

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Jernej Počkaj

**Vzpostavitev bančnega področnega podatkovnega
skladišča za vrednotenje kreditne izpostavljenosti**

DIPLOMSKO DELO

VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE
STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: doc. dr. Matjaž Kukar

Ljubljana 2014

Rezultati diplomskega dela so intelektualna lastnina avtorja. Za objavlanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Tematika naloge:

Predstavite koncepte bančnega področnega podatkovnega skladišča in jih ilustrirajte na primeru implementacije vrednotenja kreditne izpostavljenosti. Opredelite strateške prednosti področnega podatkovnega skladišča in specifične implementacije v bančništvu. Opišite konkretno implementacijo, odzive končnih uporabnikov, ter celovito ovrednotite potencialne poslovne učinke vpeljave področnega podatkovnega skladišča v prakso.

IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Jernej Počkaj, z vpisno številko 63040363, sem avtor diplomskega dela z naslovom:

Vzpostavitev bančnega področnega podatkovnega skladišča za vrednotenje kreditne izpostavljenosti

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Matjaža Kukarja,
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki »Dela FRI«.

V Ljubljani, dne 24. septembra 2014

Podpis avtorja:

Zahvaljujem se mentorju Matjažu Kukarju za vso pomoč in strokovno vodenje pri izdelavi diplomskega dela.

Zahvaljujem se svoji družini, vsem prijateljem in sodelavcem. Rad vas imam.

Kazalo

Povzetek

Abstract

1. UVOD.....	1
2. PODATKOVNA SKLADIŠČA	2
2.1. Zgodovina razvoja podatkovnih skladišč	2
2.2. Kaj je podatkovno skladišče.....	2
2.2.1. Pristop W.H. Inmona [21]	3
2.2.2. Pristop Ralpa Kimballa [21]	4
2.2.3 Razlika med operativno podatkovno bazo ter podatkovnim skladiščem	4
2.3. Cilji podatkovnega skladišča:.....	6
2.4. Arhitektura podatkovnih skladišč.....	7
2.5. Trendi podatkovnih skladišč.....	7
3. BANČNO PODATKOVNO SKLADIŠČE.....	11
3.1. Specifike in težave podatkovnih skladišč v bankah.....	12
4. IZGRADNJA BANČNEGA PODATKOVNEGA SKLADIŠČA.....	14
4.1. Področno skladišče za ugotavljanje kreditne izpostavljenosti v poslovnih bankah	15
4.1.1. Vrste izpostavljenosti	19
4.2. Implementacija kreditne izpostavljenosti v bančno podatkovno skladišče	20
4.2.1 Podatkovni model podatkovnega skladišča	20
4.3. Postopek ETL (Extract, Transform, Load)	25
4.3.1. Ekstrakcija	26
4.3.2. Transformiranje.....	26
4.3.3. Nalaganje.....	27
5. REALIZACIJA VREDNOTENJA KREDITNE IZPOSTAVLJENOSTI V PODROČNO PODATKOVNO SKLADIŠČE TER IZDELAVA POROČIL.....	28
5.1. Polnjenje tabel z ETL orodjem IBM DataStage.....	28
5.1.1. Primer polnjenja dejstva za kreditno izpostavljenost s programom DataStage.....	31
5.2. Izdelava poročil iz podatkovnega skladišča za pomoč pri odločanju ter poročanju	37
5.2.1. SAP BusinessObjects BI	37
5.2.2 Business Objects Universe	38
6. SKLEP	42
7. LITERATURA IN VIRI	45

Seznam slik

Slika 1: Osnutek koncepta bančnega podatkovnega skladišča.....	11
Slika 2: Podatkovni model za pogodbe.....	22
Slika 3: Dejstvo kreditne izpostave s pripadajočimi dimenzijami.....	25
Slika 4: DataStage osnovna maska z osnovnimi gradniki.....	31
Slika 5: Opravilo za pridobitev časovnega obdobja.....	31
Slika 6: Opravilo za pridobitev podatkov iz sumarnih in atomarnih tabel.....	32
Slika 7: Opravilo za pridobitev ključev iz dimenzij.....	33
Slika 8: Opravilo za insert v tabelo dejstva.....	33
Slika 9: Krovno opravilo za polnjenje dejstva kreditnega tveganja.....	34
Slika 10: Zaporedje polnjenja virov na atomarnem nivoju v podatkovno skladišče.....	35
Slika 11: Zaporedje polnjenja dimenzij v podatkovnem skladišču.....	36
Slika 12: Rešitve v Business Object BI.....	38
Slika 13: Primer izdelanega vesolja za kreditno izpostavljenost.....	39
Slika 14: Osnovna maska Business Object Web intelligence z pripravljenim poročilom.....	39
Slika 15: Poročilo za kreditno izpostavljenost skupine do skupine povezanih oseb z bruto izpostavo nad 8% konsolidiranega regulativnega kapitala.....	41

Povzetek

Glavni cilj diplomske naloge je bila predstavitev razvoja in funkcionalnosti bančnega področnega podatkovnega skladišča za oceno kreditne izpostavljenosti. V prvem delu smo se posvetili samim osnovam podatkovnega skladišča, opisali različne pristope gradnje ter osvetlili razlike med podatkovnim skladiščem ter operativno podatkovno bazo. Opredelili smo arhitekturo, cilje in našli nekaj sodobnih trendov podatkovnega skladiščenja. V drugem delu smo se posvetili specifikam bančnega podatkovnega skladišča ter našli nekaj slabosti in ovir pri gradnji le tega. Opisano smo ilustrirali na primeru implementacije ocenjevanja kreditne izpostavljenosti v obstoječe bančno podatkovno skladišče. Opisali smo dimenzijski podatkovni model, postopke ETL za polnjenje podatkov in izdelavo samih poročil za pomoč pri odločanju ter poročanju. V diplomski nalogi smo predstavili tudi orodje za izvajanje ETL postopkov IBM DataStage ter orodje za izdelavo poročil SAP BusinessObject BI. Delo smo zaključili z učinki in rezultati gradnje področnega podatkovnega skladišča za kreditno izpostavljenost.

KLJUČNE BESEDE

- Podatkovno skladišče
- Kreditna izpostavljenost
- Poslovno poročanje
- Dimenzijski podatkovni model

Abstract

The main objective of the thesis is the presentation of evolution and functionality of the data mart in banking data warehouse for evaluation of credit exposure. In the first part we focused on the basics of data warehouse, describe different approaches to building it and some differences between the data warehouse and operational data base. We define architecture, objectives and highlight some latest trends of data warehousing. Then we focus on the specifics of the banking data warehouse and exposure some disadvantages and barriers in the construction it. We describe and illustrate the case of implementation the values of credit exposure to existing banking data warehouse. We describe a dimensional data model, ETL processes for loading data and producing reports to assist in decision-making and reporting. We also present a tool for implementation of ETL processes IBM DataStage and tool for creating reports BusinessObject SAP BI. We conclude the work with the effects and results of the data mart construction in a banking data warehouse for evaluation of credit exposure.

KEYWORDS

- Data warehouse
- Credit exposure
- Business Intelligence
- Dimensional data model

1.UVOD

Slovenske banke imajo težave z organizacijo svojih podatkov. Ta je vsaj neoptimalna, če ne kar katastrofalno zanemarjena. Podatki so sicer na individualni ravni dosegljivi, tudi za njihovo varnost je dobro poskrbljeno. Vendar je to le en vidik, ki bankam omogoča zgolj opravljanje svoje dejavnosti, nikakor pa jim ne nudi prave opore pri njeni strateški optimizaciji. Da bi informacijski sistem banke lahko odigral tudi to vlogo, bi moral vsebovati integralno podatkovno skladišče [13].

Dandanes igra konkurenčna prednost na trgu pomembno vlogo, zato se banke vse bolj zavedajo pomembnosti podatkov, ki jih lahko s pravilno analizo in uporabo obrnejo sebi v prid. To pa je mogoče le s čistimi, dobro urejenimi in strukturiranimi podatki, česar pa v transakcijskih sistemih ni mogoče dobiti. Podatkovno skladišče je edino pravo zbirališče podatkov, iz katerega je take podatke moč dobiti. Tega se vse bolj zavedajo tudi slovenske banke. Večina, če ne že vse, so je namreč že srečala s pojmom podatkovno skladišče, ki je sicer v slovenskem bančnem prostoru bančno še relativno mlad, pa vendar nujen.

V slovenskem prostoru je trenutno o tej temi dostopno relativno malo, zato smo se odločili, da v tem diplomskem delu predstavimo temelje bančnega podatkovnega skladiščenja, njegovo zgodovino ter tudi trende, ki se komaj oblikujejo. Dotaknili se bomo tudi težav ter ovir pri izgradnji bančnega podatkovnega skladišča. Ker je ta pojem preobsežen za samo eno samo diplomsko nalogo, smo se odločili, da predstavimo le en delček, in sicer dejstvo(angl. Fact) za kreditno izpostavljenost. Predstavili bomo tudi pojem kreditna izpostavljenost ter dejstvo napolnili s podatki z IBM-ovim ETL programom DataStage preko atomarnega nivoja in dimenzij. Vmes se bomo posvetili tudi podatkovnem modelu bančnega podatkovnega skladišča. Na koncu bomo vse skupaj zaključili s predstavitvijo programa SAP Business Object BI, ki je idealen za optimizirane poizvedbe izdelave poročil ter analiz. S tem programom bomo izdelali tudi nekaj poročil, ki jih morajo banke dnevno pošiljati Banki Slovenije v pregled.

2. PODATKOVNA SKLADIŠČA

2.1. Zgodovina razvoja podatkovnih skladišč

Revolucija relacijskih baz v začetku 1980 je pripeljala do izboljšanega dostopa do dragocenih informacij, ki so shranjene globoko v podatkih. Kmalu se je izkazalo, da učinkovito optimizirane transakcijske podatkovne baze niso bile vedno najboljše za analitične potrebe ter kompleksna poročanja. V bistvu se je potreba po sistemih, ki ponujajo podporo funkcionalnosti odločanja pojavila že v začetku 1970, ko sta podjetji ACNielsen in IRI predstavili sistem za informacijsko podporo prodaje, ki je vseboval določene značilnosti sistemov, ki jim danes pravimo področna podatkovna skladišča. Današnja praksa, znana kot skladiščenje podatkov, se je pojavila konec 80-ih let, ko je IBM System Journal objavil članek »An architecture for a business and information systems«, kjer se je prvič pojavil izraz »business data warehouse« [19]. Podoben termin je sicer že leta 1970 uporabljal Bill Inmon. Leta 1990 je pomemben korak pri razvoju skladiščenja podatkov naredilo podjetje Red Brick System z razvojem sistema RedBrick Warehouse za delo s podatkovnimi bazami, specializiranega za izvedbo podatkovnih skladišč. Podjetje Prism je leta 1991 predstavilo programsko opremo »Prism warehouse manager« za izgradnjo podatkovnih skladišč. Istega leta je Bill Inmon, znan kot oče podatkovnih skladišč, objavil knjigo namenjeno podatkovnim skladiščem »Building the Data Warehouse«, ki velja za prvo objavljeno knjigo o tej temi. Leta 1995 je bil ustanovljen inštitut »The Data Warehouse Institute« namenjen promoviranju skladiščenja podatkov. Leto kasneje je Ralph Kimball objavil knjigo »The Data Warehouse Toolkit«, v kateri je strnil veliko uporabnih znanj povezanih s podatkovnim skladiščem. V njej je predstavil enega izmed najbolj znanih pristopov gradnje podatkovnega skladišča [19].

2.2. Kaj je podatkovno skladišče

Podatki so za vsako podjetje izredno dragoceni, zaradi česar morajo biti pravilno shranjeni ter zlahka dostopni za čase, ko se jih potrebuje. Dostopnost ter črpanje pomembnih podatkov pa lahko zaradi njihove velike količine, kmalu postane izredno oteženo ali celo nemogoče.

Podatkovno skladiščenje je pojav, ki je zrasel iz ogromne količine elektronskih podatkov shranjenih v preteklih letih in zaradi nuje po urgentni uporabi teh podatkov za doseganje ciljev, ki presegajo redne naloge in dnevne obdelave. Skladiščenje podatkov je zbirka metod, tehnik in orodij za podporo delavcem, ki upravljajo z znanjem (npr. višje vodstvo, direktorji,

menedžerji in analitiki,...) in izvajajo različne analize, ki pomagajo pri izvajanju procesov odločanja in izboljšanje informacijskih virov [20].

Leta 1996 je R. Kimball učinkovito povzel nekaj trditev uporabnikov klasičnih informacijskih sistemov [20]:

- »Imamo kup podatkov, vendar do njih ne moremo dostopati!«
- »Kako imamo lahko isti pogled na podatke, vendar vedno dobimo različne rezultate?«
- »Želimo zbrati, združiti in obračati podatke na vse možne načine.«
- »Pokažite mi samo, kar je pomembno.«
- »Vsi vemo, da je s podatki nekaj narobe.«

Iz prejšnjega seznama problemov in težav lahko izvlečemo seznam ključnih besed, ki so razpoznavni znak in bistvo podatkovnega skladišča [20]:

- **Dostopnost** za uporabnike, ki ne poznajo dobro IT-ja ter podatkovne strukture.
- **Integracija** na podlagi standardnega modela podjetja.
- **Prožnost poizvedb** za povečanje informacije, pridobljene iz obstoječih podatkov.
- **Jedrnatost** podatkov, ki omogoča ciljno usmerjene in učinkovite analize.
- **Večdimenzionalna predstavitev**, ki daje uporabnikom intuitivni prikaz informacij.
- **Pravilnost in popolnost** integriranih podatkov.

Poznamo dva temeljna pristopa izgradnje podatkovnega skladišča:

2.2.1. Pristop W.H. Inmona [21]

Pristop Inmona »od zgoraj-navzdol« za oblikovanje podatkovnega skladišča, je pristop reševanja na nivoju podjetja in ne samo posameznih oddelkov, ki imajo različne zahteve, standarde ter integracije. Identificiral je potrebo po vključitvi podatkov iz različnih sistemov v centralizirano skladišče, kjer se lahko podatki uporabijo za izdelavo strateškega odločanja. Po Inmanu naj bi bili podatki sistemizirani kot predmetno-orientirani, integrirani, časovno vodeni in neizbrisljivi. Podatki morajo biti dosegljivi na ravni zrn z vrtanjem navzdol ali sumarizirani proti vrhu. Oblika Inmonovega podatkovnega skladišča je odvisna od tretje normalne oblike, ki omogoča razdrobljenost podatkov za kar največjo prilagodljivost podjetja. To se lahko izkaže kot zelo koristno, če se perspektiva organizacije za skladiščenje spremeni. Relacijsko oblikovano podatkovno skladišče se lahko uporabi za podporo raznoliki strukturi

podatkov OLTP (online transaction processing) baz, raziskovalnim skladiščem in skladiščem namenjenih za podatkovno rudarjenje. Oblika od zgoraj navzdol se začne z obravnavo pridobivanja podatkov iz operativnih podatkovnih virov. Podatki so nato vstavljeni v fazo črpanja in čiščenja, naprej v fazo transformiranja, integriranja ter konsolidacije in potem v operativno shrambo podatkov. Na koncu se podatki prenesejo v področna podatkovna skladišča (angl. data marte), kjer postanejo dosegljivi končnim uporabnikom.

