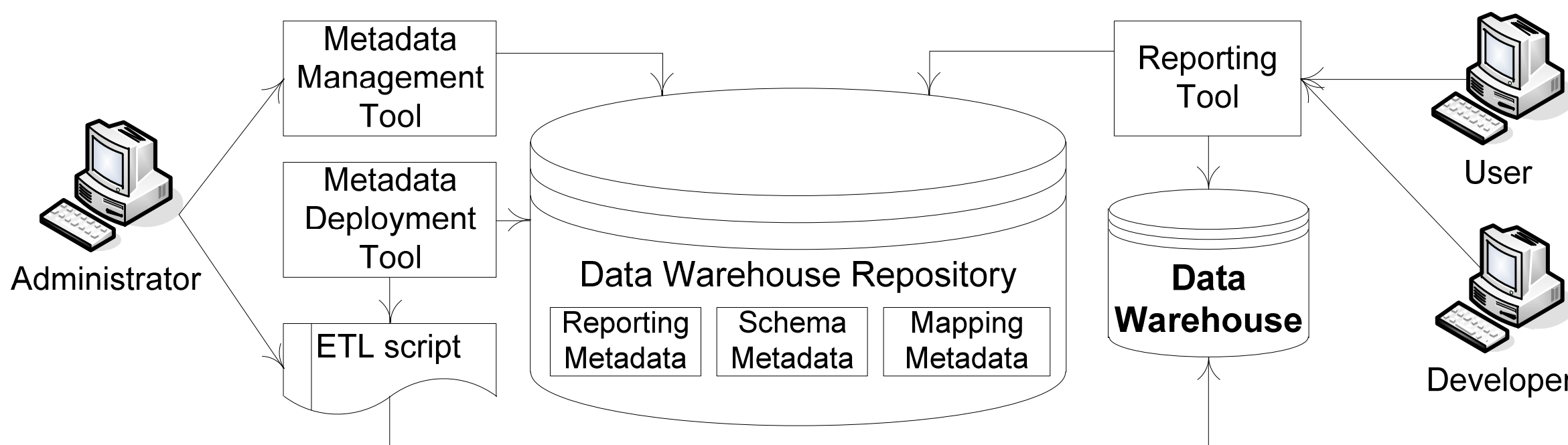
- Izmaiņas var notikt datu avotos, darījumphrasībās
- Sekas: izmaiņas datu noliktavu shēmās, ielādes procesos, eksistējošās atskaitēs.

Pētījuma mērķis: jānodrošina izmaiņu korekta apstrāde datu noliktavā

Rezultāti:

- Izstrādāta datu noliktavas evolūcijas arhitektūra, nedefinēts, kāda tipa izmaiņas tā atbalsta.
- Izstrādāts daudzversiju datu noliktavas modelis, kas balstīts uz vairāku līmeņu metadatiem
- Izstrādāts atskaišu rīks, kas atbalsta atskaišu izpildi uz vairākām versijām (balstīts uz piedāvāto datu noliktavas shēmas daudzu versiju apstrādes algoritmu).

Datu noliktavas evolūcijas arhitektūra



- Datu noliktavas evolūcijas arhitektūra sastāv no izstrādes vides un lietotāju vides.
- Izstrādes vidē atrodas datu noliktavas metadatu repozitorijs un citi komponenti, notiek datu ieguves, pārveidošanas, ielādes procesi un izmaiņu apstrāde.
- Lietotāju vidē datu noliktavas izstrādātāji un lietotāji definē un izpilda atskaites uz vienu vai vairākām datu noliktavas shēmas versijām.

Datu noliktavu pētniecības grupa

Tēmas

1. Datu noliktavu evolūcija
2. Datu noliktavu personalizācija
3. Procesu datu noliktavas

Dalībnieki

Laila Niedrite
Darja Solodovņikova
Natālija Kozmina
Aivars Niedritis
Jānis Zuters
Māris Treimanis

Procesu datu noliktavas

Problēma:

- Datu noliktavu datu modeļa viens no pamatelementiem ir fakts (notikums) un mērījumi par faktiem. Procesu datu noliktavas tiek veidotas mērījumu par procesiem analīzei. Nepieciešams izstrādāt metodes procesu datu noliktavu modeļu izstrādei, realizācijai un pielietošanai biznesa procesu mērīšanai.

Pētījuma mērķi:

- izstrādāt procesu mērīšanas ietvaru un metodi, lai nodrošinātu biznesa procesu mērīšanu, izmantojot datu noliktavu,
- izstrādāt arhitektūru, kas nodrošinātu adaptējamu individuālu lietotni, un kas ietvertu biznesa procesu mērīšanas sistēmu.

