



Procesu datu noliktavas

Laila Niedrīte
LU, 01.06.2010



Eiropas Sociālā fonda projekts
"Datorzinātnes pielietojumi un tās saiknes ar kvantu fiziku"
Nr.2009/0216/1DP/1.1.1.2.0/09/APIA/VIAA/044

Prezentācijas plāns

- ❖ **Datu noliktavu pētniecības grupas tēmas**
- ❖ **Procesu datu datu noliktavas –**
 - pētījumu virzieni pasaulē
- ❖ **Mērķu orientētā procesu mērīšana datu noliktavu vidē –**
 - Metodes
 - Modeļi
 - Pielietojumi

Datu noliktavu pētniecības grupas tēmas

❖ **Procesu datu noliktavas**

- Autori: L.Niedrīte, A.Niedrītis, M.Treimanis, D.Solodovņikova
- Virzieni:
 - » Mērķu orientēta metode procesu datu noliktavas modeļu izveidei
 - » Biznesa procesu mērīšanas iniciatīvas LU (e-studijas, reģistrācijas procesi)

❖ **Datu noliktavu evolūcija**

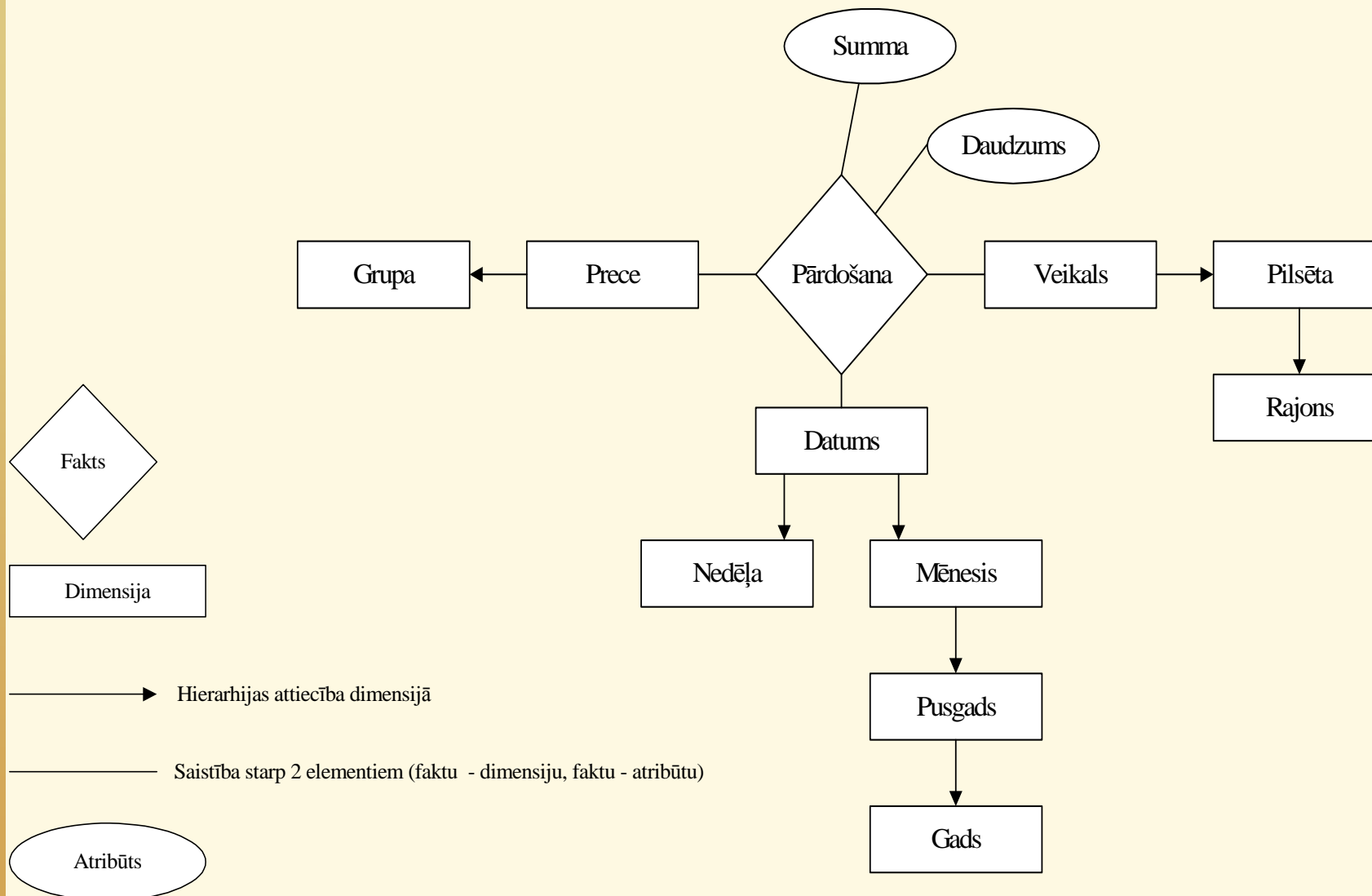
- Autori: D.Solodovņikova, L.Niedrīte
- Virzieni:
 - » Shēmu versijas datu noliktavu evolūcijas atbalstam
 - » Metamodeļa bēzēts atskaišu rīka izstrāde daudzu versiju datu noliktavai

❖ **Datu noliktavu personalizācija**

- Autori: N.Kozmina, L.Niedrīte
- Virzieni:
 - » Lietotāju iestatījumu bāzēta metode datu noliktavas personalizācijai