2.2.2. Pristop Ralpa Kimballa [21]

Kimball je prestavil koncept dimenzijskega modeliranja, ki premošča prepad med relacijskimi bazami ter večdimenzionalnimi bazami. Zasnoval je koncept podatkovnega skladišča »od spodaj-navzgor« s povezovanjem področnih podatkovnih skladišč ter bus strukturo. Področna skladišča so skladna z bus strategijo dimenzij, domeno ter kazalniki, so dimenzionirani modeli oddelčnih podatkov, ki jih lahko vidimo v podatkovnem skladišču (unija področnih podatkovnih skladišč) podjetja. Ta ureditev naredi podatkovno skladišče podobno virtualni resničnosti, zato ta struktura omogoča fleksibilnost kreiranja področnih skladišč na različnih strežnikih v celotnem podjetju. Kimballov pristop je v nasprotju s pristopom Inmona. Pri tem postopku se podatki pridobljeni iz obstoječih podedovanih sistemov in nato konsolidirani, preverjeni ter črpani. Podatki so potem z postopki ETL vneseni v skladišče. Podatkov vnesenih v skladišče nikoli ne brišemo, ampak jih samo popravljamo oz. dodajamo nove. Ko skladišče vsebuje nove kopije podatkov, se ti integrirajo in transformirajo v strukturo področnih podatkovnih skladišč. Podatki so nato agregirani, sumirani in na voljo končnim uporabnikom za analizo ter strateško odločanje.

2.2.3 Razlika med operativno podatkovno bazo ter podatkovnim skladiščem

Operativni podatki po navadi pokrivajo kratko časovno obdobje, saj večina poslov vključuje najnovejše podatke. Podatkovno skladišče pa mora omogočati tudi analize za več let nazaj. Iz tega razloga se skladišče ažurira iz operativnih podatkov in vedno bolj raste. Če bi bili podatki predstavljeni grafično, bi jih lahko videli kot slike shranjenih operativnih podatkov. Zaporedje teh fotografij bi bilo shranjeno v podatkovnem skladišču, rezultat pa bi bil film, ki razkriva stanje v podjetju od njegove ustanovitve do danes.

Načeloma se podatki iz skladišč ne brišejo, posodobitve pa se običajno izvajajo, ko so izklopljena. oziroma, ko na njih nihče ne dela poizvedb. To pomeni, da so podatkovna skladišča v bistvu namenjena samo branju (read only).

	Transakcijska baza podatkov	Podatkovno skladišče
Uporabnikov	več tisoč	več sto
Obremenitev	transakcije	analitične poizvedbe
Dostopi	Na stotine zapisov, pisanje in branje	Milijoni zapisov, večinoma samo branje
Cilji	Odvisno od aplikacije	Podpora odločanju in poročanju
Podatki	Podrobni, numerični in nenumerični	Povzeti, predvsem numerični
Kvaliteta	celovitost	Dosledno sledenje začrtanih smernic in pravil
Časovna pokritost	trenutni podatki	trenutni in zgodovinski podatki
Posodobitev	stalna	periodična
Model	normaliziran	denormaliziran, multidimenzionalen
Optimizacija	Za OLTP dostop do dela zbirke podatkov	Za OLAP dostop do večino baze

Tabela 1- Razlika med transakcijsko bazo ter podatkovnim skladiščem (Povzeto po [18])

2.3. Cilji podatkovnega skladišča:

- Podatkovno skladišče mora narediti podatke organizaciji lahko dostopne.

Vsebina podatkovnega skladišča mora biti razumljiva ter smiselna. Podatki morajo biti intuitivni ter očitni za poslovnega uporabnika in ne le za razvijalce, saj želijo tudi poslovni uporabniki ločevati in združevati podatke v skladišču v neskončnih kombinacijah. Orodja, ki dostopajo do podatkov v skladiščih, morajo biti preprosta za uporabo in morajo vrniti rezultate poizvedb v najkrajšem možnem času.

- Podatkovno skladišče mora predstaviti podatke dosledno in kredibilno.

Podatki v podatkovnih skladiščih morajo biti kredibilni, skrbno sestavljeni iz različnih virov, očiščeni, kvalitetni in prikazani samo kadar ustrezajo uporabnikom. Informacije iz enega poslovnega procesa se morajo ujemati s podatki drugega. Če imata dva atributa iz različnih procesov enako ime, mora biti tudi njuna vsebina enaka, sicer se ne smeta imenovati enako.

- Podatkovno skladišče mora biti prilagodljivo ter odporno na spremembe.

Potrebe uporabnikov, pogoji poslovanja, podatki ter tehnologija so neizogibno podvrženi spremembam. Podatkovno skladišče mora biti narejeno tako, da z lahkoto kljubuje tem spremembam in pri tem ne spreminja obstoječih podatkov. Obstoječi podatki ter aplikacije morajo biti zasnovani tako, da ko uporabniki zastavijo nova vprašanja ali zahtevajo nove podatke, to tudi z lahkoto dobijo.

- Podatkovno skladišče mora biti varno in mora ščititi informacije Organizacije.

Podatki organizacije so za organizacijo zaklad. Najmanj, kar skladišče vsebuje, so informacije o tem, kaj prodajajo, komu in po kakšni ceni. Če ti podatki pridejo v roke napačnih ljudi, je to lahko škodljivo. Podatkovno skladišče mora biti varno in mora varovati zaupne podatke.

- Podatkovno skladišče mora biti temelj za izboljšanje odločitvenih sistemov.

Podatkovno skladišče mora vsebovati prave podatke, če želimo da nam pomaga pri odločitvah. Odločitve, ki so narejene na podlagi informacij iz podatkovnega skladišča morajo biti predstavljene z dokazi.

- Poslovna skupnost mora sprejeti podatkovno skladišče, če želimo da bo uspešno.

Nič nam ne pomaga, če zgradimo eleganto rešitev, ki je poslovna skupnost ne sprejme in je ne uporablja aktivno. Uporaba podatkovnega skladišča velikokrat ni nujna, kar pripelje do tega, da se ji uporabniki izognejo.

2.4. Arhitektura podatkovnih skladišč

Sledeče lastnosti so bistvene za sistem podatkovnega skladišča [18]:

- **Ločitev:** Analitična in transakcijska obdelava morata biti čim bolj ločeni.
- **Nadgradljivost:** Strojna in programska arhitektura mora biti enostavno nadgradljiva.
- **Razširljivost:** Arhitektura mora biti sposobna gostiti nove aplikacije in tehnologije brez prenove celotnega sistema.
- **Varnost:** Spremljanje dostopov je bistveno zaradi strateških podatkov shranjenih v podatkovnih skladiščih.
- **Upravljanje:** Podatkovnega skladišča ne sme biti prezahtevno upravljati.

2.5. Trendi podatkovnih skladišč

Podatkovno skladiščenje ni več le nova ideja za študij oziroma eksperimentiranje. Postalo je nujnost. Sicer ni še prisotno v vsaki trgovini, pisarni, ordinaciji, ... vendar tudi ni več posvečeno samo poslovanju visoko dobičkonosnih podjetij. Več kot polovica podjetij v Ameriki se je zavezala uporabi podatkovnim skladiščem. Približno 90 odstotkov multinacionalk ima podatkovno skladišče ali pa ga namerava zgraditi v kratkem. Vsako podjetje, ki ima podatkovno skladišče, se zaveda ogromnih prednosti, ki jih le-ta prinaša. Nekateri strokovnjaki menijo, da tehnologija pelje podatkovno skladiščenje naprej in da smo prišli do točke, ko smo priča pomembnemu napredku v programski opremi. V naslednjih letih je v skladiščenju podatkov pričakovati velike korake v programski opremi, predvsem za optimiziranje poizvedb, indeksiranje velikih tabel, kompresijo in širjenje dimenzij modeliranja (zbirka tehnik in konceptov v podatkovnem skladišču) [11].

Nekaj sodobnih trendov v podatkovnem skladiščenju [11]:

- Skladiščenje podatkov v realnem času

Klasični poslovno inteligentni sistemi in podatkovna skladišča so bili uporabni predvsem za strateška odločanja. Podatkovna skladišča so bila ločena od operacijskih sistemov. Zadnje čase pa podjetja stremijo k uporabi poslovne inteligence tudi za taktično odločanje na dnevni osnovi. Tradicionalno je skladišče pasivno in zagotavlja zgodovinske trende, skladiščenje v realnem času pa je dinamično in zagotavlja zadnje (up-to-date) poglede poslovanja. Podjetja se soočajo z velikim pritiskom zagotavljanja podatkov na dnevni osnovi oz. v realnem času, kar pa je za podatkovno skladiščenje velik izziv, saj informacije pridobljene v realnem času povečujejo produktivnost.

Primer prodajalca v trgovini: Zanje bi bilo najbolje, če bi stranko lahko obravnavali nenehno. Med nakupom v trgovini ali na spletni strani, bi želeli spremljati kupčeva ravnanja, hkrati pa bi imeli na voljo tudi vzorec zgodovine nakupov kupca, da ga bo moč primamiti z ustreznimi izdelki.

- Različni tipi podatkov

Običajno so, ko se zgradi prva iteracija podatkovnega skladišča, podatkovni tipi osnovni (Numeric,...), vendar se kmalu izkaže, da to ni dovolj. Zato se je začelo v skladišča vključevati tudi nestrukturirane tipe (slike, video, audio,...). Primer: Recimo, da odločevalec izvaja analizo, da bi našel tip najbolj prodajanega produkta. Ko pride do določene vrste izdelkov, ki obstajajo v času analize, želi te izdelke videti, kar mu pomaga pri nadaljnih odločitvah. V glasbeni ter filmski industriji so audio ter video tipi podatkov v podatkovnem skladišču lahko zelo uporabni.

- Vizualizacija podatkov

Ko uporabnik piše poizvedbe in dobi rezultate samo v obliki preglednic ter seznamov, ti rezultati niso več ažurni. Rezultate je potrebno prikazati v obliki slik in grafov, kar tudi vsak uporabnik pričakuje. Vizualizacija podatkov omogoča, da lahko vidimo različne trende v času, kar nam pomaga, da si lažje in hitreje razlagamo podatke.

- Orodje za poizvedbe

V podatkovnem skladišču je zelo pomembno tudi poizvedbeno orodje., zato se le-ta vse bolj razvijajo. Orodje mora biti enostavno za uporabo z možnostjo predstavitve rezultatov na spletu v različnih formatih (SAP BI BusinessObject, IBM Cognus, AQT,..).

- Podatkovno rudarjenje (Data mining)

Podatkovno rudarjenje se ukvarja z odkrivanjem skritega znanja v podatkih ter iskanju nepričakovanih vzorcev in pravil na veliki količini podatkov. Dobro zgrajeno podatkovno skladišče predstavlja izhodišče za podatkovno rudarjenje. Če si znamo pridobljeno informacijo razložiti in jo lahko uporabimo v korist podjetja, je bilo rudarjenje uspešno. Veliko podjetij se je šele začelo zavedati prednosti podatkovnega rudarjenja in izkoriščanja potencialov, ki ga nudi. Podjetja, ki rudarjenja še ne uporabljajo, bodo prej ali slej prisiljena v njegovo uporabo.

CRISP-DM (Cross Industry Standard Process for Data Mining) je procesni model, ki opisuje pogoste postopke, ki jih podatkovni rudarji uporabljajo za reševanje problemov [17].

CRISP-DM razdeli proces podatkovnega rudarjenja na šest glavnih faz [17]:

- Poslovno razumevanje
- Podatkovno razumevanje
- Priprava podatkov
- Modeliranje
- Vrednotenje
- Predaja končnemu uporabniku

- Veliki podatki (Big data)

Dandanes lahko podjetja hitreje in iz več virov zberejo veliko več podatkov, kot je bila praksa v preteklosti. Veliki podatki je izraz, ki opisuje zbiranje, organiziranje, analiziranje in upravljanje podatkov. To je naslednja generacija poslovne inteligence, podatkovnega skladiščenja ter analitike podatkov, ki zahteva nove načine za organizacijo podatkov, nova orodja za analizo podatkov ter vsekakor ekipo ustreznih ljudi, ki bodo razumeli podatke.

Veliki podatki so veliki, saj je njihov obseg veliko večji od tipičnih baz podatkov. Vendar pa je pomembno vprašanje, kako te podatke uporabiti za povečanje produktivnosti, konkurenčnosti, zmanjšanje stroškov, povečanje prodaje, izboljšanje procesov..., in ne njihova količina. Glavna trenda velikih podatkov sta vedno večja količina podatkov ter zmanjšanje razkoraka med zbiranjem podatkov in analizo.

Uspešno izkoriščanje velikih podatkov lahko pomagajo banki doseči tri ključne cilje [22]:

- Optimiziranje ponudbe

Biti sposoben predvideti potrebe kupcev, kar lahko vodi do povečanje prodaje, zmanjšanja stroškov ter večjega zadovoljstva strank.

- Ugotavljanje in odkrivanje goljufij

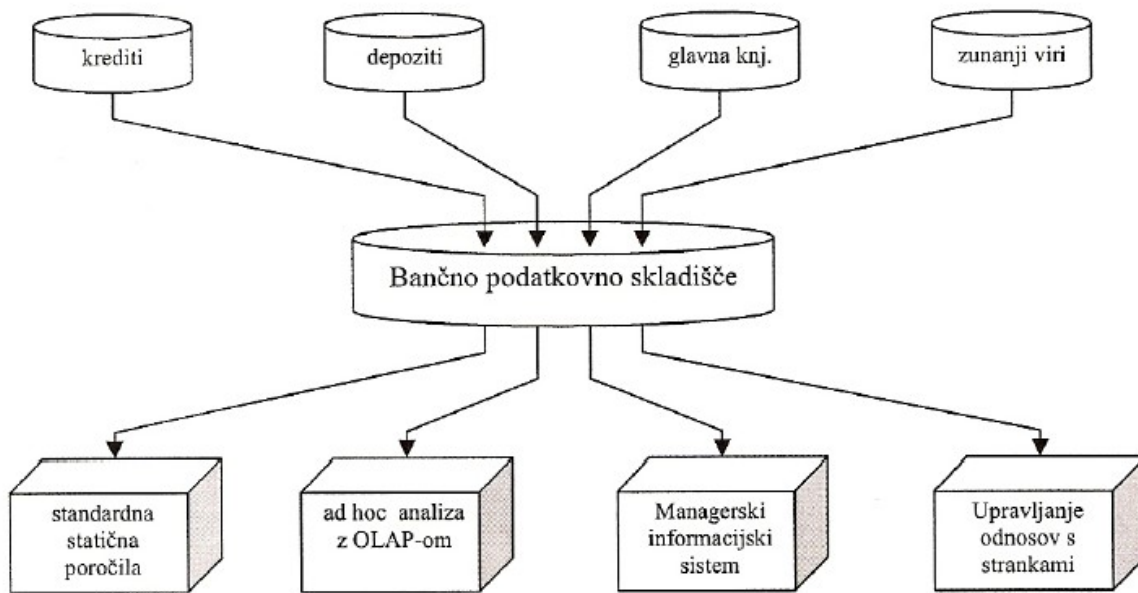
Odkrivanje goljufij in prevar ter skrb za varnost podatkov in transakcij je eno izmed zelo pomembnih področij v bančništvu.