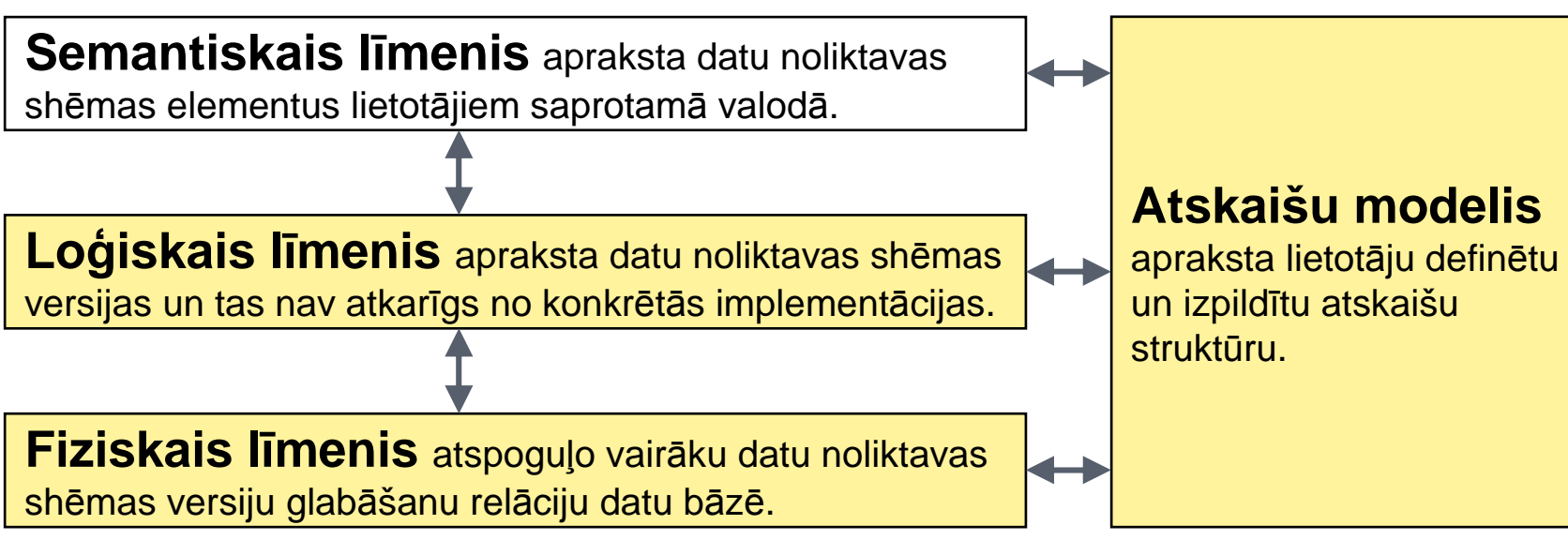
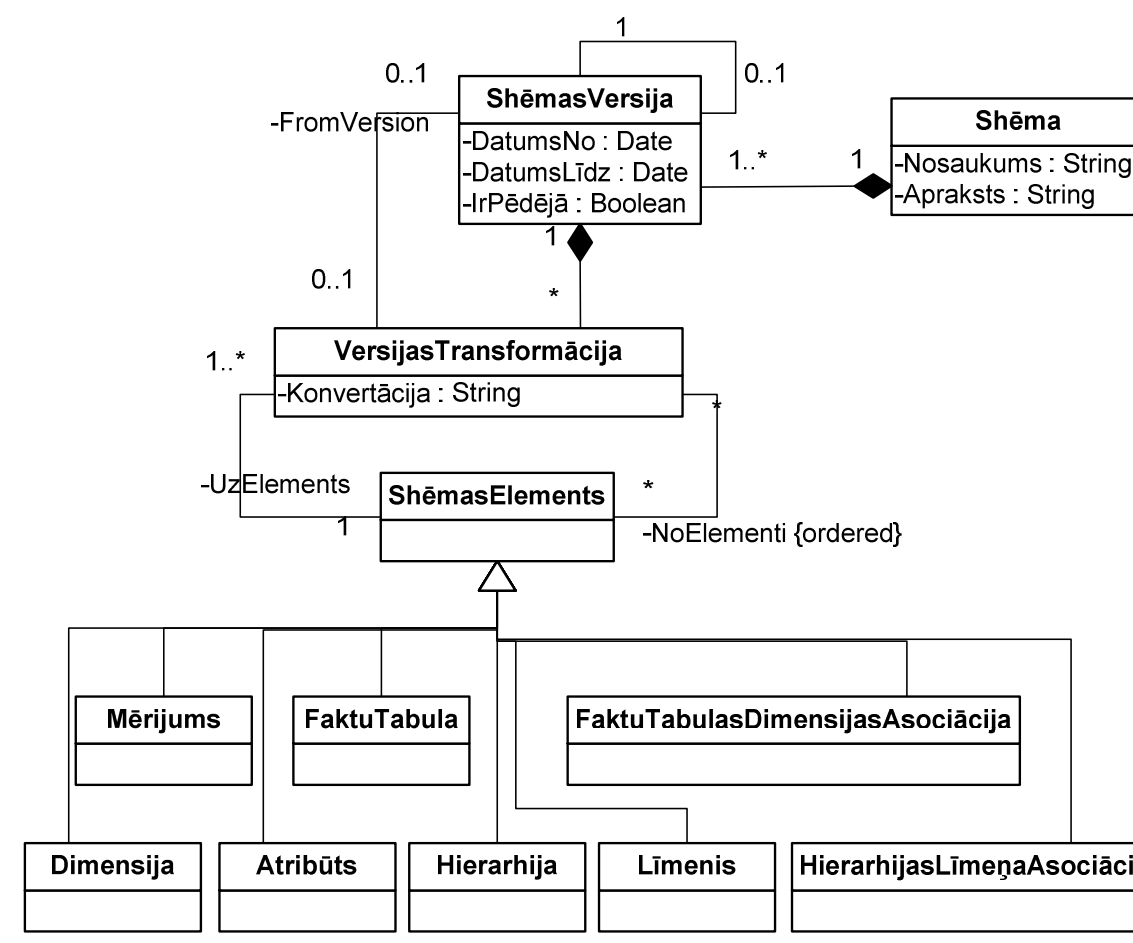
Rezultāti:

- izstrādāts procesu mērīšanas ietvars un veiktspējas indikatoru modelis,
- izstrādāta tīmekļa informācijas sistēmu adaptācijas arhitektūra, kas ietver procesu mērīšanas sistēmu.

Daudzversiju datu noliktavas modelis un metadati

- Izstrādāts daudzversiju datu noliktavas modelis, kas balstīts uz vairāku līmeņu metadatiem un to saistību
- Modeļa fragments un atbalstītās izmaiņas – nākošos slaidos.

Loģiskā līmeņa metadati – versiju metadati

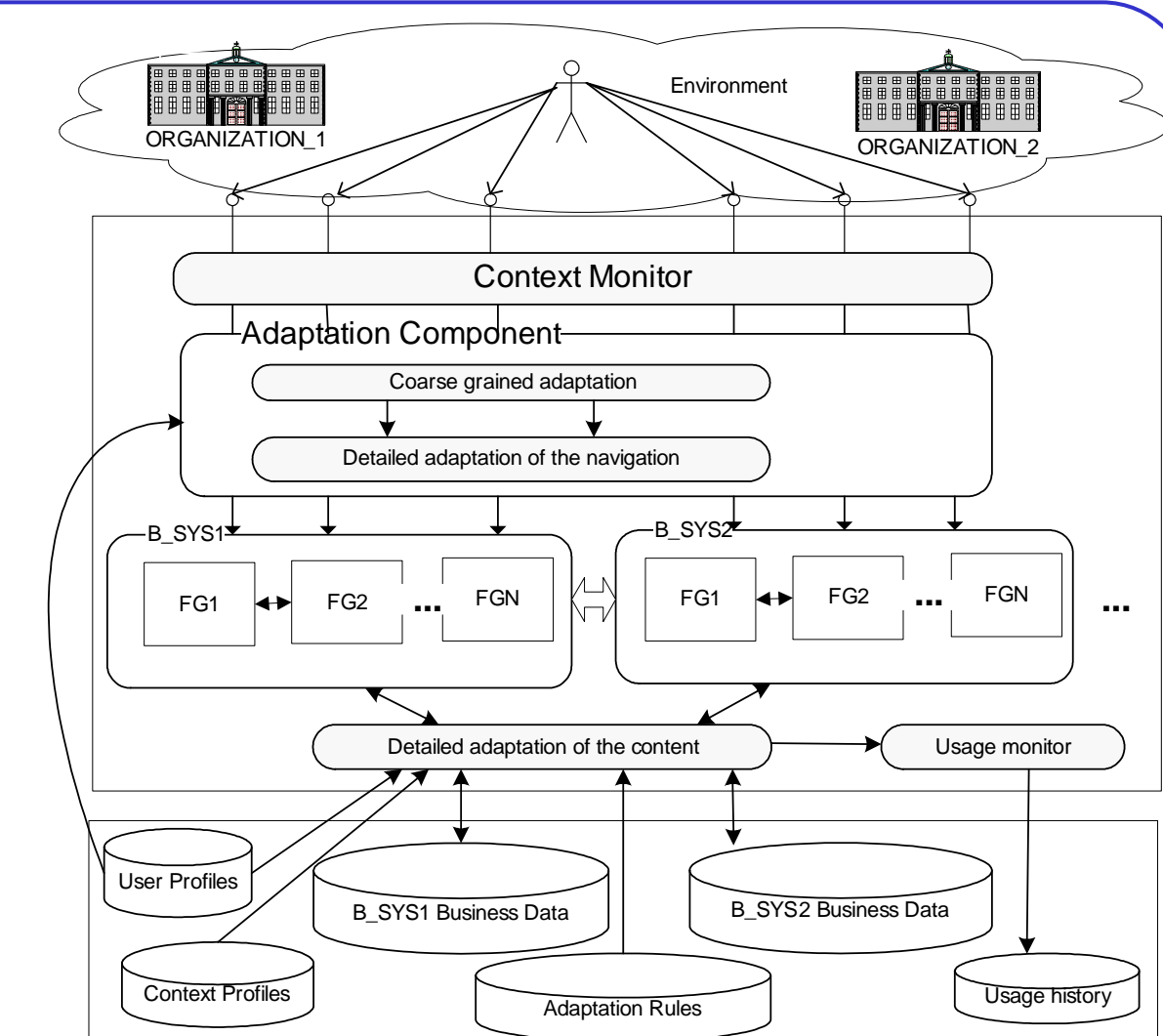


Piedāvātā sistēmas adaptācijas arhitektūra

- Balstās uz daudzu nomnieku vienas instances risinājumu programmatūras kā servisa pieejā,
- Papildināta ar divu līmeņu adaptācijas pieeju,

Komponenti:

- Konteksta monitors (Context monitor)
- Datu slānis (Data layer)
- Adaptācijas komponents (Adaptation component)
- Lietošanas monitors (Usage monitor)- procesu mērīšanai