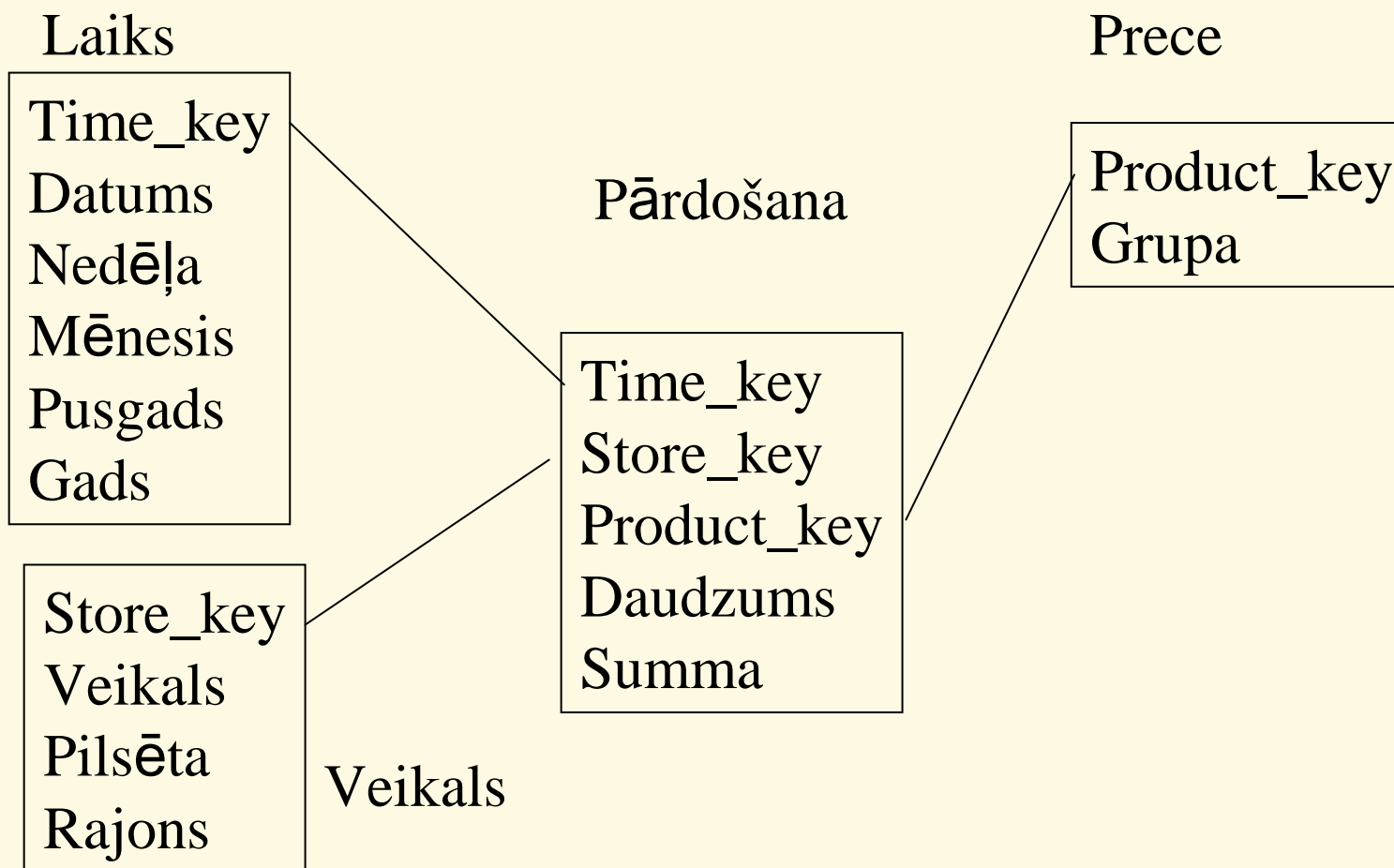
Datu noliktavas konceptuālais modelis (piemērs) ME/R notācijā

Procesu datu noliktavas



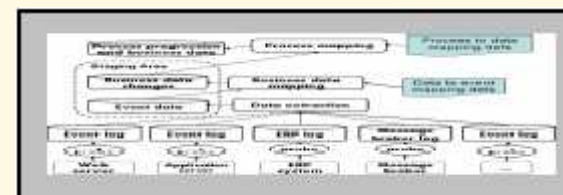
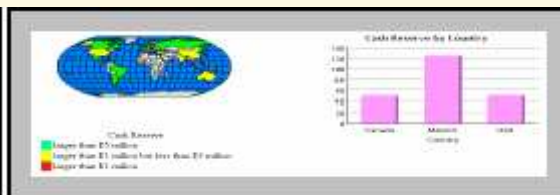
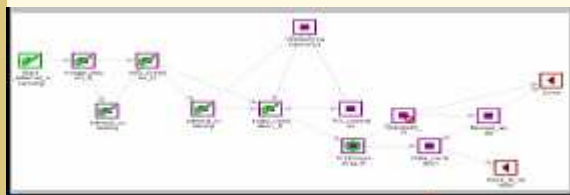
Piemērs – Zvaigznes shēma

Procesu datu noliktavas

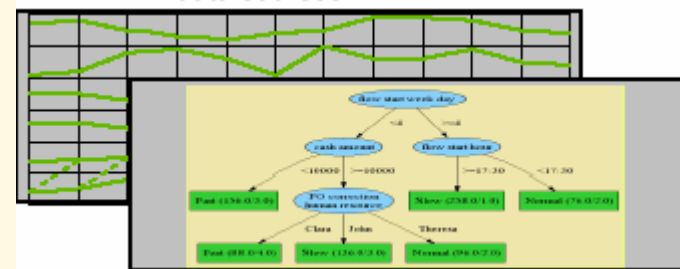
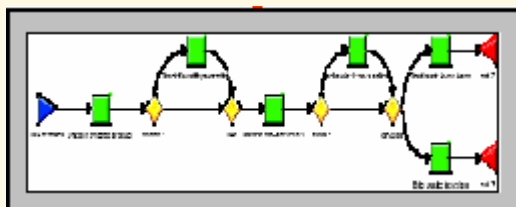


Darījumprocesu analīze - saistītās pētniecības jomas

Procesu datu noliktavas



- Procesu modelēšana
- Metriku noteikšana
- Datu noliktavu modeļi, ETL procesi
- Datu analīze, datizrace, procesu uzvedības modeļu atrašana
- Darījumprocesu pārbūve



Metriku veidu piemēri

❖ **Procesu metrikas**

- izpildes laiks, izpildes ilgums, izpildes ceļš

❖ **Resursu metrikas**

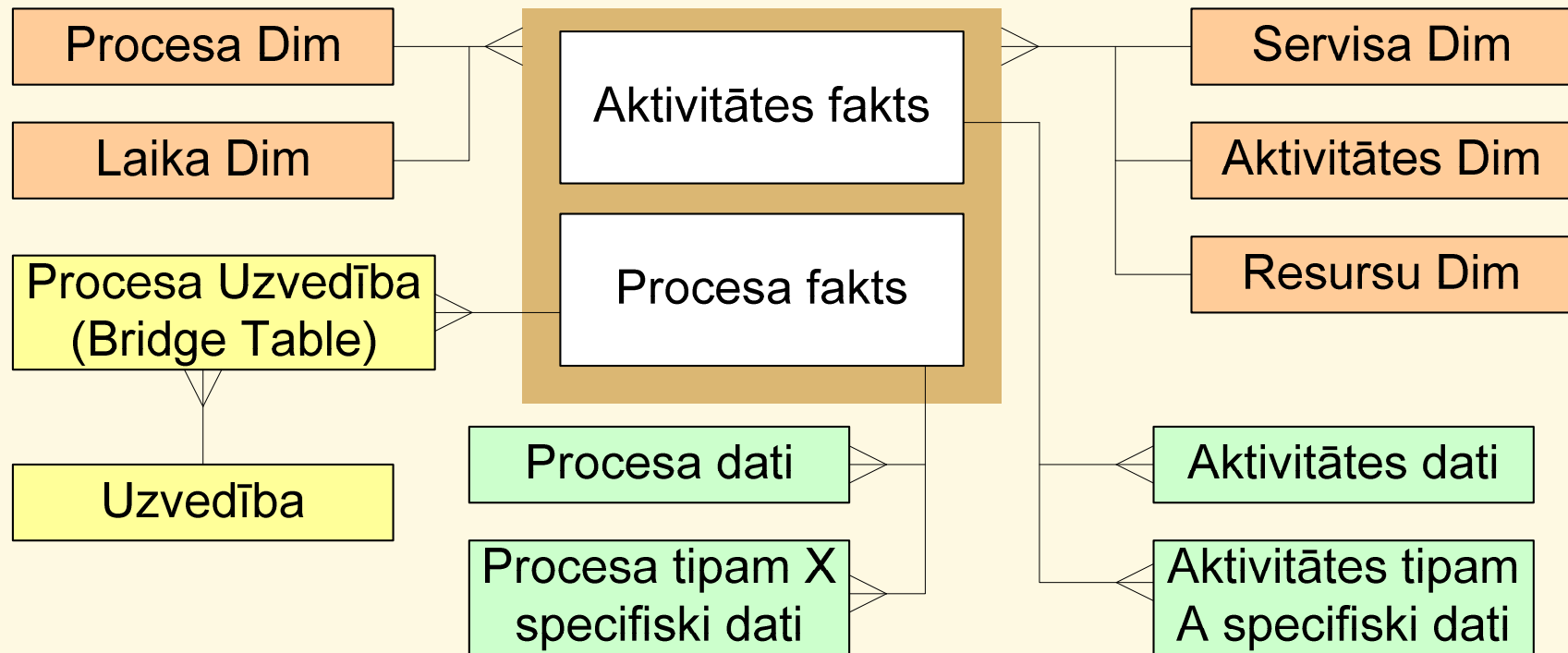
- Cilvēku resursi, procesu veicēji,
- automatizētie procesi – IS izmantošana procesu soļu izpildē