- Upravljanje s kreditnimi tveganji

Eksplozija količine podatkov potrebnih za oceno kreditnega tveganja.

3. BANČNO PODATKOVNO SKLADIŠČE

Poslovne banke se vse pogosteje srečujejo z vse večjimi zahtevami po podatkih o bančnih poslih. Na eni strani jim Banka Slovenije, država ter tudi Evropska unija z zakoni in predpisi narekujejo določen tempo ustvarjanja poročil, na drugi strani pa si uprava ter analitiki želijo globlji vpogled v samo strukturo poslov, ki presegajo okvire računovodskih informacij znotraj bilance stanja in izkaza uspeha [12]



Slika 1: Osutek koncepta bančnega podatkovnega skladišča[12]

Pri gradnji bančnega podatkovnega skladišča so ključni naslednji dejavniki [12]:

- Skladišče moramo graditi od zgoraj navzdol v skladu s celotno strukturo banke.
- Graditi ga moramo interaktivno. Začnemo z manjšim naborom podatkovnih polj, ki jih nato širimo.
- Prva konkretna faza mora biti končana v štirih mesecih. Uporabniki morajo videti konkurenčne prednosti, ki se lahko takoj vpeljejo v proces odločanja s ciljem povečanja dobička.
- Uprava in višji management morata podpirati projekt. Lahko celo ustanovita odbor za bančne podatkovne vire.
- Ključno je tudi sodelovanje uporabnikov ter razvijalcev. Dobro je, da so fizično in organizacijsko skupaj.

- Projekt naj vodi končni uporabnik – nekdo, ki pozna ključne poslovne probleme.

Prednosti uveljavljanja bančnega podatkovnega skladišča so večplastne. Analitiki izdelajo poročila hitreje, saj lažje in hitreje pridejo do podatkov, informatiki so manj obremenjeni z izdelavo poizvedb in izpisov, saj lahko analitiki sami naredijo poročila, banka pa pozna stranko v celoti in razume, katera stranka je bolj donosna. Te koristi so dejansko merljive [12].

Bančna podatkovna skladišča nudijo zagotavljanje jasne vizije, odkrivanje novih priložnosti (konkurenčna prednost), zagotavljajo pravočasne informacije, zmanjšujejo poslovni riziko, izboljšujejo sodelovanje in razumevanje med uporabniki, omogočajo boljše upravljanje sredstev banke,...

Podatki zbrani v podatkovnem skladišču v večji meri nastanejo v transakcijskih sistemih. Pri polnjenju se transformirajo (očistijo, dopolnijo, poenotijo) in napolnijo v skladišče prek ETL procesov. Tu pride do preskoka v kakovosti podatkov. Podatke lahko tudi dodatno opremimo z različnimi informacijami. Na ta način se podatkom poveča vrednost za različne analize poslovanja banke, kar predstavlja boljšo osnovo za poslovne odločitve.

3.1. Specifike in težave podatkovnih skladišč v bankah

Najpogostejša slabost podatkovnega skladiščenja v bankah v primerjavi s podatkovnim skladiščenjem v drugih panogah je vsekakor večje število transakcijskih sistemov, kar je rezultat postopnega nastajanja sistemov za podporo poslovanja po posameznih produktih (kreditih, depoziti, TRR-ji), ki med seboj v večini primerov niso povezani. Tako prihaja do podvojenih podatkov v skladiščenju (npr. v aplikaciji za depozite je stranka iz Ljubljane, v aplikacije za kredite pa je ta ista stranka iz Maribora). Ta razdrobljenost in nepovezanost vodi v slabo kakovost podatkov na ravni banke, saj ti podatki ne zagotavljajo dobrih poslovnih odločitev.

Velika ovira pri izgradnji bančnega podatkovnega skladišča je pomanjkanje resursov in predvsem znanja. Veliko bank se zaveda pomembnosti takšnega projekta, a si v trenutnih pogojih zaostrenega boja tega ne morejo privoščiti. Od velikosti banke je odvisno, kako velik tim stalnih sodelavcev sestavlja izgradnjo skladišča. Kimball v svojih delih govori o deset in več različnih vlogah, ki jih je potrebno pokriti, med drugim vodjo projekta, podatkovnega

skrbnika, programerje, administratorja baze, načrtovalca podatkovnih modelov,... Gre za profil z zahtevanim visokim strokovnim znanjem na področju informatike in vsebinskih področjih bančništva. Takšen profil pa je na trgu delovne sile pomanjkljivo zastopan, zato največkrat bankam ostane izgradnja internega kadra [14].

Težava nastane tudi pri dnevnem polnjenju, oziroma v nekaterih primerih skoraj on-line polnjenju skladišča. Nemalokrat se zgodi, da nekateri viri zamujajo in zato poročevalci ne dobijo pravih podatkov v poročilih. Banka mora identificirati in spremeniti vse poslovne procese in poslovno prakso na tak način, da bodo vsi postopki knjiženj in obdelav podatkov zaključeni še isti dan oz. do konca dneva. Ta zahteva je ena izmed težjih in predstavlja veliko oviro za popolnoma uspešno implementacijo podatkovnega skladišča.

Pri polnjenju podatkov v podatkovno skladišče je velika težava tudi kvaliteta podatkov, saj se skladišče polni s podatki iz različnih virov, aplikacij in transakcijskih sistemov. Potrebno je poskrbeti, da se vsi vpleteni začnejo zavedati problematike kvalitete podatkov. Kljub dejstvu, da enolična definicija kvalitete podatkov ni mogoča, lahko povzamemo naslednje opredelitve [25]:

- Kvaliteta podatkov - kriterij za določanje kvalitete podatkov je lahko spremenljiv glede na tip podatkov, način uporabe, poslovne potrebe, tehnološke možnosti ter še sprejemljivo toleranco do napačnih, pomanjkljivih in neenotnih podatkov.
- Kvaliteta podatkov in poslovni procesi - poslovni procesi z vidika kvalitete podatkov se nanašajo na zagotavljanje kvalitete podatkov, na katero vplivajo uporabniki preko različnih aplikacij. Proces se nanašajo na standardizacijo od samega vnosa podatkov do definiranja procesov, kjer se podatki pregledajo in ročno popravljajo ter nazadnje do splošnega pristopa pri upravljanju in nadzoru podatkov.

4. IZGRADNJA BANČNEGA PODATKOVNEGA SKLADIŠČA

Za banko je izgradnja podatkovnega skladišča strateškega pomena, kar pomeni, da je tudi odločitev za izgradnjo skladišča strateška odločitev najvišjega vodstva, ki poskrbi za pravilno pozicioniranje podatkovnega skladišča in njegovih rezultatov v banki. Izgradnje se moramo lotiti kot kontinuiranega dolgoročnega procesa, za katerim stoji uprava banke. Pri izgradnji in procesu dograjevanja skladišča morajo sodelovati tudi izvršni direktorji v banki, ki se skupaj z upravo odločajo o prioritetah v razvoju in dolgoročnih ciljih bančnega skladišča ter posredno skrbijo tudi za financiranje. Sodelovati morajo tudi končni uporabniki, saj bodo le na ta način zbrani pravi podatki, ki bodo odgovarjali na poslovna vprašanja.

Glavna zahteva izgradnje podatkovnega skladišča je vsekakor razumevanje poslovnih potreb in ciljev. Prav tako je velikega pomena za uspeh skladišča določanje poslovnih prioritet ter upoštevanje zastavljenih časovnih okvirov.

Cilji so vsekakor višanje dobička z nižanjem operativnih stroškov, boljše poslovne odločitve in višanje učinkovitosti poslovanja. Zavedati se moramo, da bančno skladišče ni poceni in pomeni veliko porabo sredstev ter časa, po drugi strani pa pomeni tudi velik prihranek ter primerjalno prednost (poznavanje strank, prilagoditve ponudbe,...). Tako je bančno skladišče investicija in ne le strošek.

Ko gradimo bančno podatkovno skladišče naletimo na kar nekaj perečih vprašanj [14]:

- Ali skladišče kupiti ali ga izdelati?

Strokovnjaki so precej enotni pri zagovarjanju mnenja, da se bančnega podatkovnega skladišča ne more kupiti. Bančna teorija je tako obsežna, da kljub dobri razdelanosti ne obstaja enoten konsenz o tem, kateri pristop je na kakšnem področju najboljši. Poleg tega ima vsaka banka svojo organizacijsko kulturo, svoj pogled na to, kaj je ključnega pomena in vsaka uprava svoj pristop k managementu. Tako pridemo do naslednjega vprašanja.

- Za katero področje začeti graditi skladišče?

To odločitev mora sprejeti management in zanjo prevzeti vso odgovornost, kar pa ne pomeni, da banka ne more kupiti »know-how« na področju podatkovnega skladiščenja v obliki strokovnjakov, ki so na tem področju že nabrali določene izkušnje. Lahko najame in si pomaga tudi z izkušenimi programerji na tem področju, prav tako lahko kupi pomoč pri projektnem vodenju samega skladišča. V bistvu velja, da lahko kupuje ljudi in njihovo znanje, nikakor pa ne sme kupovati produkta.

Verjetno pa bi se v teh kriznih časih banke vseeno poenotile in za začetek gradnje podatkovnega skladišča izbrale področje kreditne izpostavljenosti. Nenazadnje je to ena izmed glavnih težav oz. težnja, kako to tveganje čim bolj zmanjšati. Konec koncev pa tudi Banka Slovenije ter Evropska unija od bank strogo zahtevata vsakodnevno poročanje o izpostavljenosti.

4.1. Področno skladišče za ugotavljanje kreditne izpostavljenosti v poslovnih bankah

Iz kreditnega tveganja izhaja več kot polovica bančnih problemov, zato lahko trdimo, da je to osrednje mesto tveganja v bankah. O kreditnem tveganju govorimo, ko posojilodajalec ni več zmožen vračati dolga posojilodajalcu. Opredelimo ga lahko kot verjetnost, da ob zapadlosti glavnica posojila in/ali obresti niso v celoti poplačane [3].

Banka pri vsakem kreditiranju prevzame tveganje, da ji posojilodajalec kredita ne bo vrnil ali ga ne bo vrnil v dogovorjenem roku. Tako mora banka vnaprej računati možnost izpada pri vračanju kredita in obresti. Bančni posel se zaključi šele, ko je kredit vrnjen v celoti in brez kakršnihkoli izgub. Da zniža kreditno tveganje, mora banka omejiti slaba posojila tako, da pred odobritvijo analizira boniteto potencialnega kreditodjemalca. Pozorno mora spremljati tudi vračanje kredita in razprševati svoj kreditni portfelj. To je začetek postopka aktivnega nadzora kreditnega tveganja s strani banke.

Bessis deli kreditno tveganje na tri različne sklope [3]:

- Tveganje neplačila, ki ga ugotavljamo kot verjetnost, da v določenem časovnem obdobju pride do zamud pri vračanju kredita.

- Tveganje izpostavljenosti, ki izvira iz nepoznavanja bodoče izpostavljenosti v primeru neplačila.
- Tveganje poplačila pomeni možnost, da v primeru neplačila v nadaljnji izterjavi ali z unovčenjem zavarovanj ne bo prišlo do poplačila. Tukaj je najpomembnejša kakovost zavarovanja.

Banka se, zaradi nenehnih spreminjanj sredstev ter finančnih virov pri kreditnih odnosih s svojimi komitenti, ne sme opreti izključno na raven kreditne sposobnosti ob odobritvi kredita. Dejavniki tveganja so nenehno prisotni in se spreminjajo. Tako ne moremo govoriti o poskusu omejevanja in odpravljanja kreditnih tveganj brez predhodne sistemizacije in analize vsake rizične skupine.

Glede na vzroke ločimo naslednje skupine kreditnih izpostavljenosti [16]:

- Prevelika koncentracija kreditne izpostavljenosti .
- Nepravilnosti pri določanju optimalnega kreditnega obdobja.
- Nepravilnosti pri uporabi odobrenega kredita – ne nameska poraba.
- Pomanjkljiva ocena določenih parametrov kreditne sposobnosti in bonitetne poslovanja.
- Institucionalna osnova kot determinanta tveganja in stabilnost gospodarjenja.
- Inflacija in kreditno tveganje.

Bistvena sestavina skrbnega poslovanja banke je upravljanje s kreditnim tveganjem. Preudarno upravljanje s kreditnim tveganjem vključuje razmerje med tveganjem in donosom ter nadzor in znižanje tveganja. Podrobnosti upravljanja se razlikujejo med bankami glede na naravo in zapletenost njihovih kreditnih funkcij in portfeljev, vendar mora splošen program bank upravljanja s tveganji prepoznati obstoječe in potencialno kreditno tveganje, ki mu je banka izpostavljena. Razviti in uporabljati mora postopke za učinkovito odobravanje kreditov, njihovega dokumentiranja ter izterjave.

Merjenje kreditnega tveganja je podrejeno trem glavnim komponentam [6]:

- Izračunu verjetnosti, da posojilojemalec ne bo pravočasno ali v celoti poravnal svojih obveznosti,
- Izračunu izpostavljenosti banke iz naslova kreditnega tveganja v določenem časovnem obdobju, ter

- Izračunu dejanske izgube v procentih, ki jo banka realizira potem, ko je že uveljavila zavarovanja za odobreno posojilo oziroma začela postopek izterjave neplačanega dolga.

Banke morajo učinkovito in pozorno spremljati izpostavljenost do posameznih oseb, skupin povezanih oseb, bank in držav.

Izpostavljenost banke do posamezne osebe je vsota vseh terjatev, naložb banke v kapital in vrednosti naložb v vrednostne papirje te osebe in pomeni prikaz največje možne izgube banke v primeru, da ta oseba ne poravnava vseh svojih obveznosti.

Za posamezno osebo se štejejo v izpostavljenost vse klasične zunaj bilančne in aktivne bilančne postavke. Pri izračunu izpostavljenosti za posameznega komitenta se štejejo tudi vse osebe, ki so med seboj povezane in predstavljajo za banko eno samo tveganje.

25 odstotkov kapitala banke je največja dopustna izpostavljenost banke do posamezne osebe. Vsota vseh velikih izpostavljenosti banke ne sme presegati 800 odstotkov višine kapitala banke [15].