Evolūcijas atbalsts – atbalstītās izmaiņas

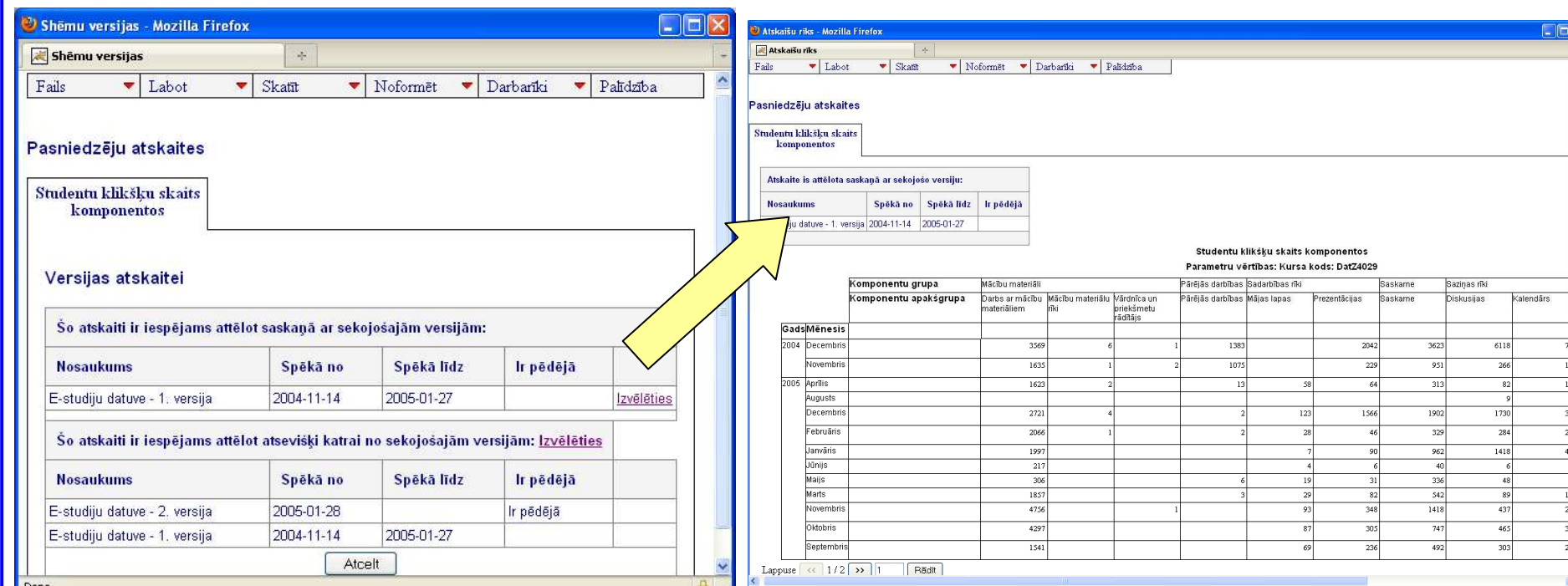
Fiziskās izmaiņas

- Darbojas ar datubāzes objektiem (tabulām, kolonām)
- Rezultātā tiek mainīti gan loģiskā, gan fiziskā līmeņa metadati
- Piemēri: jaunais dimensijas atribūts, dimensijas atribūta datu tipa maiņa, dimensijas atribūta dzēšana, jaunā fakta tabula u.c.

Loģiskās izmaiņas

- Maina shēmas metadatus
- Rezultātā tiek mainīti loģiska līmeņa metadati vai var tikt pievienotas atslēgas un atslēgu kolonas
- Piemēri: dimensijas piesaistīšana fakta tabulai, jaunā dimensijas hierarhija, hierarhijas dzēšana u.c.