❖ **Darījuma (biznesa) metrikas:**

- darījuma summa, preču skaits u.c.
- darījuma metriku korelācija ar procesa metrikām – procesa izpildes efektivitāte, kvalitāte

Datu modelis

Procesu datu noliktavas



A. Bonifati, F. Casati, U. Dayal, M.-C. Shan, "Warehousing Workflow Data: Challenges and Opportunities", VLDB 2001.

Procesu izpildes tipiskie modeļi (Patterns)

❖ Uzvedība (dimensija) un tās atribūti (piemēri):

- a1: Procesa izpilde ir starp 10% ātrākajām instancēm
- a2: Procesa aktivitāšu skaits < vidējo visiem procesiem
- a3: Procesa izpilde ir starp 10% lēnākajām instancēm
- a4: Procesa aktivitāšu skaits > vidējo visiem procesiem

– Atribūtu vērtības ir 0 vai 1

– Iespējamie modeļi, kas varētu interesēt (piemēri):

M1: 1 0 0 0

M2: 1 1 0 0

M3: 1 0 0 1

....

❖ Procesa uzvedība (fakts):

- katram jaunam procesa izpildes faktam, var noskaidrot, kādam “uzvedības modelim” šī jaunā instance atbilst.
- Sākotnēji dati šai faktu tabulai tiek izrēķināti no citiem faktiem par procesu izpildi

e-Universitātes iniciatīva – risinājumi

❖ Biznesa procesu pārbūve un mērīšana

- e-universitātes iniciatīvas ietvaros 2 procesi tika:
 - mērīti,
 - pārbūvēti
 - jaunas formālas procedūras
 - IS pielāgošana
 - leviesti
 - Kontrolēti un atkārtoti mērīti
 - » Medvedis, I., Niedrite, L., Niedritis, A., Treimanis, M., & Voitkans, A. (2008) **e-University Initiative: Approach, Solutions and Case Studies** , DBIS2008.
 - » Medvedis, I., Niedrite, L., Niedritis, A., Treimanis, M., & Voitkans, A. (2008) **Approach and Technical Solutions for e-University Initiative**, In Haav, H.M., & Kalja, A. (eds). *Frontiers in Artificial Intelligence and Applications*, IOS Press, pp. 155-166

Procesu kontrole (Process monitoring)

Procesu datu noliktavas

- ❖ **Izmanto reālā laikā datus no procesu izpildes sistēmas**
 - piemēram, darba plūsmas sistēmas, žurnālu faili,
- ❖ **Procesu dati tiek vākti par pašu procesu izpildi:**
 - procesa instanču skaits, kas dotajā momentā izpildās,
 - procesa izpildē iesaistītais lietotāju skaits,
 - vidējais katra procesa tipa izpildes ilgums dotajā laika periodā;
- ❖ **Dati tiek attēloti reālā laikā,**
- ❖ **Dati parasti netiek uzkrāti papildus citās datu bāzēs**

Procesu mērišana (Process Measurement)

- ❖ **Kombinē datus par procesu izpildi ar darījumu datiem, indikatoru piemēri:**
 - studentu skaits, kas pierēģistrējas uz semestri Internetā;
 - studentu skaits, kurus pierēģistrēja dažādu fakultāšu lietveži;
- ❖ **Darījumu dati var tikt vākti procesa izpildes laikā vai tikt paņemti, piemēram, no jau eksistējošas datu noliktavas.**

Procesu datu noliktava

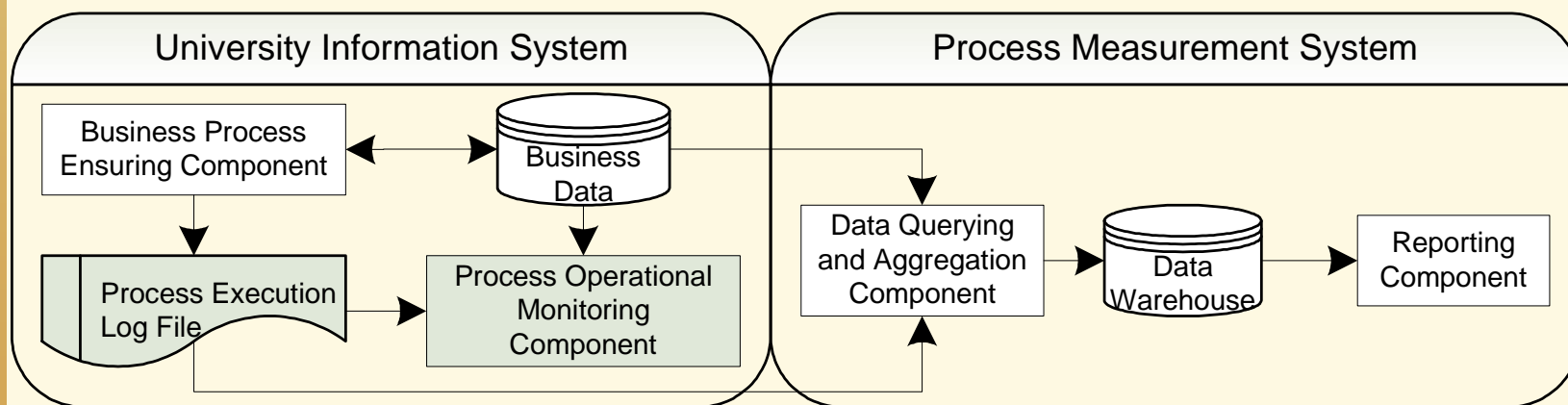
❖ Procesu kontroles komponents LUIS

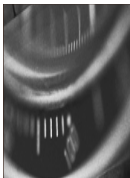
- Procesu izpildes žurnālfails
- Procesu operatīvās izpildes modulis