Kot veliko izpostavljenost banke do posamezne osebe opredeljuje Zban v 81. členu vse izpostavljenosti višje od 10 odstotkov kapitala banke. Za vse take osebe je za sklenitev posameznega pravnega posla potrebno pridobiti soglasje nadzornega sveta banke. Soglasje nadzornega sveta je potrebno tudi za sklenitev pravnega posla, zaradi katerega se velika izpostavljenost banke do posamezne osebe poveča tako, da doseže oziroma preseže 20 odstotkov kapitala banke[15].

Navadno ugotavljamo izpostavljenost banke do vsake posamezne osebe ter vsake odvisne družbe do vsake posamezne osebe.

Izpostavljenost do skupine povezanih oseb je enaka seštevku izpostavljenosti do posameznih oseb znotraj skupine povezanih oseb.

Pravila konsolidacije določajo, da se pri ugotavljanju izpostavljenosti bančne skupine do vsake posamezne osebe ali bančne skupine do skupine povezanih oseb, medsebojne terjatve in oslabitve v okviru bančne skupine ne upoštevajo. Pri ugotavljanju izpostavljenosti skupine banke se torej v stanje izpostavljenosti ne vključujejo terjatve banke in oslabitve do odvisnih družb, terjatve in oslabitve ene odvisne družbe do druge odvisne družbe ali terjatve in

oslabitve odvisne družbe do banke ne glede na to ali gre za terjatve in oslabitve iz naslova postavk sredstev ali naslova zunaj bilančnih postavk.

Posamezne osebe, ki so med sabo povezane po definiciji Sklepa o veliki izpostavljenosti bank in hranilnic, tvorijo skupino povezanih oseb. Ena posamezna oseba je lahko v več skupinah povezanih oseb. Posamezne osebe, ki tvorijo skupino povezanih oseb, so lahko tako pravne kot fizične osebe.

Navadno merimo izpostavljenost banke po produktih, vrsti finančnega instrumenta, po kriteriju ali je postavka Trgovalne knjige ali postavka Bančne knjige, po ročnosti, po kritosti, po OE (organizacijske enote) po generičnih OE,...

Za kreditno izpostavljenost navadno izračunavamo tri različne mere izpostavljenosti, ki jih izračunamo po formulah [26].

- Dejanska izpostavljenost

$$AE(c, t) = \max\{0, V(c, t)\}$$

Dejanska izpostavljenost (AE) je maksimum neke pogodbe (c) na določen časovni trenutek (t) med 0 ter vrednostjo pogodbe na določen časovni trenutek (V(c,t)), ki bo izgubljen, če dolg ne bo vrnjen.

- Potencialna izpostavljenost

$$PE(c, t) = \max\{0, \max_{t < \tau \leq T} \{PV_t[V(c, \tau)] - V(c, t)\}\}$$

Potencialna izpostavljenost pogodbe je maksimum dodatne vrednosti dejanske izpostavljenosti, ki bo izgubljena za neko časovno obdobje (ne pa na določen časovni trenutek), kjer je $PV_t[*]$ funkcija, ki transformira vrednosti v prihodnosti na sedanjo vrednost. Potencialna izpostavljenost mora biti definirana z upoštevanjem spreminjanja trga, to pa zato, ker se vrednost pogodbe glede na trg spreminja.

- Celotna izpostavljenost

$$TE(c, t) = AE(c, t) + PE(c, t)$$

Celotna izpostavljenost je seštevek dejanske in potencialne izpostavljenosti. To je maksimalna vrednost, ki bo izgubljena po pogodbi, če komitent v katerikoli točki na določen časovni trenutek z upoštevanjem sprememb na trgu odpove oz. neha plačevati kredit.

4.1.1. Vrste izpostavljenosti

Izračunavamo naslednje vrste izpostavljenosti:

- Bruto izpostavljenost (knjigovodska vrednost postavke sredstev pred upoštevanjem oslabitev in vrednost zunaj bilančne postavke)

Odštejemo oslabitve (za postavke sredstev) in rezervacije (za zunaj bilančne postavke) in dobimo:

- Neto izpostavljenost

Upoštevamo možna izvzetja (delno ali v celoti) in zmanjšanja zaradi učinkov kreditnih zavarovanj in dobimo:

- Izpostavljenost po Sklepu o veliki izpostavljenosti, Izpostavljenost po sklepu

Primer izračuna bruto ter neto izpostavljenosti [26]:

PRIMER ANALITIČNA STANJA GK po IDPOGODBE

konto	vrsta	stanje breme	stanje dobro	stanje
xxx	Glavnica (nezapadla)	1200000	0	-1200000
xxx	Glavnica (nezapadla)	2800000	1200000	-1600000
xxx	Terjatve in obveznosti za obresti (nezapadle)	7610,09	3867,42	-3742,67
xxx	Terjatve za opravnine za obdelavo vloge (nezapadle)	1400	1400	0
xxx	Stroški ob začetnem pripoznavanju FI	408,49	2200	1791,51
xxx	Popravek vrednosti obresti	36,74	72,3	35,56
xxx	Popravek vrednosti za opravnine za obdelavo vloge	13,3	13,3	0
xxx	Popravek vrednosti glavnice	11400	38000	26600

BRUTI IZP. = [(Glavnica (nezapadla)+ Terjatve in obveznosti za obresti (nezapadle)) + (Stroški ob začetnem pripoznavanju FI)] *(-1)

BRUTI IZP. = [(-1200000 + -1600000 + -3742,67)) + (1791,51)] * (-1) = 2801951,16

POPRAVKI, REZERVACIJE = [(Popravek vrednosti glavnice + Popravek vrednosti obresti)] = 26635,56

NETO IZPOSTAVA = BRUTO – POPRAVKI = 2775315,6

4.2. Implementacija kreditne izpostavljenosti v bančno podatkovno skladišče

Ker je kreditna izpostavljenost eden izmed glavnih oziroma osrednjih poslanstev bank in njihovega poslovanja, se bomo v diplomskem delu posvetili izdelavi področnega podatkovnega skladišča za področje kreditne izpostavljenosti.

4.2.1 Podatkovni model podatkovnega skladišča

Podatkovni model je abstraktna vendar lahko razumljiva predstavitev potrebnih podatkov celotne organizacije oziroma njenega dela, kar pomeni predstavitev same organizacije oz. njenega dela z vidika potrebnih podatkov.

Modeliranje podatkov je zaporedje postopkov, s katerimi opredelimo entitete, razporedimo attribute (podatke) med entitete in preverimo ali atributi res sodijo k tem entitetam.

Podatkovne modele uporabljamo za predstavitev konceptualnega in zunanega nivoja podatkovne baze, ne pa tudi fizičnega nivoja. Podatkovni model opisuje okolje s termini kot so **entitetni tipi, atribut, relacija, ključ**.

Entitetni tipi ali predmet podatkov je objekt opazovanja. Objekt opazovanja pa je vse, kar lahko enolično identificiramo, izoliramo od okolice in opišemo. Entitetne tipe identificiramo tako, da jim določimo ime, lastnosti entitete in vrednost lastnosti. Entitetni tipi so npr.: delavec, proizvod, kupec, naročilo,...

Atribut - entiteto določene vrste opisujemo z lastnostmi oz. atributi. Atribut ali podatkovni element ali element podatkov je tisti podatek, ki z vidika vsebine, torej z vidika uporabnika, ni več členljiv oz. ga ni možno ali smiselno razstaviti na manjše dele. Atribut je npr.: datum rojstva, matična številka,...

Povezave med entitetami - **relacije** ponazarjajo odnose med predmeti podatkov v smislu števila povezanih predmetov podatkov.

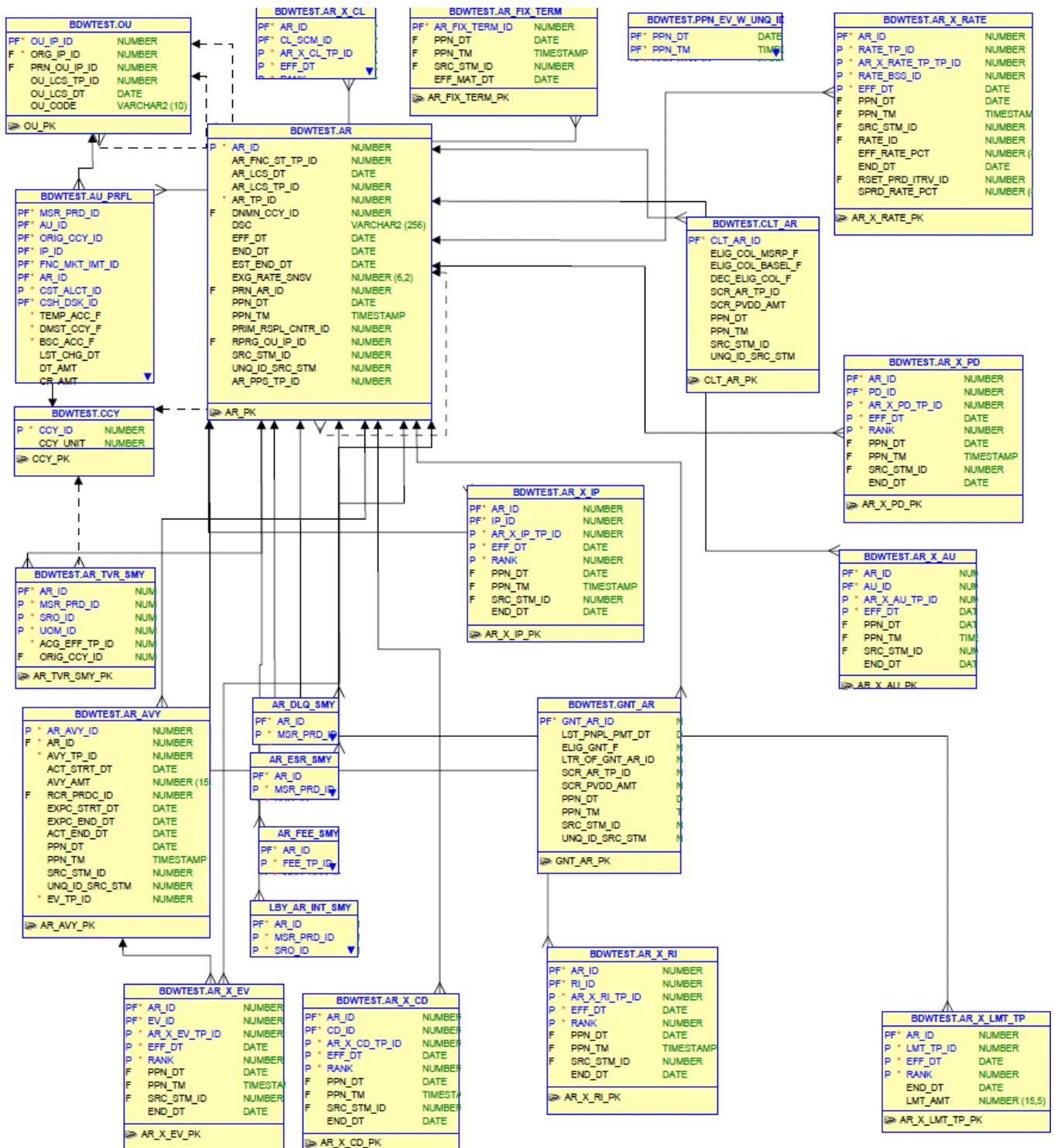
Poznamo dva glavna koncepta podatkovnih modelov.

- ER model (entity-relationship)
- Dimenzijski podatkovni model

Ker je zadnje čase v podatkovnem skladiščenju vse bolj razširjen dimenzijski podatkovni model bomo predstavili le-tega.

4.2.1.1. Atomarni nivo

Najprej moramo podatke iz transakcijskih sistemov spraviti v podatkovno skladišče na atomarni nivo. Atomarni nivo je v bistvu namenjen podatkom na najnižjem nivoju detajlov. Ponuja osnovne podatke za vse podatkovne transformacije, iz katerih potem polnimo dimenzije ter dejstva. Na atomarnem nivoju so shranjeni razni šifranti (stroškovna mesta, valute, države, mesta,...), registri (pogodbe, osebe, konti...) in vse povezave med njimi. Na tem nivoju je zelo pomembno, da imamo kar se da dobro razdelan podatkovni model. Slika 2 predstavlja pogodbe ter odvisne relacije, ki so pomembne v odnosu do pogodb. Že iz slike vidimo kako komplicirano in zahtevno je postaviti dober podatkovni model na atomarnem nivoju.



Slika 2 Podatkovni model za pogodbe

4.2.1.2. Sumarne tabele ter agregati

To so načeloma tabele, v katerih so združeni ali povzeti podatki zaradi performančnih razlogov. Podatki so združeni po pravilih in/ali združujejo veliko količino podatkov. Podatke lahko združimo po eni ali več dimenzijah, ki jih podatki vsebujejo. V podatkovnem skladišču se z vsako summarizirano tabelo rezervira določen prostor, tako se poveča hitrost pridobivanja podatkov, saj se te tabele dnevno osvežujejo.

Glavni cilj oblikovanja sumarnih tabel in agregatov je zmanjšanje količine podatkov, do katerih se dostopa, in število tabel, s katerimi se te podatke združi. V bistvu so to neke vrste vmesne oz. pomožne tabele, s katerimi razvijalci pohitrijo polnjenje podatkov v dejstva. Največkrat jih oblikujejo oziroma kreirajo kar razvijalci sami, s čimer lahko bistveno pohitrijo delovanje podatkovnega skladišča.

4.2.1.3. Dimenzijski podatkovni model

Dimenzionalni model je relativno nov koncept in ni dokončno definiran, predvsem ko ga primerjamo z drugimi tehnikami npr. ER model. Dimenzionalni model je tehnika za konceptualizacijo ter vizualizacijo podatkovnih modelov kot niz ukrepov, ki so opisani v splošnih vidikih posla. To je še posebno koristno za preurejanje in povzemanje podatkov ter za predstavitev pogledov za podporo analize podatkov. Osredotoča se predvsem na numerične podatke kot so vrednosti, štetje, uteži stanja in dogodki.