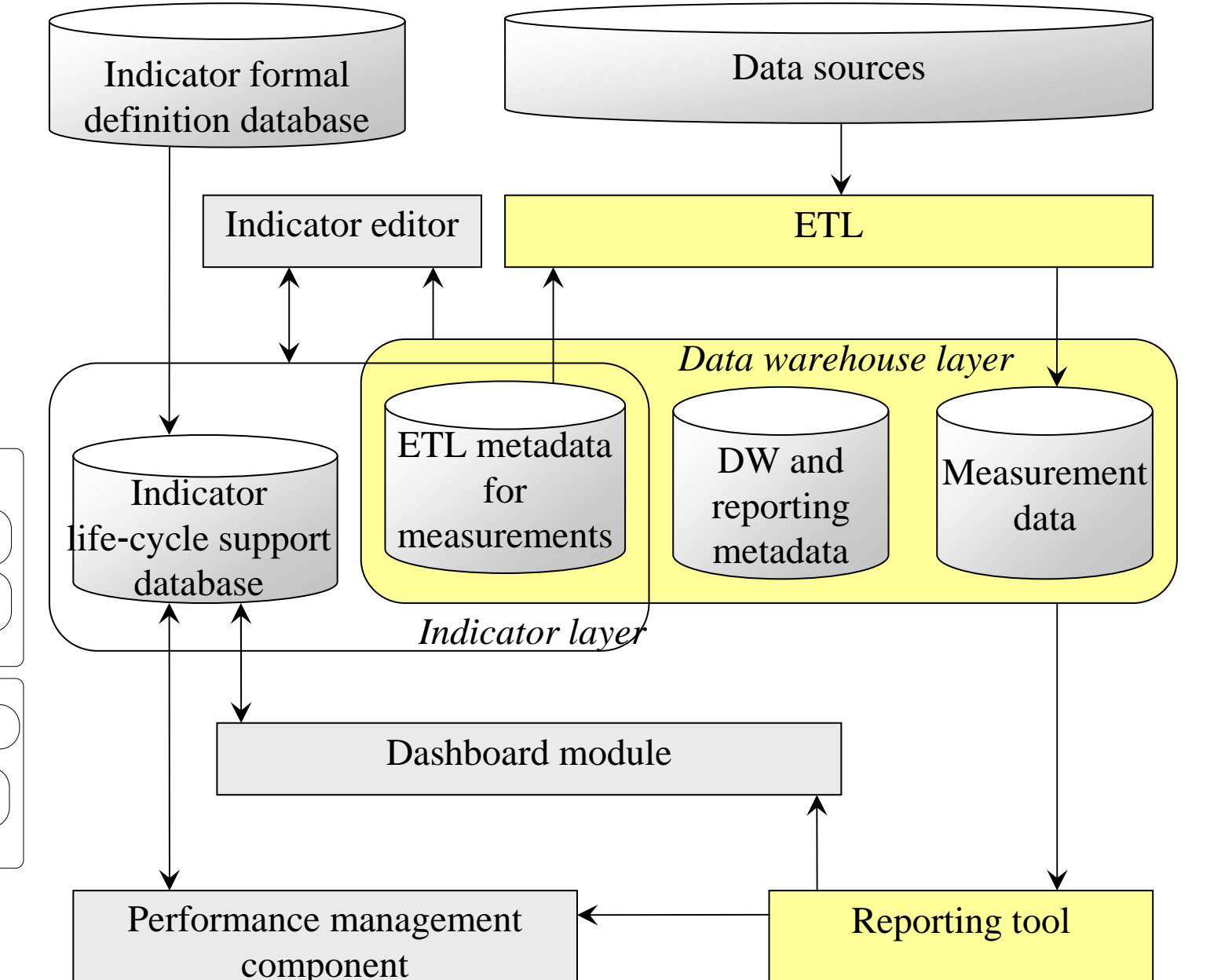
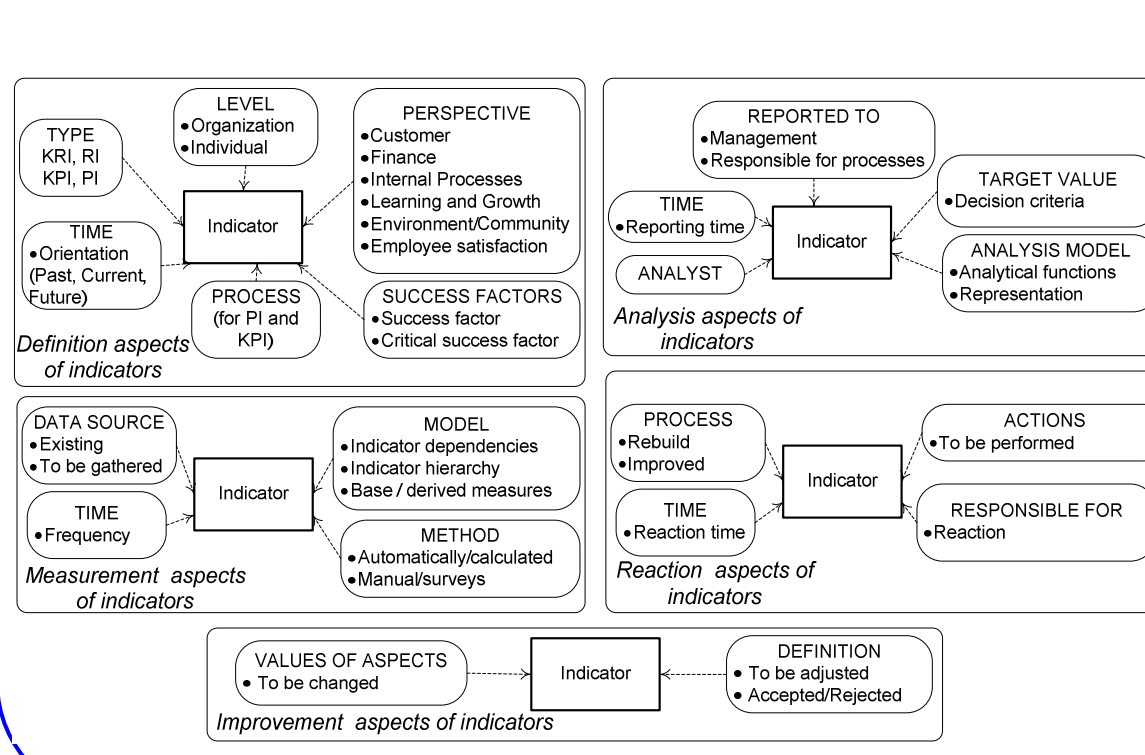
Atskaišu attēlošanas rīks



- Vienai un tai pašai atskaietei var būt dažādas versijas, ja datu noliktavas shēmā notikušas izmaiņas.
- Lietotājs var izvēlēties attēlot atskaiti saskaņā ar vienu konkrētu shēmas versiju vai attēlot atskaites datus atsevišķi divās atskaitēs katrai shēmas versijai.
- Atskaišu rīka prototipa aprobācija veikta LU e-studiju datuvē, realizējot e-studiju datu izmaiņu apstrādi rīka metadatos, versiju definēšanai izmantojot piedāvāto rīka metamodeli.