❖ *Procesu mērīšanas sistēma*

- Balstās uz datu noliktavas modeļiem

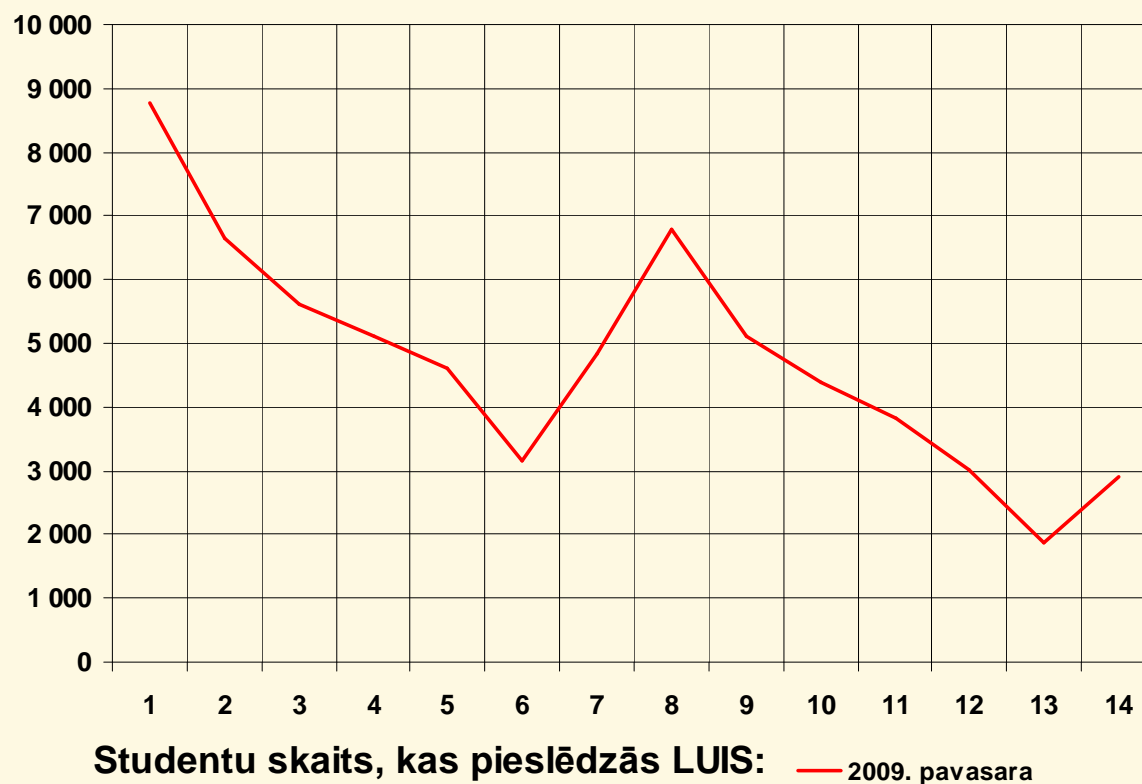
Procesu datu noliktavas





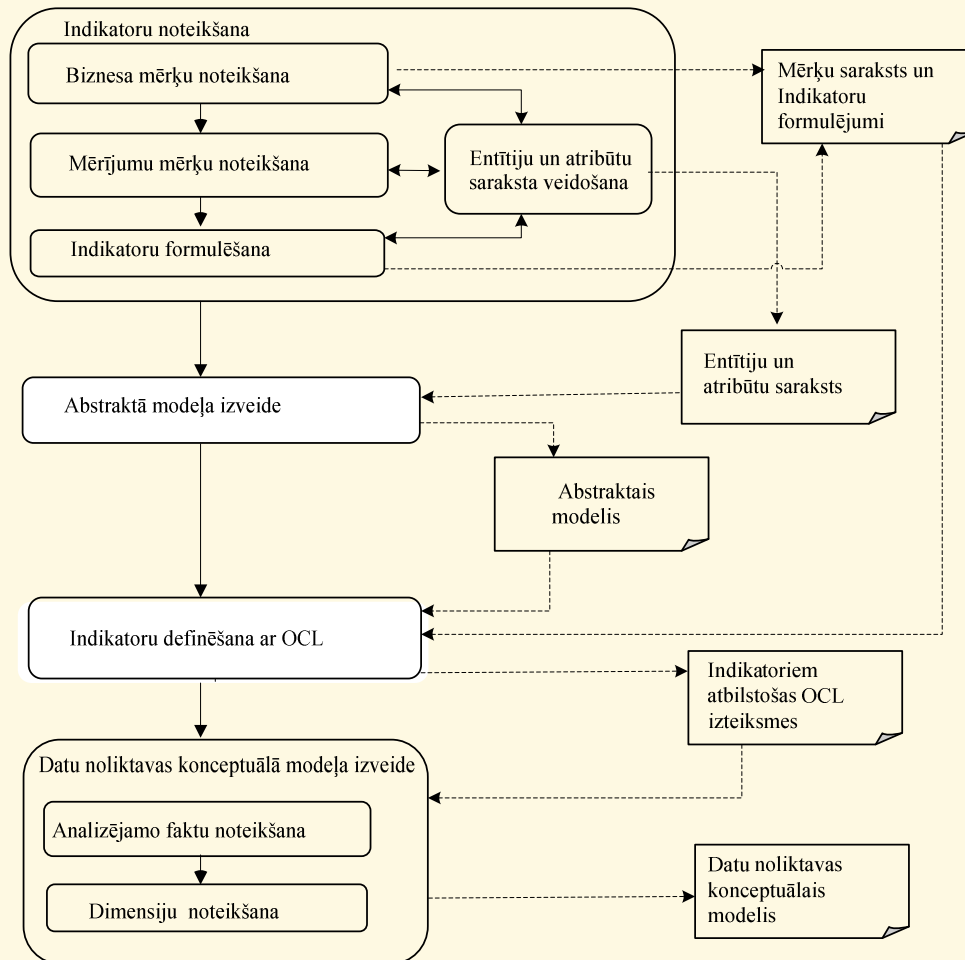
- ❖ **Reģistrēšanās uz semestri un reģistrēšanās uz kursiem**
- ❖ **Studenti, kas pieslēdzās LUIS pa dienām**