Dimenzionalni model ima več osnovnih gradnikov:

- Dejstva (Facts)

Dejstva so zbirka povezanih podatkovnih elementov sestavljenih iz ukrepov in kontekstnih podatkov. So centralne tabele dimenzijskega modela. Vsebujejo izmerjene podatke o rezultatih poslovanja v obliki t.i. merljivih dejstev. Zapisi v tabeli dejstev predstavljajo posamezna dejstva, pomembna za poslovni proces. Vsako dejstvo običajno predstavlja en poslovni element, poslovno transakcijo ali dogodek, ki se lahko uporablja pri analizi poslovanja. V skladiščih so dejstva implementirana v osrednjih tabelah, v katerih so vsi shranjeni podatki numerični. Zapis v dejstvu je sestavljen iz dveh delov, meritev ter večdelnega ključa. Ti tuji ključki so kategorije, po katerih analiziramo podatke, mere pa predstavljajo vrednost ali velikost podatka, ki ga analiziramo. Najuporabnejše so

aditivne numerične mere, katerih vrednost se v odvisnosti od dimenzij spreminja. Omogočajo nam poizvedbe, ki iz številnih zapisov pripravijo seštevke ter omogočajo primerjave. Nivo podrobnosti v tabeli dejstev moramo določiti pri načrtovanju dimenzijskega modela. Primer dejstva je lahko kreditna izpostavljenost banke (slika 3)

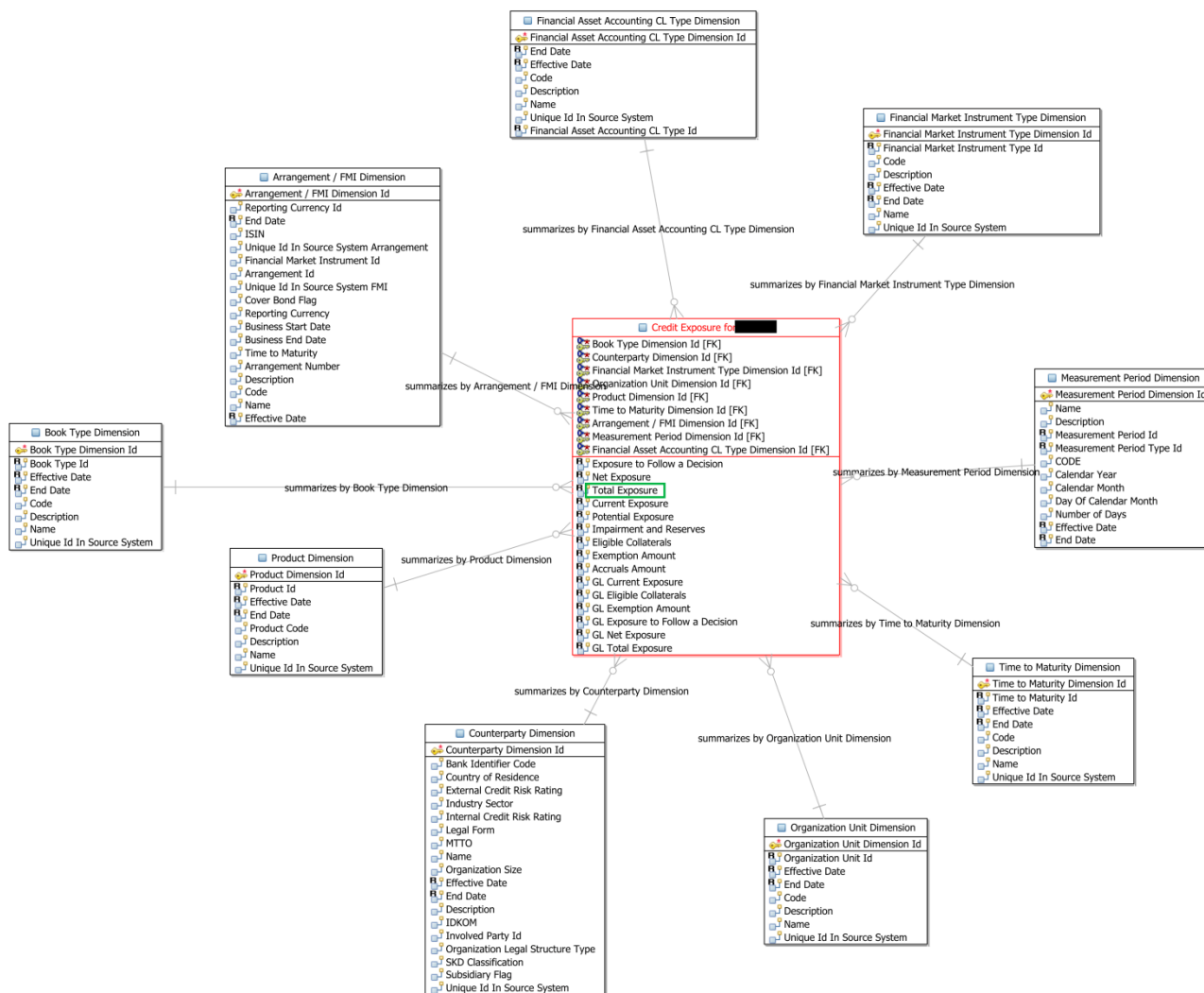
- Dimenzije

Dimenzije so zbirke članov ali enot iste vrste pogledov. V diagramu so običajno predstavljene kot osi. V dimenzionalnem modelu je vsak zapis v dejstvih povezan z natanko enim zapisom enega tipa dimenzije. To pomeni, da dimenzije določajo kontekstno ozadje dejstev. Dimenzije se običajno preslikajo v nenumerične subjekte kot so komitenti, zaposleni, pogodbe, produkti, organizacijske enote, ... (Slika 3)

- Mere

Mere so numerični atributi dejstev, ki predstavljajo relativno vedenje dimenzije. Dejanske številke so imenovane spremenljivke. Na primer, mera je prodaja izražena v denarju, količina prodaje, vrednost transakcije in tako naprej. Mere so določene s kombinacijami členov dimenzije, ki se nahajajo v dejstvih.

Primer: (Slika 3) Total exposure v dejstvu credit exposure je mera.



Slika 3: Dejstvo kreditne izpostave s pripadajočimi dimenzijami.

4.3. Postopek ETL (Extract, Transform, Load)

Namen skladiščenja podatkov je sprememba mase razpršenosti podatkov v koristne informacije. Postopki ETL omogočajo smiselno konverzijo in integracijo transakcijskih podatkov v obliko primerno za tvorbo podatkovnih skladišč in podporo orodjem za poslovno odločanje.

Proces ETL je sestavljen iz treh delov:

- Ekstrakcija podatkov iz različnih izvornih sistemov. (Extract)
- Transformiranje podatkov v koristne informacije. (Transform)
- Smiselno nalaganje podatkov v skladišče. (Load)

Ekstrakcija, transformiranje in nalaganje podatkov zajemajo vsa področja pridobivanja in shranjevanja podatkov. To so procesi, ki pokrivajo zajem podatkov iz virov, vključujejo vse funkcije, procedure in spreminjanje podatkov v enotno obliko primerno za shranjevanje v skladišče ter končno tudi fizično nalaganje v skladišče.

Ko želimo pretvoriti podatke iz vira v poslovne informacije v podatkovnem skladišču, ima vsak od naštetih ETL postopkov pomembno vlogo.

4.3.1. Ekstrakcija

Ekstrakcija je faza, katere namen je ustvarjanje ločene zbirke podatkov, ki so ločeni od izvornega okolja na nek časovni presek. Ta zbirka podatkov naj bi bila na ločenem sistemu podatkovnega skladišča in naj ne bi spreminjala vsebine. V bistvu gre za sliko izvornega oziroma transakcijskega sistema. Zaradi lažje in hitrejše obdelave se v tej fazi prenesejo le tisti podatki, ki so za podatkovno skladišče pomembni. Ko imamo podatke shranjene na ločen sistem, nas ne zanima več, kaj se dogaja na izvornih sistemih. S tem smo se tudi izognili težavam preobremenjevanja transakcijskih sistemov.

Seznam vprašanj v zvezi z ekstrakcijo [11]:

- Identifikacija virov (in njihove strukture)
- Metode ekstrakcije (za vsak vir moramo definirati, če je zajem ročni oz. vršen z orodjem)
- Frekvenca ekstrakcije (določiti moramo frekvenco zajema: dnevno, tedensko, mesečno,...)
- Časovno okno (koliko časa zajem traja)
- Urnik izvajanja (kdaj naj se določeni zajemi zaženejo)
- Lovljenje napak (določiti kaj v primeru, če podatkov ne moremo zajeti)

4.3.2. Transformiranje

Ko imamo podatke na ločenem sistemu, jih lahko poljubno preoblikujemo, ne da bi s tem kakorkoli ogrozili delovanje transakcijskih sistemov. V fazi transformacije spreminjamo obliko podatkov v skladu s predpisi in standardi, ki so določeni v podatkovnem skladišču. Tem podatkom dodajamo nadomestne ključe (surrogate key) za nadaljnje lažje združevanje. V tej fazi tudi odstranimo podvojene zapise ter morebitne napačne zapise popravimo.

Glavne naloge transformiranja [11]:

- Izbor – v tej nalogi izberemo podatke oziroma njihove dele, ki jih potrebujemo. Načeloma izločimo podatke, ki jih ne potrebujemo, vendar jih nismo znali izločiti v fazi ekstrakcije.
- Cepitev/Združevanje – V tej nalogi združimo ali cepimo podatke, ki so namenjeni za preoblikovanje ali pretvorbo. Načeloma združimo podatke iz več virov.
- Pretvorba – V tej nalogi vključujemo najrazličnejše pretvorbe atributov, da naredimo polja bolj uporabna in razumljiva uporabnikom ter da jih poenotimo oz. standardiziramo.
- Računanje – Kadar ni izvedljivo niti smiselno obdržati podatkov na najnižji ravni detajlov nastopi naloga računanja. Nekatere podatke je bolj smiselno imeti na primer na dnevni ravni, zato jih sumiramo.
- Obogatitev – Ta naloga je preureditev in poenostavitev posameznih področij za boljšo uporabnost v podatkovnem skladišču. Lahko uporabimo eno ali več področji iz istega vhodnega zapisa, da ustvarimo boljši pregled podatkov iz podatkovnega skladišča.

4.3.3. Nalaganje

Celoten proces premikanja in osveževanje podatkov v podatkovno skladišče imenujemo nalaganje.

Poznamo tri vrste nalaganja podatkov v skladišče [11]:

- Inicialno nalaganje (polnjenje tabel prvič)
- Postopno nalaganje (nalaganje sprememb)
- Polno osveževanje (popolnoma izbrišemo vsebino tabele in osvežimo z novimi podatki).

Nalaganje podatkov v skladišče je izredno dolgotrajno, zato povzroča veliko skrbi. Med nalaganjem mora biti podatkovno skladišče izključeno in neoperativno.

5. REALIZACIJA VREDNOTENJA KREDITNE IZPOSTAVLJENOSTI V PODROČNO PODATKOVNO SKLADIŠČE TER IZDELAVA POROČIL

5.1. Polnjenje tabel z ETL orodjem IBM DataStage

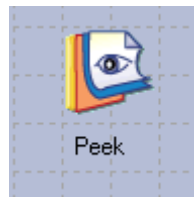
Ko imamo enkrat dobro izdelan dimenzijski podatkovni model, je čas za polnjenje tabel s postopki ETL. V diplomskem delu bomo predstavili IBM-ov program InfoSphere DataStage, ki je namenjen prav razvijanju postopkov po konceptu ETL. Gre za orodje, ki za izgradnjo ETL postopkov uporablja vizualni zapis. Orodje omogoča povezovanje podatkov iz več sistemov vzporedno in je optimizirano za procesiranje velike količine podatkov. Omogoča branje podatkov iz različnih virov (DB2, Oracle, csv, txt...) v podatkovno skladišče, kjer jih lahko nato preoblikujemo, zbiramo, preverjamo, odkrivamo napake, združujemo,...ipd. Pri izdelavi ETL-a imamo v orodju na voljo kar nekaj koristnih gradnikov (stage-v), lahko pa si jih poljubno zgradimo tudi sami. Opisali bomo nekaj osnovnih gradnikov, ki se tudi večinoma uporabljajo. Gradnike in osnovno masko je mogoče videti na sliki 4.

Osnovni gradniki :

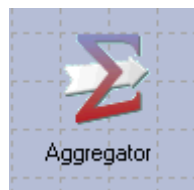
- **Transformer** je gradnik, ki ponuja raznolik nabor funkcij. V njem je mogoča pretvorba med različnimi podatkovnimi tipi, omogoča uporabo logičnih funkcij, filter, obravnavo NULL vrednosti, možna je tudi uporaba matematičnih funkcij. Poleg vsega naštetega lahko definiramo in uporabljamo lokalne spremenljivke, ter raznorazne funkcije, ki smo jih sami kreirali.



- **Peek** je gradnik, ki se uporablja predvsem v razvojnem okolju za odkrivanje napak ter pregled dobljenega izhoda.



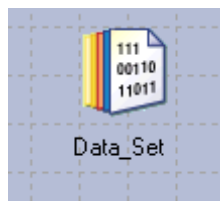
- **Aggregator** gradnik je namenjen štetju vrstic po ključu, ki mu ga določimo. V gradniku aggregator lahko uporabljamo tudi funkcije max, sum in podobne.



- **Funnel** je namenjen zlivanju podatkov iz različnih virov. Zlivanje je mogoče samo za attribute, ki so definirani v vseh virih z enakim podatkovnim tipom.



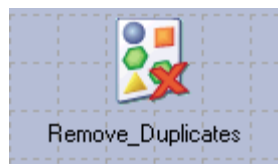
- **Data set** je datastage-va oblika shranjevanja vmesnih rezultatov. V bistvu gre za interne datoteke podobne csv-jem.



- **Slowly Changing Dimension (SCD)** je namenjen zaznavanju razlik med podatki, ki so na izvoru in podatki, ki so že v podatkovnem skladišču. Tukaj imamo na voljo SCD1, ki je namenjen spreminjanju starih zapisov, in pri katerem se ne vodi zgodovina, pri SCD2 se vodi zgodovina, kar pomeni, da se novi zapis odpre z današnjim dnem, stari zapis pa se zapre z enim dnevom nazaj. V gradniku SCD imamo tudi možnost definiranja nadomestnega oz. surogatnega ključa.



- **Remove duplicates** je namenjen odstranjevanju dvojnikov po danem ključu.



- **Sort** je namenjen sortiranju zapisov padajoče ali naraščajoče po atributih kot mu določimo.



- **Sequential file** je namenjen branju ter zapisovanju v CSV ali TXT obliki.