Indikatoru dzīves cikls un procesu mērīšanas sistēmas shēma

1. Indikatoru definēšana
2. Mērīšana
3. Analīze
4. Reakcija
5. Uzlabošana



Datu noliktavu personalizācija

Problēma:

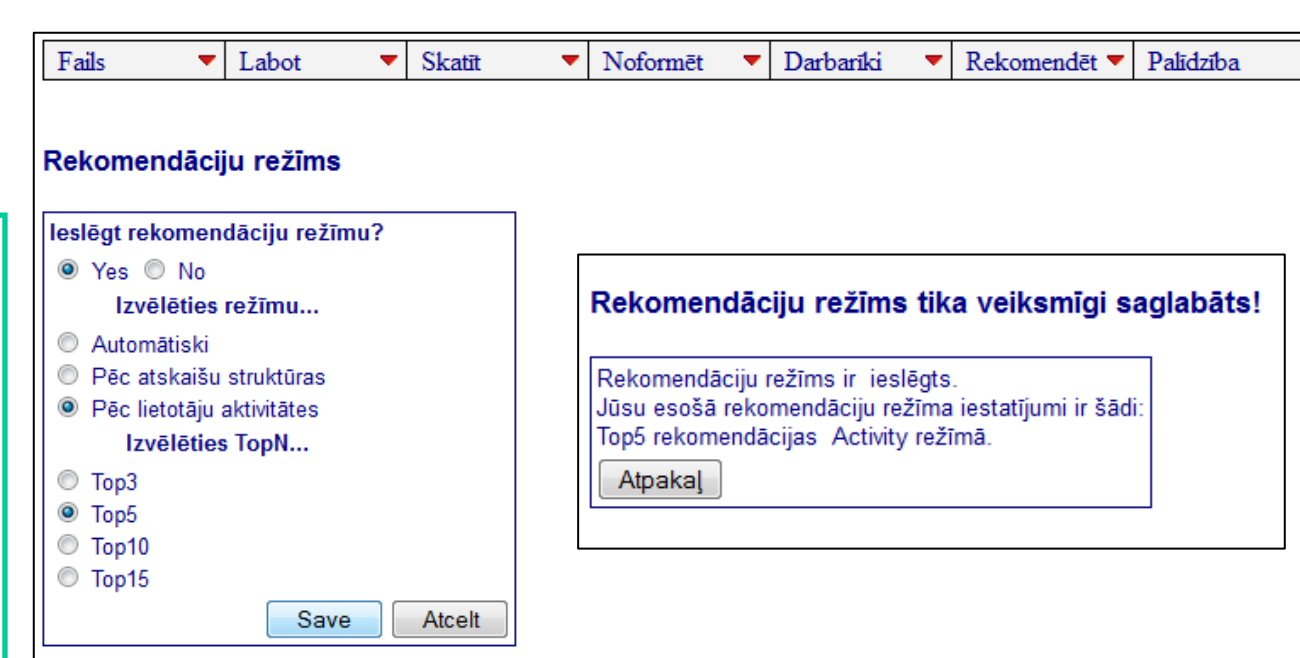
Liels datu apjoms datu noliktavās, liels skaits atskaišu, tiek veidotas jaunas, grūti atrast vajadzīgo informāciju

Pētījuma mērķis: izstrādāt rekomendāciju metodes atbilstošu datu noliktavas atskaišu ieteikšanai lietotājam

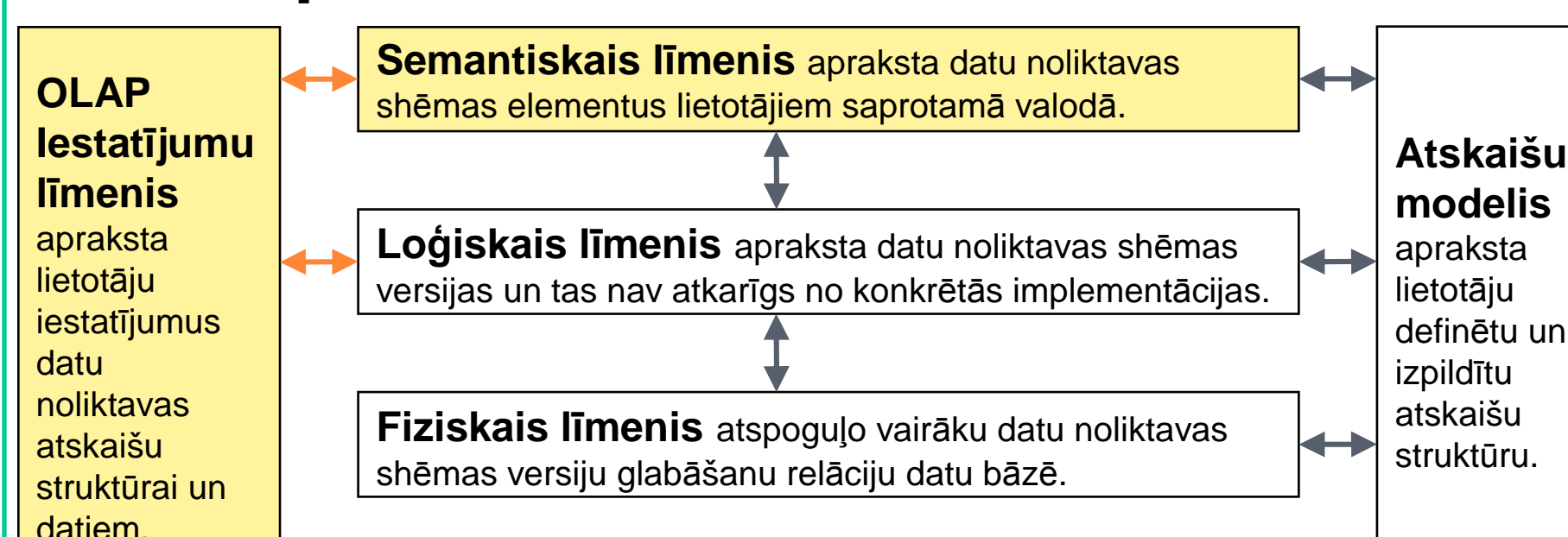
Rezultāti:

- Paplašināts metadatu modelis atskaišu rīkam
- Izstrādātas 2 metodes: Hot-start un Cold-start, kas piedāvā potenciāli interesantas atskaites, balstoties uz lietotāju netiešajiem iestatījumiem OLAP shēmas elementiem un pieļaujamām agregācfunkcijām.
- Izstrādāts atskaišu rīkam rekomendāciju komponents, kas ļauj izvēlēties dažādus rekomendāciju režīmus

Atskaišu rīka rekomendāciju režīmi un atskaišu rekomendācijas



Paplašinātais metadatu modelis



Nozīmīgākās publikācijas

Procesu datu noliktavas

- Niedritis, A., Niedrite L. The Adaptation of a Web Information System: a Perspective of Organizations. Proceedings of the 19th International Conference on Information Systems Development (ISD2010), Prague, Czech Republic, Springer, September 2011.
- Niedritis, A., Niedrite, L., Kozmina, N. Performance Measurement Framework with Formal Indicator Definitions. In: J.Grabis, M.Kirikova (Eds.), Perspectives in Business Informatics Research, LNBIP, Springer, 2011.
- J.Zuters, Near Real-Time Data Warehousing with Multi-Stage Trickle and Flip. In: J.Grabis, M.Kirikova (Eds.), Perspectives in Business Informatics Research, LNBIP, Springer, 2011.
- Niedritis, A., Niedrite, L., Zuters, J. Performance Measurement Framework with Indicator Life-cycle Support. In: A.Caplinskā, G.Dzemyda, A.Lupeikiene, O.Vasilecas (Eds.), Databases and Information Systems. 10th Int. Baltic Conf. on Databases and Information Systems. Local Proceedings, Materials of Doctoral Consortium. Vilnius, Žara, pp. 115-127, (2012).

Datu noliktavu evolūcija

- D.Solodovņikova, L. Niedrite. (2011) Evolution-Oriented User-Centric Data Warehouse. J.Pokorny, V.Repa, K.Richta et.al. (Eds), Proceedings of the 19th International Conference on Information Systems Development, Springer, pp. 721 – 734.
- Promocijas darbs Darja Solodovņikova, „Uz datu noliktavas shēmas evolūciju orientēts vaicājumu definēšanas un attēlošanas rīks”, aizstāvēts 2011. gada 9. februārī.
- Solodovņikova, D, Niedrite, L., and Niedritis, A. Query-Driven Method for Improvement of Data Warehouse Conceptual Model, Proceedings of the 21st International Conference on Information Systems Development (ISD2012), Prato, Italy, Springer (in print)

Datu noliktavu personalizācija

- N. Kozmina, D. Solodovņikova. Towards Introducing User Preferences in OLAP Reporting Tool. In: Niedrite L, et al. (eds.) BIR 2011 Workshops. Springer, Heidelberg, LNBIP 106:209-222
- N. Kozmina, D. Solodovņikova. On Implicitly Discovered OLAP Schema-Specific Preferences in Reporting. Tool In: Proceedings of BIR'11, Scientific Journal or Riga Technical University, Computer Science: Applied Computer Systems, 46:35-42.
- N. Kozmina, L. Niedrite. OLAP Personalization with User-describing Profiles. In: Forbrig P, Günther H (eds.) BIR 2010. Springer, Heidelberg, LNBIP, 64:188-202
- N. Kozmina, L. Niedrite. Research Directions in OLAP Personalization. In: Proceedings of ISD'10, Springer Science+Business Media, pp 345-356