Procesu datu noliktavas



Mērķu vadīta metode datu noliktavas modeļa izveidei

Procesu datu noliktavas

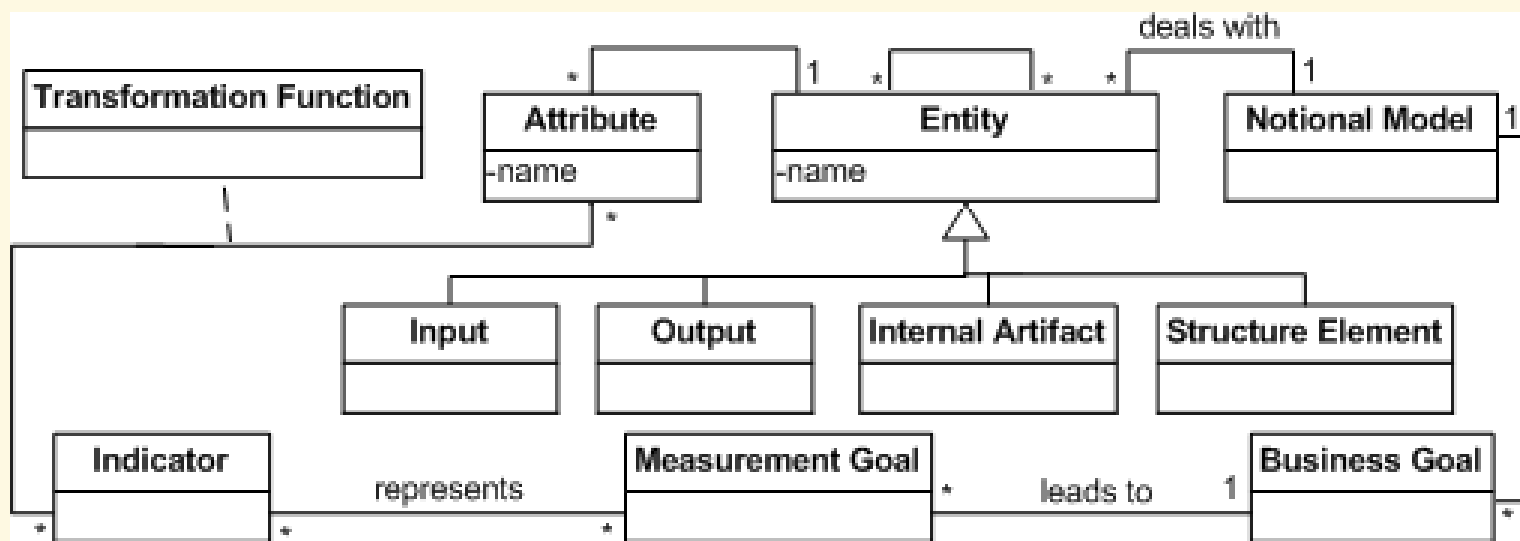


- ❖ **Indikatoru noteikšana.** Sākumā nosaka biznesa mērķus, tad mērījumu mērķus, tad indikatorus, pielietojot GQ(I)M metodi.
- ❖ **Modeļa izveide.**
- ❖ **Indikatoru definēšana ar OCL.**
- ❖ **Datu noliktavas konceptuālā modeļa izveide.**

Niedrite L., Solodovnikova D., Treimanis M., Niedritis A., „The Development Method for Process-Oriented Data Warehouse”, In: WSEAS Transactions on Computer Research, Issue 2, Vol. 2, February 2007.

Indikatoru noteikšana

- *Goal-Question-Indicator-Measure* metode
- Metamodeļa elementi
 - » 4 types of Entities, Attributes, Notional Model
 - » Business Goal, Measurement Goal, Indicator
 - » Transformation Function



Indikatoru noskaidrošana

Procesu datu noliktavas

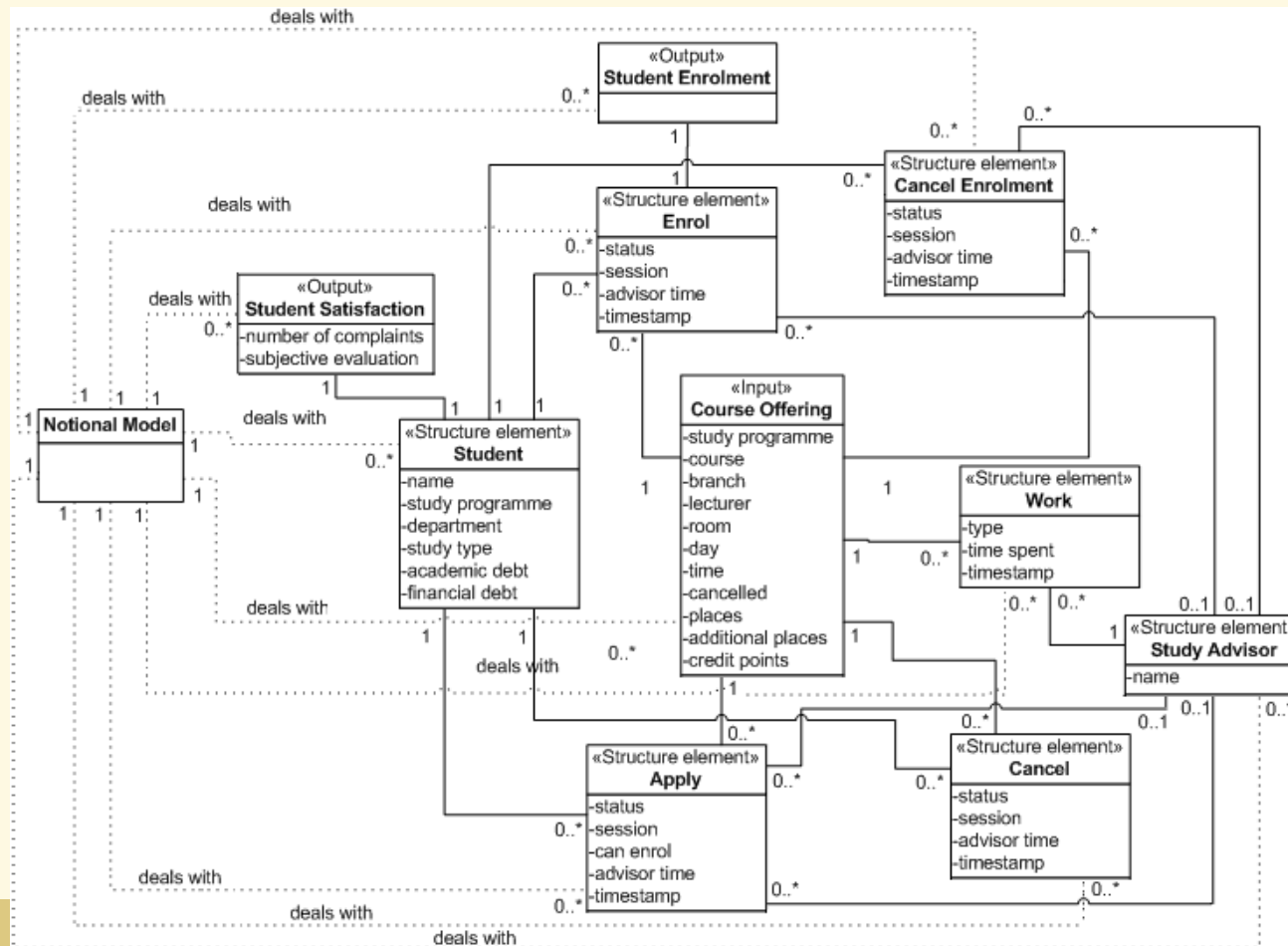
- ❖ Pielieto GQ(I)M metodi: noskaidro mērķus, jautājumus, (indikatorus) mērījumus
- ❖ Piemērs – process “Studentu reģistrācija kursiem”
 - Mērķis: Uzlabot reģistrācijas procesu no studenta skatu punkta
 - Jautājumi un indikatori katram mērķim (piemēri)

Q3	Kāda ir kursu atcelšanas kārtība, ja piesakās mazs studentu skaits?	I8	Atcelto kursu skaits
Q4	Cik studentiem nebija pieejams reģistrācijas process izmantojot Internetu un kāpēc?	I10	Studentu skaits ar finansu parādiem
		I11	Studentu skaits ar akadēmiskajiem parādiem

Modeļa izveide

❖ Instanču diagramma atbilstoši indikatoru definēšanas metamodelim

- Elementi tiek identificēti GQ(IM) metodes lietošanas laikā intervijās



Indikatoru definēšana ar OCL

- ❖ Indikatori tiek formulēti ar OCL vaicājuma operācijām, izmantojot klases, atribūtus un asociācijas no iepriekš izveidotā modeļa.