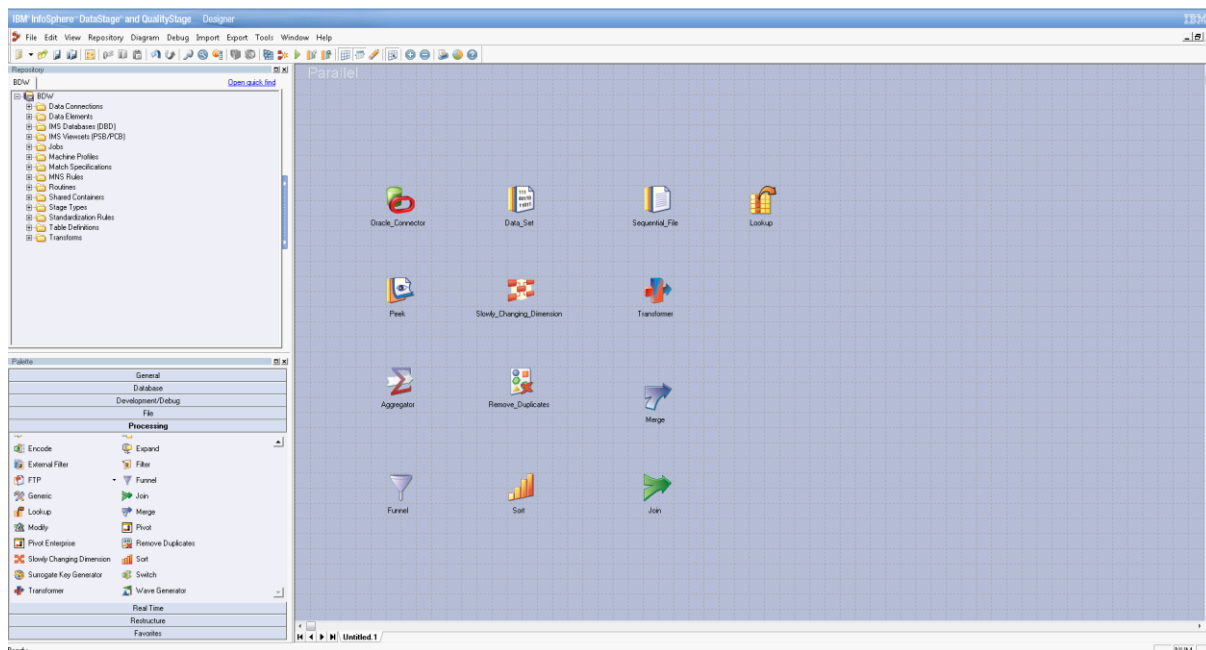


- **Join** in **Lookup** sta namenjena združevanju podatkov različnih tabel po ključu. Join je počasnejši, vendar se izkaže pri veliki količini zapisov. Lookup pa ima to prednost, da lahko v istem gradniku združimo več tabel hkrati. Kjer potrebujemo združitev samo dveh tabel, navadno uporabimo join, pri več tabelah pa lookup.



- **Merge** je namenjen združevanju podatkov, ki imajo isti ključ. Pri merge-u moramo paziti, da nimamo več zapisov z istim ključem, saj nam bo v tem primeru enega naključno zavrgel.

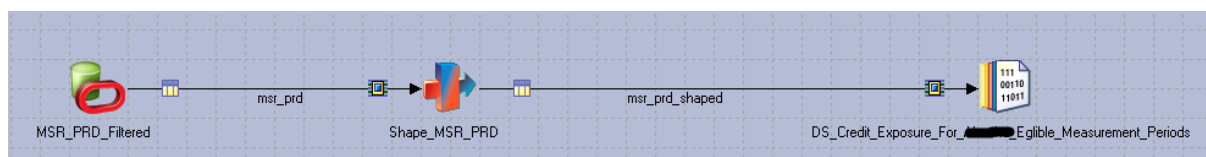




Slika 4: DataStage osnovna maska z osnovnimi gradniki

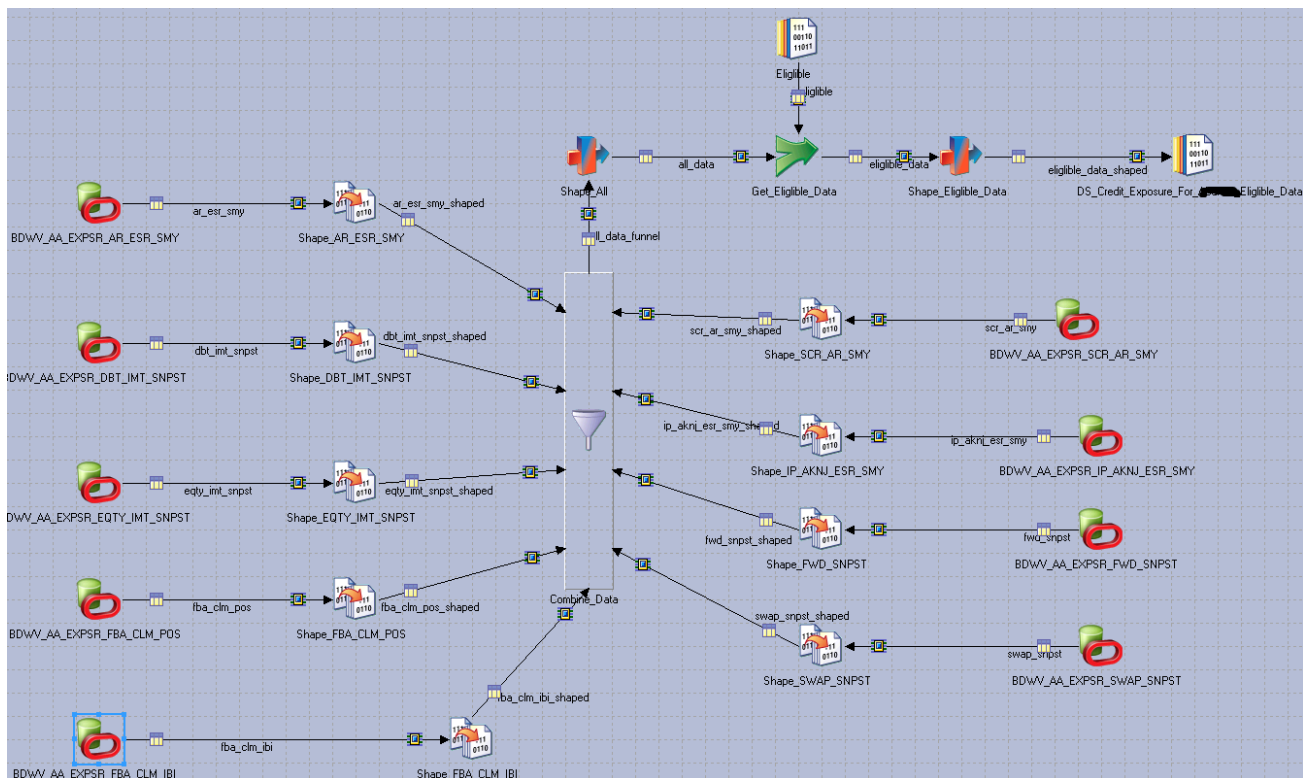
5.1.1. Primer polnjenja dejstva za kreditno izpostavljenost s programom DataStage

Na sliki 5 imamo opravilo (job), v katerem najprej preberemo iz tabele MSR_PRD obdobje, za katerega bomo polnili dejstvo, ga oblikujemo v ustrezno obliko in shranimo v dataset za poznejšo uporabo.



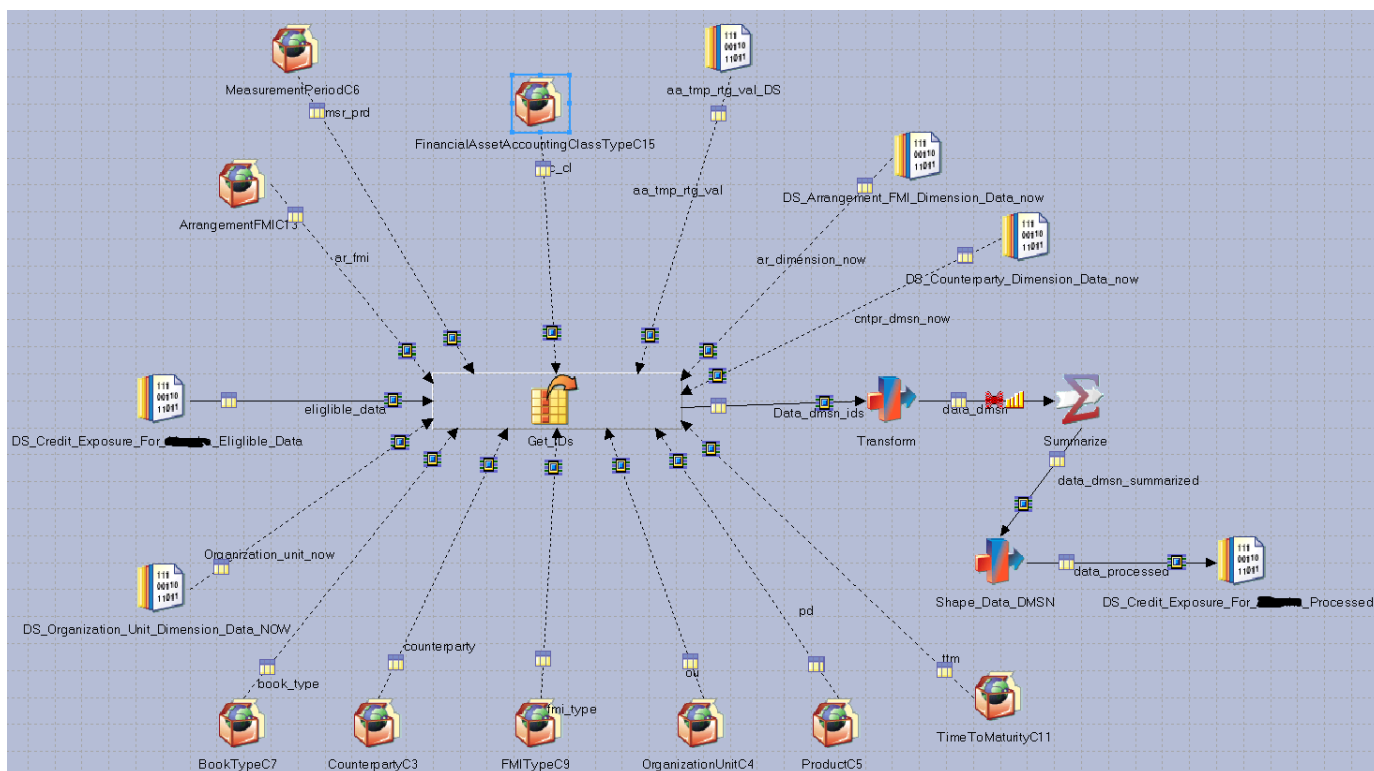
Slika 5: Opravilo za pridobitev časovnega obdobja

V drugem opravilu `credit_exposure_for_PreparedData` (Slika 6) preberemo pripravljene podatke iz sumarnih ter atomarnih tabel (priloga 2: poizvedba `BDWV_AA_EXPSR_AR_ESR_SMY` iz sumarne tabele) podatke iz pogodb, depozitov, portfelja delnic, terjatev,... in jih združimo v en izhod za lažjo obdelavo z stage-m funnel. Podatke nato združimo (inner join) z obdobjem, ki smo ga dobili v prejšnjem opravilu ter jih nato začasno shranimo v dataset (`DS_Credit_Exposure_For_Eligible_Data`). V transformerju (`shape_eligible_Data`) prej poskrbimo, da v podatkih ni praznih vrednosti.



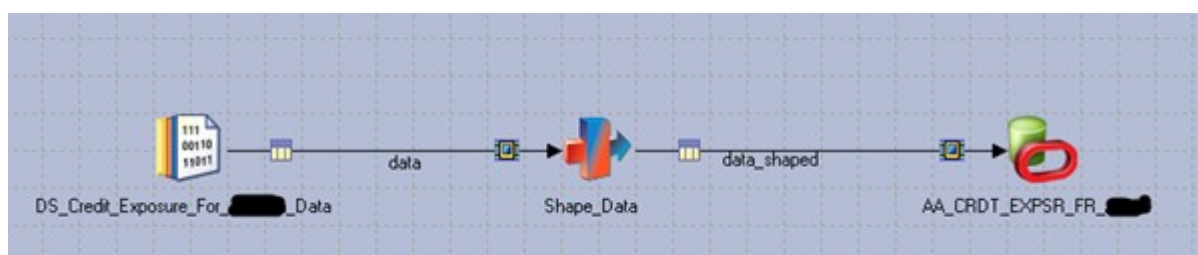
Slika 6: Opravilo za pridobitev podatkov iz sumarnih in atomarnih tabel

Opravilo `Credit_Exposure_for_Process_Data` (Slika 7) je namenjeno pridobivanju ključev dimenzij. V opravilu pridobimo ključne dimenzije za osebe (Counterparty), organizacijsko enoto (OrganizationUnit), produkta (Product), finančnega instrumenta (FMIttype), pogodbe (Arrangement),... V primeru, da za določeno vrednost ne dobimo veljavnega ključa dimenzij za to obdobje, poskrbimo, da vzamemo za to vrednost zadnjega veljavnega (Transform). Po teh ključih izračunamo v stage-u Summarize skupno vsoto različnih vrst izpostav (trenutna izpostavljenost, potencialna izpostavljenost, ...).



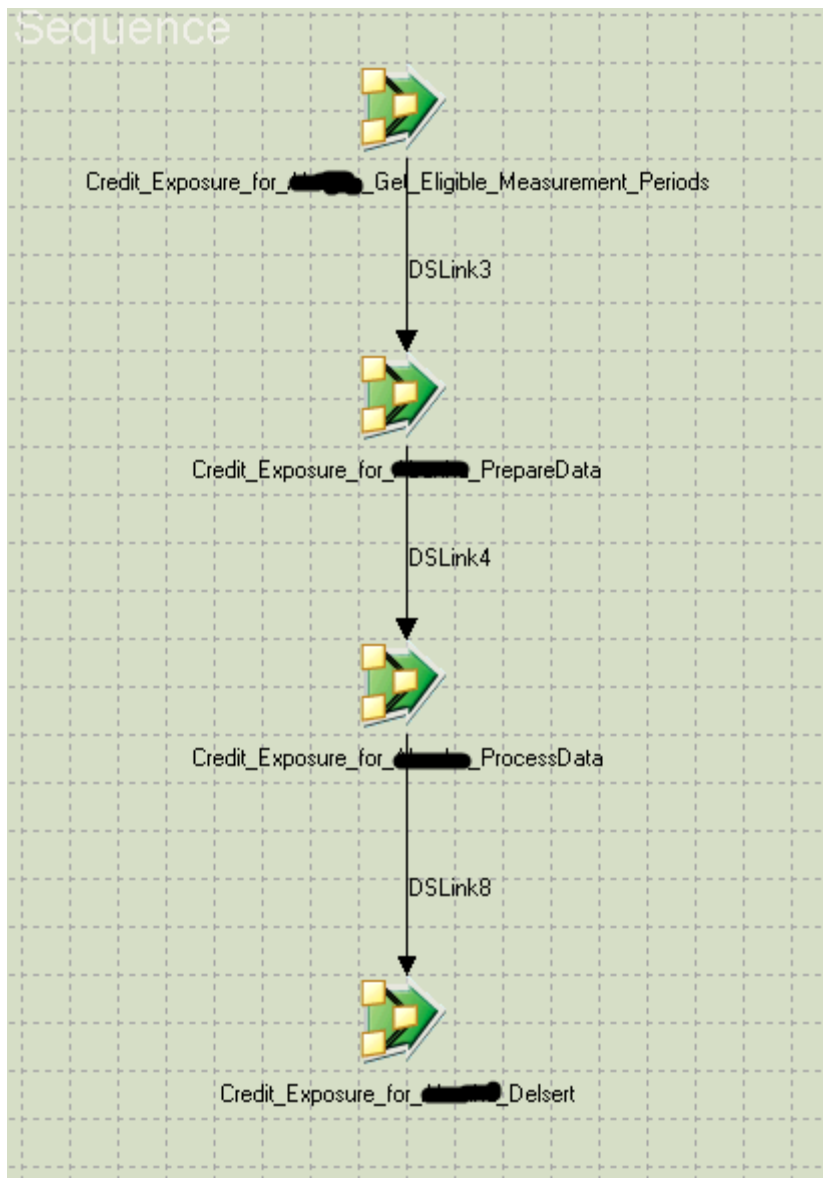
Slika 7: Opravilo za pridobitev ključev iz dimenzij

V zadnjem opravilu `Credit_exposure_for_Insert` (slika 8) shranimo podatke v osnovno dejstvo za kreditno izpostavljenost. Pred tem podatke, če za to obdobje že obstajajo, izbrišemo.