Piemērs

I11 “Studentu skaits ar akadēmiskajiem parādiem”

```
Student->select(academic debt='Yes')-  
>size()
```

Datu noliktavas konceptuālā modeļa izveide – faktu un dimensiju noteikšana no OCL

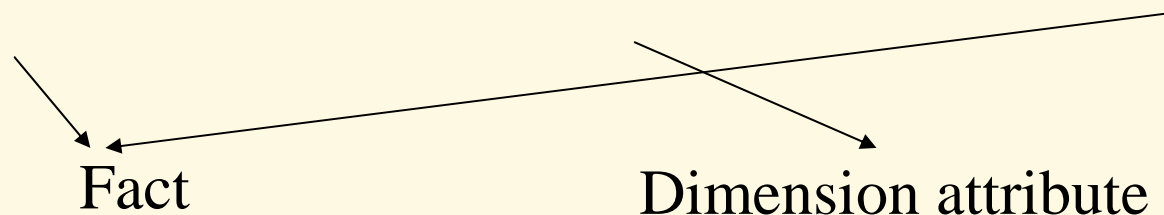
❖ Fakti

- If a result of an operation is numerical, for example, `sum()`, `size()`, `round()`, multiplication, division, we consider such values as potential facts

❖ Dimensiju atribūti

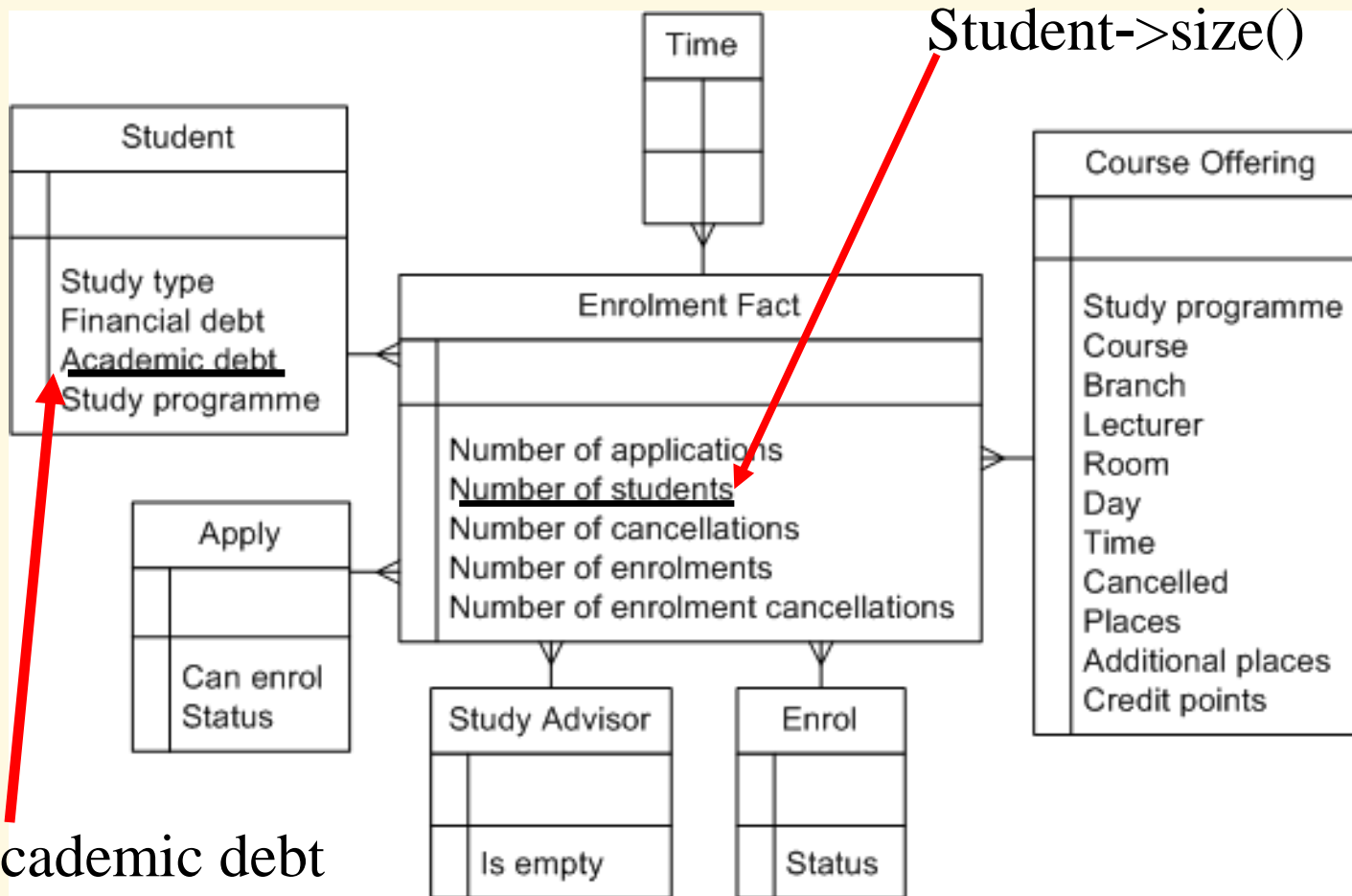
- classes, which appear in *context* clause of OCL query operations excluding the class Notional Model
- other dimension attributes are derived from class attributes used in *select* clause of OCL query operations. These attributes are grouped into dimensions corresponding to classes that contain these attributes

❖ Student->select(academic debt='Yes')->size()



Datu noliktavas konceptuālā modeļa izveide

Procesu datu noliktavas



select(academic debt = 'Yes')

Student->size()



Sistēmu lietošanas paradumu noteikšana

Turpinājums 2009. gadā

Ko un Kā mērit? – Esošās pieejas

❖ Informācijas sistēmas mērīšana

- 6 neatkarīgas dimensijas
 - » sistēmas, informācijas un servisa kvalitāte, sistēmas lietošana, lietotāju apmierinātība, ieguvumi
- “Mērāmo parametru izvēle bieži ir funkcija no pētījuma mērķa un organizācijas konteksta”
- “Pēc iespējas jāsamazina dažādo lietoto mērījumu skaits”

❖ Izvēle un skaits – pirms uzsākt mērīšanu

DeLone, W.H., & McLean, E.R. (2002) Information Systems Success Revisited. In Proceedings of the Hawaii International Conference on System Sciences Hawaii

Ko un Kā mērit? - Divu līmeņu mērīšana

❖ **Veids un Objekts (Kā? un Ko?)**

- 1. augsta līmeņa mērīšana
 - sistēma kā mērījumu objekts
- 2. datalizēta līmeņa mērīšana
 - Biznesa procesi (funkcijas, kas to atbalsta).