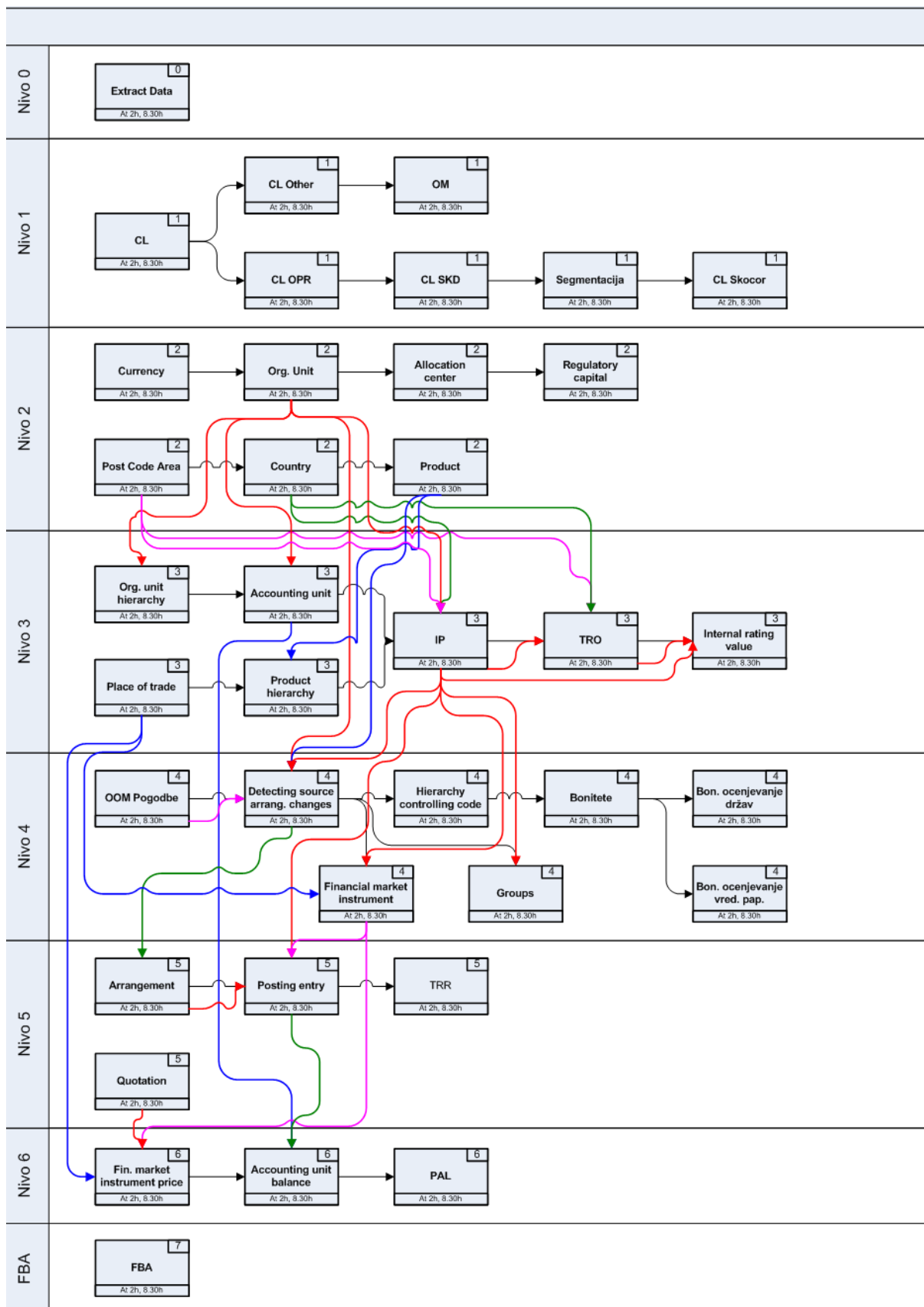
Slika 8: Opravilo za insert v tabelo dejstva

Vse skupaj v sekvenci povezuje (slika 9) krovno opravilo, ki tudi določa vrstni red izvajanja.

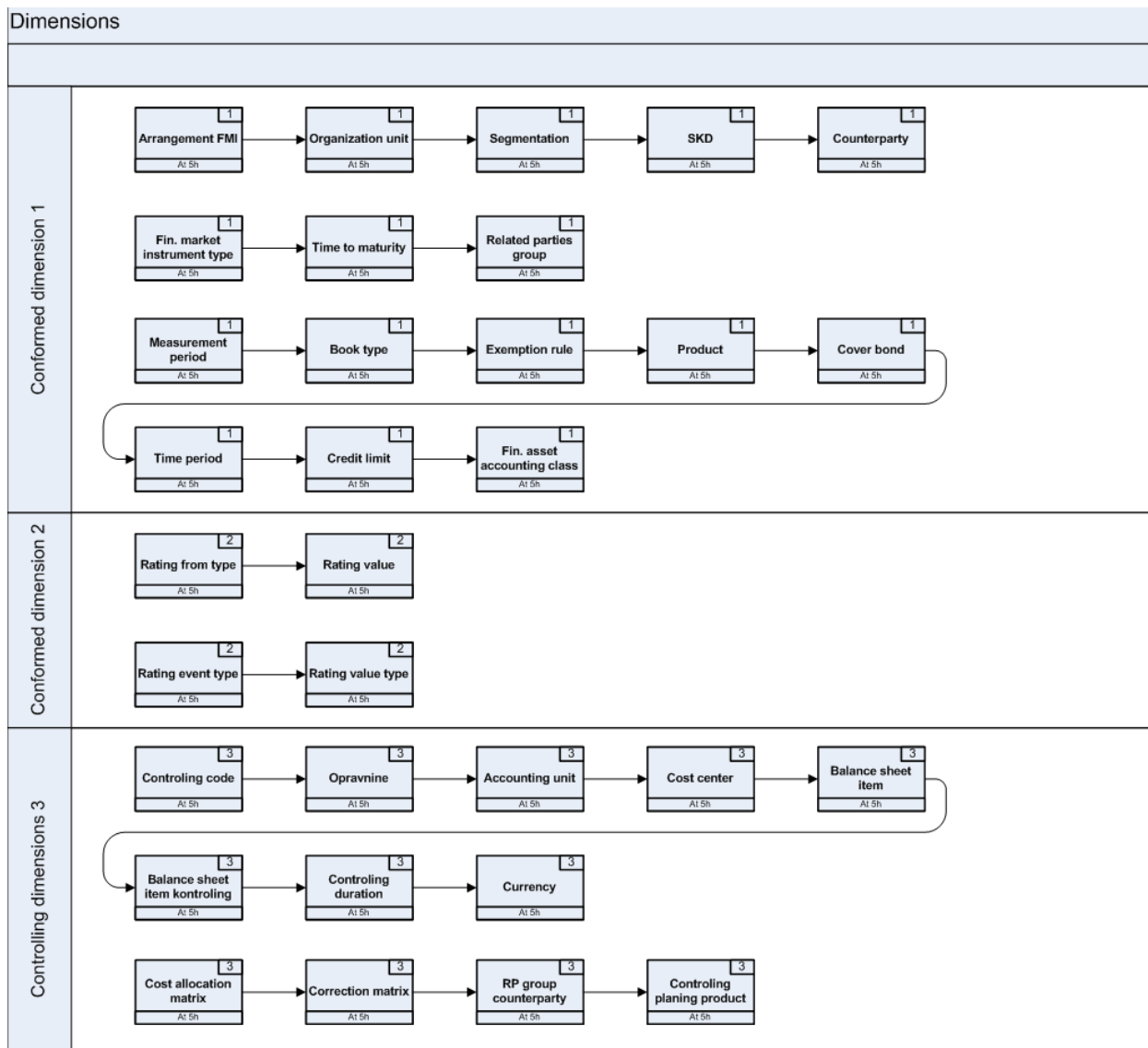


Slika 9: Krovno opravilo za polnjenje dejstva kreditnega tveganja

Ko govorimo o polnjenju dejstva za kreditno izpostavljenost, je potrebno predhodno napolniti vse atomarne, sumarne ter dimenzijske tabele. Takih opravil se v srednje veliki oz. veliki banki v Sloveniji nabere med 500 in 1000 in to samo, ko govorimo o polnjenju enega dejstva s pripadajočimi dimenzijami in atomarnimi tabelami. Res pa je, da je mogoče napolnjene tabele pozneje uporabljati tudi za druga področja v banki.



Slika 10: Zaporedje polnjenja virov na atomarnem nivoju v podatkovno skladišče



Slika 11: Zaporedje polnjenja dimenzij v podatkovnem skladišču

Zgornje slike (10 in 11) prikazujejo zaporedja polnjenja virov ter dimenzij v bančno podatkovno skladišče. Vsak kvadraterk predstavlja en vir oz. eno dimenzijo, v kateri se skriva od 5 – 10 kreiranih opravil z programom Datastage.

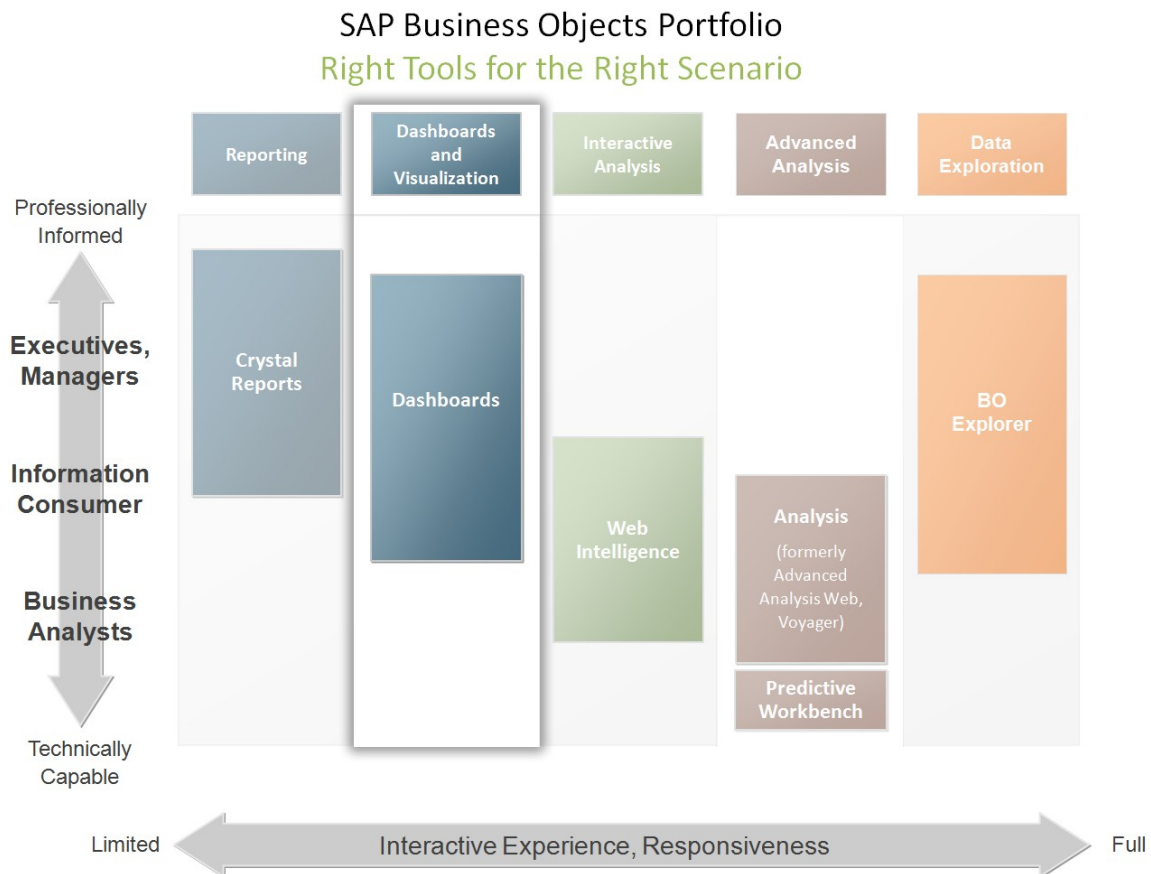
5.2. Izdelava poročil iz podatkovnega skladišča za pomoč pri odločanju ter poročanju

Ko so podatke napolnjeni v podatkovnem skladišču, nam brez pravih poizvedb še nič ne povedo. Poizvedbe je potrebno narediti tako, da nam pridobljeni podatki, kaj povedo, da jih znamo uporabiti banki v prid. Poznamo kar nekaj orodij za izdelovanje poročil (IBM Cognos BI, Birst, Oracle Reports, SAP BusinessObjects BI,...)[24].

Ker je SAP BusinessObject Bi eno izmed najbolj razširjenih orodij, ga bomo v naslednjem poglavju na kratko predstavili.

5.2.1. SAP BusinessObjects BI

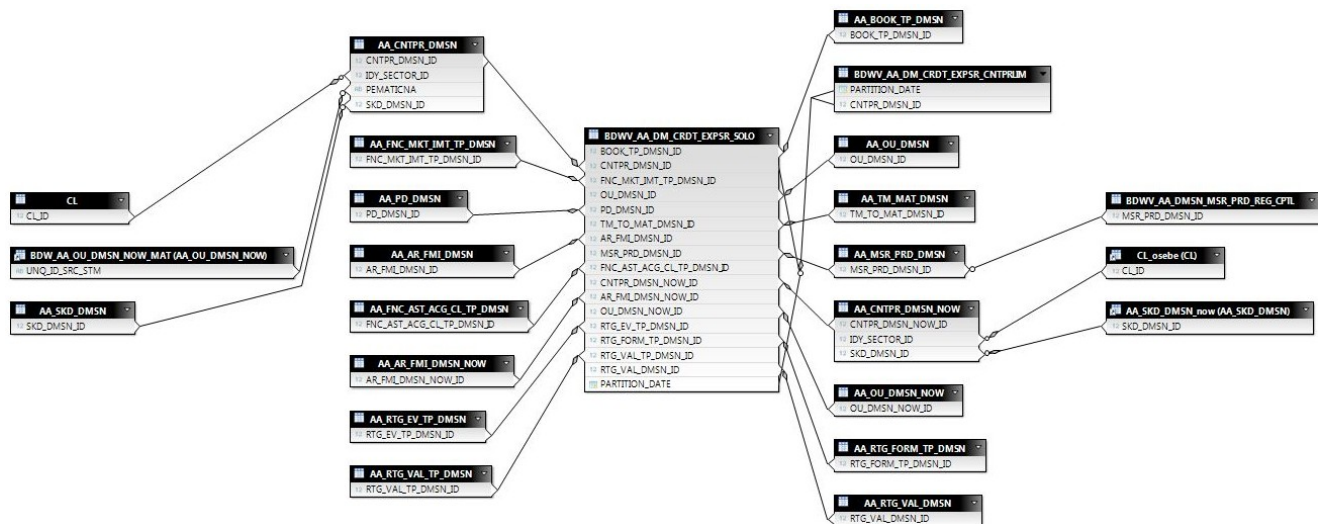
Orodje SAP BusinessObjects BI je celovit programski paket za hitrejše, zanesljivejše ter kakovostnejše odločanje. Vsi v organizaciji so med seboj povezani preko prilagojenega in poenotenega ogrodja, v katerem so integrirane vse poslovne funkcije. Poslovnim uporabnikom je s tem paketom orodij omogočen celoten spekter funkcionalnosti. Mogoče je izdelovati napredne analize, vizualizacije, interaktivne analize, raziskovanje podatkov ter poročanje [13].