❖ **Kad un kāpēc? (1. līmenim)**

- Jauns pakalpojums
- Esošo sistēmu uzlabošana, paplašināšana
- Kritiskās sistēmas, prioritātes

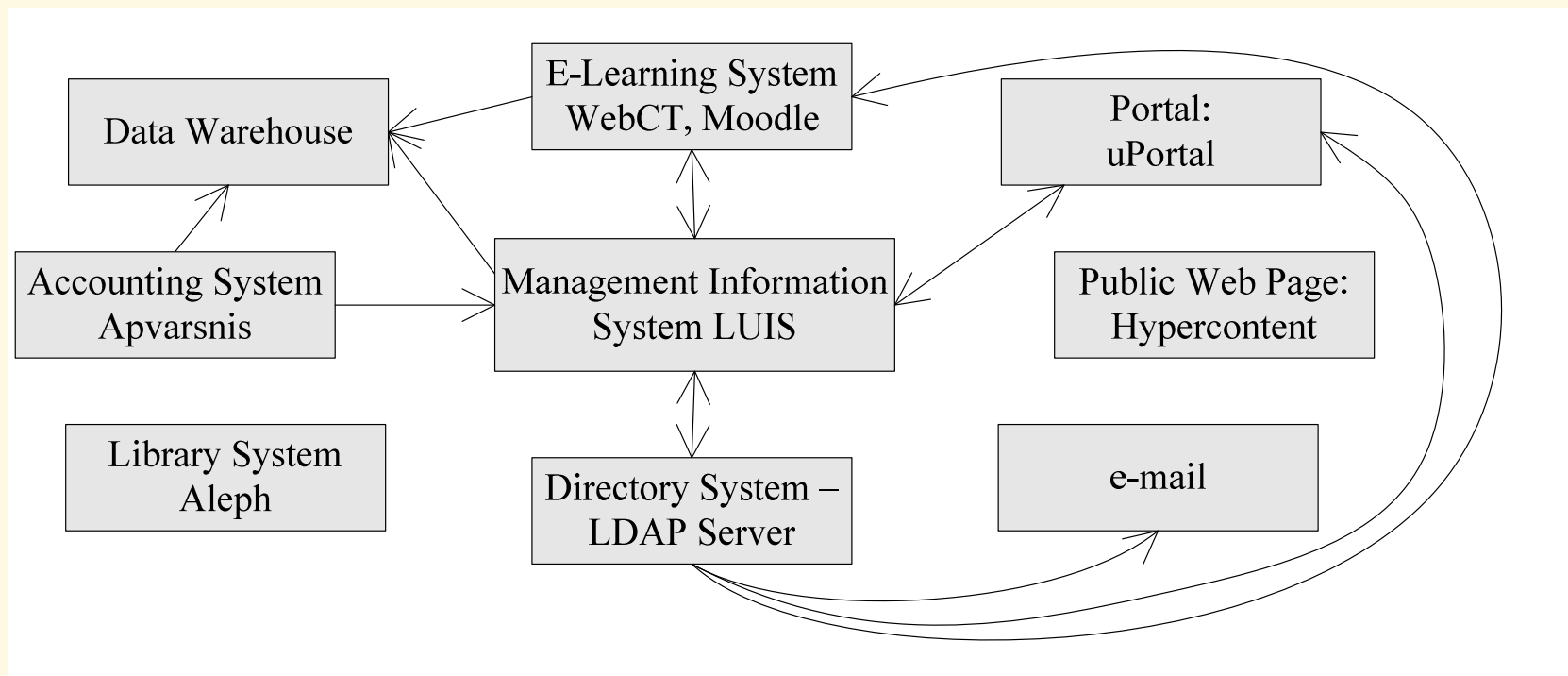
Sistēmu līmeņa mērišana

- ❖ **Sistēmas darbības efektivitātes mērišanai**
- ❖ **Problēma – vai var salīdzināt mērijumu rezultātus?**
- ❖ **Īpašības, kas ietekmē un iepriekš jāanalizē:**
 - Sistēmas lietošanas obligātums
 - Pieejas kontroles veids un ierobežojumu eksistence

Universitātes sistēmas

❖ deviņas galvenās informatīvās sistēmas:

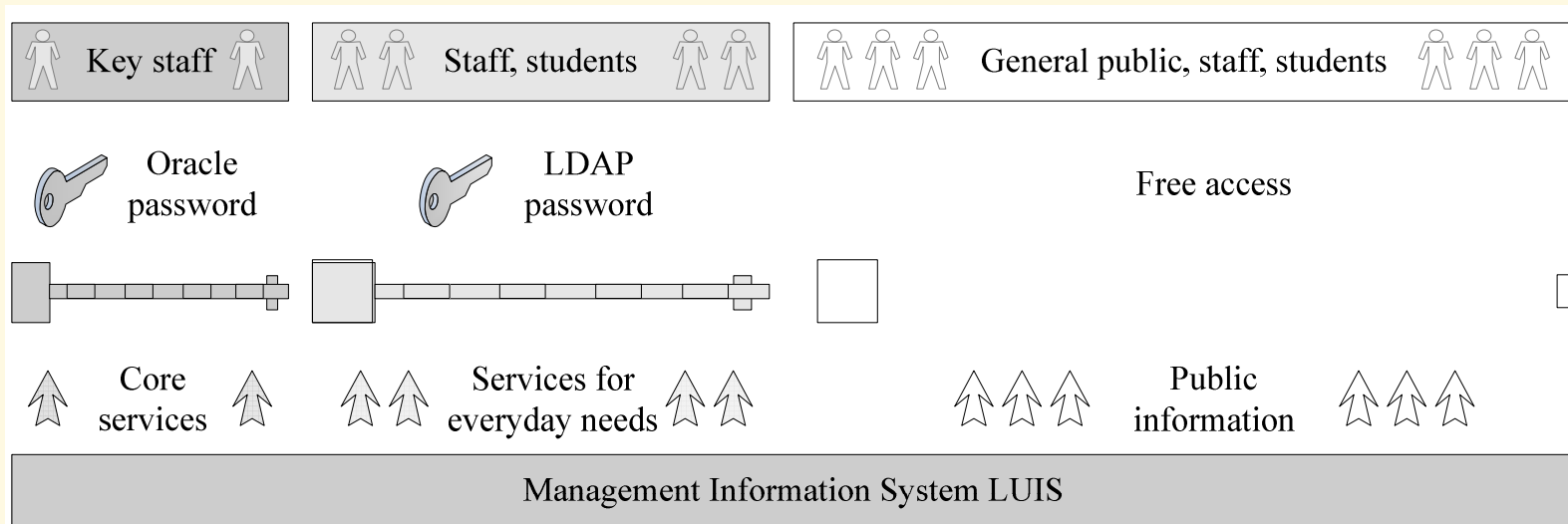
- Atšķirīgs lietošanas mērķis
- ar dažādu savstarpējo atkarību



LUIS izmantošanas trīs atšķirīgas grupas

- Darbinieki, kas nodrošina pamatservisus – vairāk kā 1000 definēti Oracle DB lietotāji.
- Studentu un darbinieku pieeja personīgajai informācijai - funkcijas ar LDAP pieeju (50000 lietotāji)
- Publiskas funkcijas bez pieejas kontroles.

Procesu datu noliktavas



Sistēmu īpašību analīze (1)

❖ **LUIS – 3 daļas :**

- Kodols (autorizētā daļa) – obligāts (nepieciešams darba funkciju izpildei, nav alternatīvas)
- Pašapkalpošanās daļa (LDAP daļa) – nav stingri obligāts, jo pieļauj zināmu alternatīvu pastāvēšanu
 - Aiziet pie lietveža un palūgt, lai izdara (reģistrēt uz semestri)
 - Daļa no funkcijām nav obligātas – savu datu apskate
- Publiskā daļa – pilnībā brīvprātīgs

Sistēmu īpašību analīze (2)

- ❖ **Moodle - ja e-kursam ir tikai atbalsta funkcija, studentam pastāv iespēja to nelietot**
 - lieto LDAP autorizāciju
- ❖ **Portālam ir 2 daļas :**
 - 1) nepieciešama LDAP autorizācija,
 - 2) publiska – visa lietošana neobligāta
- ❖ **Salīdzināšanai izvēlētas :**
 - LUIS pašapkalpošanās funkciju daļa, Moodle un portāla autorizētā daļa
 - līdzīgs autorizācijas mehānisms (LDAP),
 - visos gadījumos parādās zināma izvēles brīvība

Sistēmu lietošana – mērījumu rezultāti

❖ Lietotāji - Moodle 6861, LUIS, Portal – 20478

- Unikālie lietotāji mēnesī (4 mēnešu mērīšanas periods),
 - Moodle: Min 3150 (46%), Max 5256 (77%)
 - Portal: Min 5702 (28%), Max 6227 (30%)
 - LUIS: Min 12083 (59%), Max 14976 (73%)
- % no potenciālo lietotāju skaita

❖ Salīdzināt nevar “absolūtos rādītājus”



Tālākie pētījumu virzieni

Turpinājums 2010. gadā

Pieejas elastīgu un izmaksu efektīvu tīmekļa informācijas sistēmu izstrādei

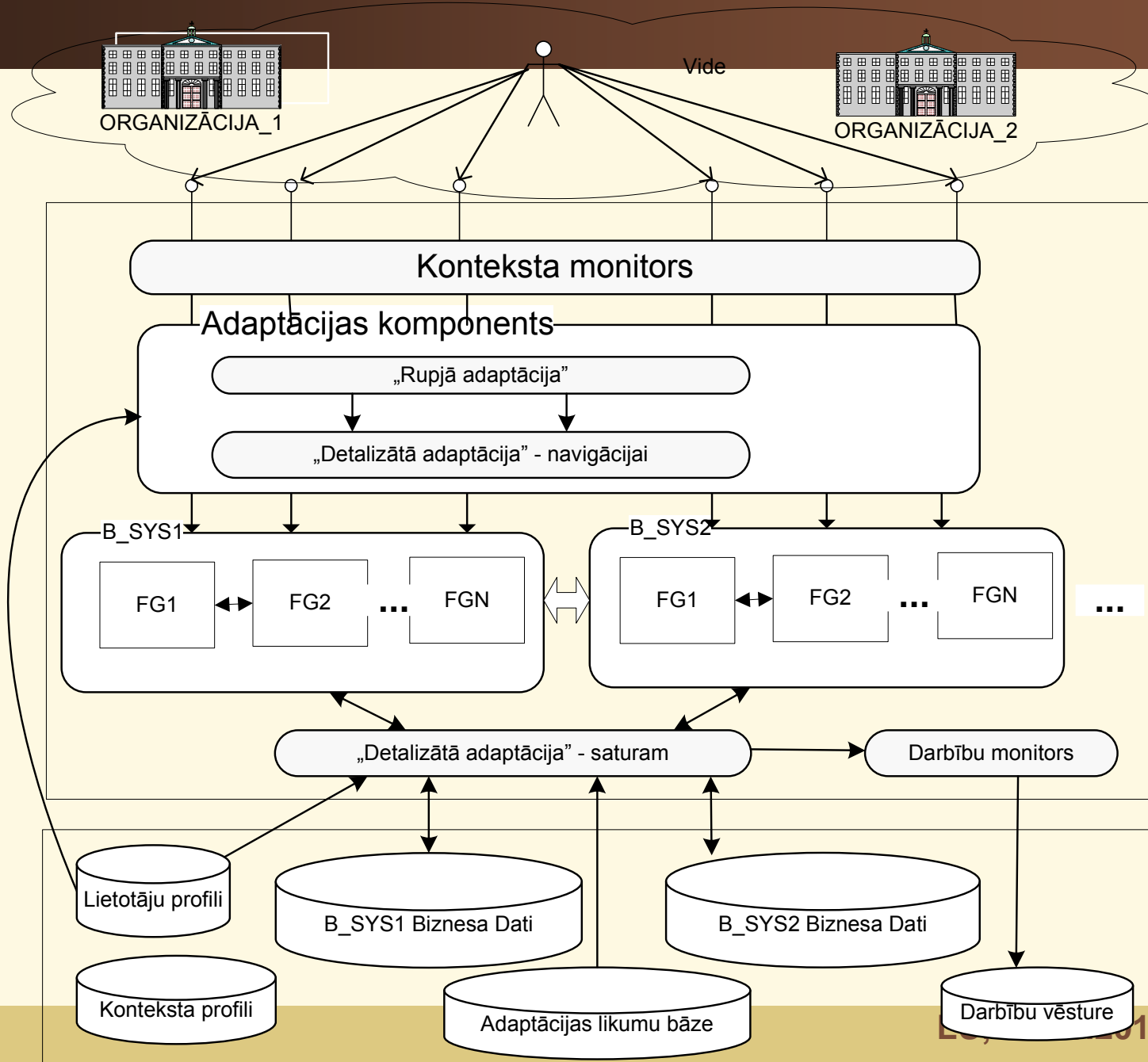
- ❖ **Software as a service (SaaS):** “software deployed as a hosted service and accessed over the Internet”
- ❖ **Piedāvā kā pakalpojumu dažāda lieluma organizācijām, lai nodrošinātu šīm organizācijām līdzīgus biznesa procesus:**
 - piemēram, klientu attiecību pārvaldība
- ❖ **Organizācijas piesakās lietot un maksā par lietošanas faktu;**
- ❖ **Pakalpojuma sniedzējs izvieto SaaS programmatūru uz saviem serveriem, uztur gan programmatūru, gan infrastruktūru.**
- ❖ **Dažādi realizācijas veidi, piemēram :**
 - Pielāgotas instances katram klientam uz pakalpojuma piegādātāja servera
 - Konfigurējams daudzu nomnieku (multi-tenant) vienas instances risinājums uz pakalpojuma piegādātāja servera.

Pieejas elastīgu un izmaksu efektīvu tīmekļa informācijas sistēmu izstrādei

- ❖ Lai efektīvi risinātu problēmas, kas saistītas ar konfigurēšanu un daudzu nomnieku pārvaldību, SaaS lietotnēm nepieciešamas piemērotas arhitektūras
- ❖ Šīm arhitektūrām jānodrošina arī pielāgošanu daudzo nomnieku vajadzībām
- ❖ Piedāvājums šādai arhitektūrai, kurā izmantoti arī procesu kontroles komponenti, aprakstīts ISD2010 iesniegtā rakstā

Procesu datu noliktavas

Arhitektūra



Paldies par uzmanību!