Slika 12: Rešitve v Business Object BI [13]

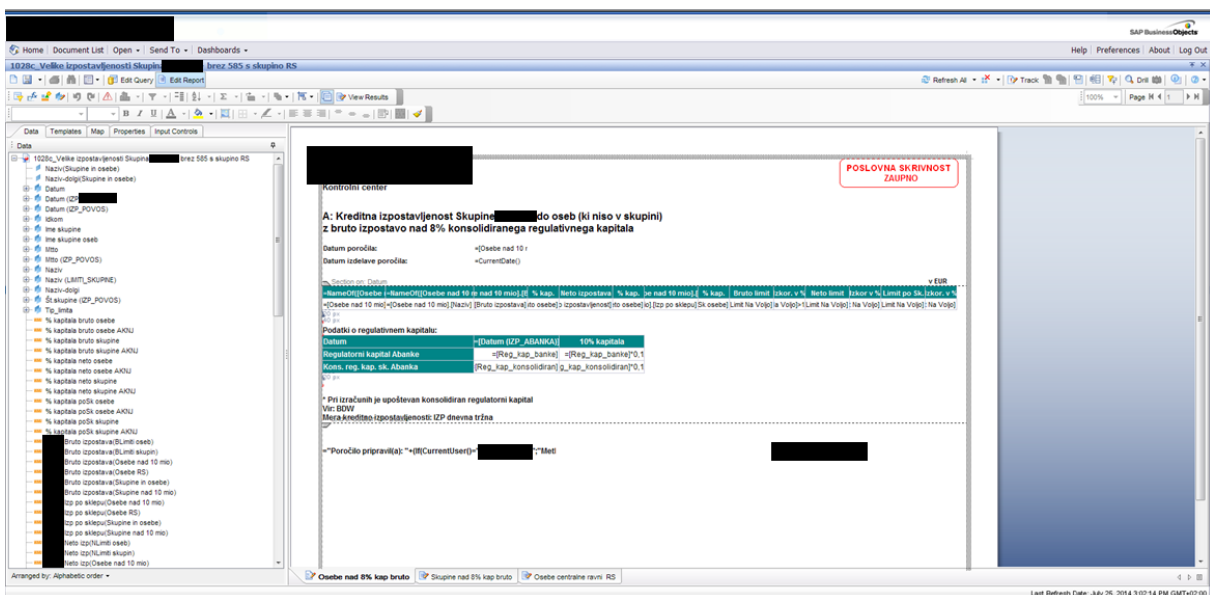
5.2.2 Business Objects Universe

Business Objects Vesolje (Universe) je v bistvu priprava podatkov med podatkovnim skladiščem in poslovnim uporabnikom. Je poslovna predstavitev podatkov skladišča ali transakcijske baze. Uporabnikom omogoča, da dostopajo do podatkov, brez da bi poznali zapletenost ter prepletenost tabel in njihovih relacij v podatkovnem skladišču, niti jim ni potrebno vedeti, kje so ti podatki shranjeni. Vesolja so ustvarjena v znani poslovni terminologiji za opisovanje poslovnega okolja in omogočajo uporabniku natančno pridobivanje podatkov, ki jih zanima. Vesolja, ki so ustvarjena, morajo biti primerno izdelana za potrebe končnega uporabnika. Najpomembnejši vidik vesolja je, da je enostaven. Vseboval naj ne bi predmetov, ki so nepotrebni ali ki bi lahko povzročili zmedo med končnimi uporabniki. Imena objektov morajo biti jasno in ustrezno navedena, da je nedvoumno, kateremu končnemu poslovnemu uporabniku so namenjena [23].



Slika 13: Primer izdelanega vesolja za kreditno izpostavljenost

Izdelana vesolja so poslovnim uporabnikom v veliko pomoč, saj iz njih dokaj enostavno in hitro izdelajo raznorazna poročila.



Slika 14: Osnovna maska Business Object Web intelligence z pripravljenim poročilom

Slika 14 predstavlja osnovno masko Business Object-a Web Intelligence, v katerem imamo v naprej pripravljena in oblikovana poročila, ki si jih navadno sami uporabniki, vsaj napredni, pripravijo. Nato samo izbirajo željeni datum, za katerega želijo poročilo, ter vklapljuje razne omejitve ter dodajajo pogoje. Izdelano poročilo si bolj natančno lahko pogledamo na Sliki 15.

Gre za poročilo kreditne izpostavljenosti skupine do skupine povezanih oseb z bruto izpostavo nad 8 odstotkov konsolidiranega regulativnega kapitala. Poročilo je bilo izdelano na

dan 27.07.2014 za 25.07.2014. Na ta dan je imela banka bruto izpostavljenost večjo od 8 odstotkov konsolidiranega regulativnega kapila do 21 skupin. V stolpcih lahko vidimo odstotek izpostavljenosti do teh skupin ter bruto, neto ter izpostavo po sklepu. Podatek imamo tudi o neto limitu in njegovi izkoriščenosti. Razberemo lahko, da ima banka do nekaterih skupin že zelo kritično oziroma rizično izpostavljenost (glede na to, da ima banka 197.466.027 regulatornega kapitala, vidimo, da je skupina 3 zelo tvegana, saj je odstotek izpostavljenosti skupine 3 proti kapitalu banke kar 32-odstotni). Tako poročilo in še veliko drugih podobnih poročil morajo banke v Sloveniji dnevno pošiljati Banki Slovenije.

Navadno napredni uporabniki izdelujejo poročila z Business Objectom kar sami, saj najboljše poznajo podatke, zahteve ter najbolj vedo, kaj želijo kot rezultat. Ko enkrat obvladujejo in dovolj dobro in podrobno poznajo podatke, je sama izdelava poročila zelo enostavna in hitra. Izdelavo preprostih poročil lahko merimo tudi v minutah.

Kontrolni center

B: Kreditna izpostavljenost Skupine xy banke do skupine povezanih oseb z bruto izpostavo nad 8% konsolidiranega regulativnega kapitala

Datum poročila: 23.7.2014

Datum izdelave poročila: 25.07.2014

v EUR

Št.skupine	Ime skupine	Bruto izpostava	% kap.	Neto izpostava	% kap.	Izp. po Sklepu	% kap.	bruto limit	zkor. v %	Neto limit	zkor. v %	Limit po Sk	zkor v %	Op.
1	Skupina 1	81.386.484,14	41,38%	27.207.740,09	13,83%	27207740,09	13,83%	30.000.000	90,69%	30.000.000	90,69%			
2	Skupina 2	75.381.632,71	38,33%	75.381.498,06	38,33%	375680,49	0,19%	80.000.000	94,23%	30.000.000	1,25%			**
3	Skupina 3	67.209.632,63	34,17%	66.927.833,63	34,03%	63158763,63	32,11%	47.000.000	#####	47.000.000	#####			*1
4	Skupina 4	52.596.652,92	26,74%	31.941.064,72	16,24%	30519913,37	15,52%	37.000.000	86,33%	35.000.000	87,20%			
5	Skupina 5	45.603.328,03	23,19%	13.546.689,18	6,89%	13546689,18	6,89%							
6	Skupina 6	43.819.143,75	22,28%	36.920.269,58	18,77%	26858831,44	13,66%	51.000.000	72,39%	42.000.000	63,95%			
7	Skupina 7	42.618.340,01	21,67%	8.847.038,52	4,50%	8847038,52	4,50%							
8	Skupina 8	41.712.809,79	21,21%	39.904.098,48	20,29%	39720869,18	20,20%	42.000.000	95,01%	42.000.000	94,57%			
9	Skupina 9	39.209.261,87	19,94%	31.600.968,50	16,07%	31490968,51	16,01%	33.000.000	95,76%	33.000.000	95,43%			
10	Skupina 10	38.508.939,51	19,58%	23.436.896,18	11,92%	23428108,45	11,91%	27.000.000	86,80%	27.000.000	86,77%			
12	Skupina 12	35.751.072,97	18,18%	10.986.754,10	5,59%	10986754,1	5,59%							
13	Skupina 13	35.187.552,26	17,89%	20.367.560,81	10,36%	20293274,72	10,32%	21.000.000	96,99%	21.000.000	96,63%			
14	Skupina 14	35.049.628,52	17,82%	34.847.802,69	17,72%	32865908,13	16,71%	46.000.000	75,76%	46.000.000	71,45%			
15	Skupina 15	33.722.655,22	17,15%	21.633.614,96	11,00%	20955673,93	10,65%	25.000.000	86,53%	25.000.000	83,82%			
17	Skupina 17	31.093.519,96	15,81%	5.505.178,65	2,80%	5457288,65	2,77%							
18	Skupina 18	30.433.528,71	15,47%	30.433.528,71	15,47%	30433528,71	15,47%	45.000.000	67,63%	45.000.000	67,63%			
19	Skupina 19	30.100.496,33	15,30%	13.455.040,84	6,84%	12687227,06	6,45%							
20	Skupina 20	28.725.767,37	14,61%	9.321.555,76	4,74%	9321555,76	4,74%							
21	Skupina 21	28.017.943,26	14,25%	15.239.633,59	7,75%	8714687,06	4,43%							

Podatki o regulativnem kapitalu:

Datum	23.7.2014	10% kapitala
Regulatorni kapital xy banke	197.466.027,00	19.746.602,70
Kons. reg. kap. sk. xy banka	196.680.755,00	19.668.075,50

* Pri izračunih je upoštevan konsolidiran regulatorni kapital

Vir: Podatkovno skladišče

Mera kreditne izpostavljenosti: IZP dnevna tržna

Opombe:

** Pri spremiljavi izkoriščenosti neto limitov in limitov po Sklepu o veliki izpostavljenosti ni upoštevana izpostavljenost do Republike Slovenije oziroma centralne ravni države.

Slika 15: Poročilo za kreditno izpostavljenost skupine do skupine povezanih oseb z bruto izpostavo nad 8% konsolidiranega regulativnega kapitala

6. SKLEP

Projekt gradnje bančnega podatkovnega skladišča vsekakor prinese velike spremembe v organizacijo. Največ jih je seveda v informacijski tehnologiji, nekatere spremembe pa so tudi v organizaciji dela.

Izgradnja podatkovnega skladišča banke identificira nepravilnosti v podatkih, ki jih prej ni bilo mogoče na enostaven ter očitni način opaziti. V banki se kar na enkrat začne govoriti o kvaliteti podatkov, ki niso le zneski in konti. Kar naenkrat postanejo pomembne stvari, ki jih prej nismo opazili. Kar naenkrat potrebujemo ažurne informacije npr. ali je obrestni dodatek spremenljiv s sklepom banke ali ne. V sistemu morajo biti še isti dan vsi znani podatki. Za to je potrebno spreminjati procese. Zaradi izgradnje bančnega skladišča se začnejo v vseh oddelkih dogajati premiki. Vsi imajo več dela, vsi se morajo naučiti nekaj novega, bolj poglobljeno razmišljati. Poslovno prakso in vse poslovne procese mora banka identificirati in spremeniti na tak način, da bodo vsi postopki obdelav in knjiženj zaključeni še isti dan do konca dneva. Ta zahteva predstavlja veliko oviro v popolni implementaciji podatkovnega skladišča. V Sloveniji je to za enkrat skoraj nemogoče zaradi objektivnih razlogov, kot je na primer poslovna praksa knjiženja bankomatnih dvigov in drugih »off-line« transakcij, ki se knjižijo za nazaj.

Načeloma to na začetku na vsakega posameznika deluje zelo negativno. Uprava se mora začeti ukvarjati s problemi, ki jih prej niso imeli. Probleme morajo reševati direktorji, referenti pa se morajo šolati in delati zahtevnejša dela kot do sedaj. Gre za projekt, ki, dokler ni dokončan, nima veliko simpatizerjev. Ko pa so prvi rezultati vidni in ko uporabnik zagleda oz. dobi informacijo, o kateri je prej lahko le sanjal, se mu zasvetijo oči.

Uporabniki lahko pridejo do nekaterih osnovnih podatkov, ki jim pomagajo pri hitrem odločanju, tudi brez uvedbe podatkovnega skladišča. Seveda pa je možno z uvedbo podatkovnega skladišča pridobiti podrobnejše, bolj čiste in poenotene podatke veliko hitreje. Brez podatkovnega skladišča mora na zahtevo uporabnika izdelati poročilo oddelek IT. Če so podatki že primerni za poročilo, je potrebno samo izdelati poročilo, največkrat pa je potrebno podatke vnaprej pripraviti in izračunati v raznih aplikacijah in šele nato izdelati poročilo. Izdelava takih poročil na zahtevo se lahko meri v dnevih, za bolj zahtevna poročila tudi v tednih, če pa ta poročila izdelujejo zunanji izvajalci, pa lahko govorimo celo v mesecih.

Z uvedbo podatkovnega skladišča, ki naj bi vseboval »vse« podatke, ki naj bi jih uporabniki potrebovali ali bi jih mogoče lahko potrebovali, pa je stvar izdelave novih poročil stvar uporabnikov. Uporabniki sami izdelujejo poročila oz pridobivajo podatke, ki jih potrebujejo. Občasno vsekakor potrebujejo tudi pomoč IT oddelka.

Hitrost se po uvedbi podatkovnega skladišča lahko meri v minutah, saj je uporabnik tudi izvajalec, tako da ne prihaja do nesporazumov med zahtevami in izvedbo. Izključene seveda niso napake, ki so lahko posledica nepoznavanja pomena podatkov ali napačne uporabe orodij.

Razvoj podatkovnega skladišča je proces, ki ga želimo oz. moramo vpeljati v banko kot nekaj novega in stalnega, nekaj, kar ne bo izginilo, temveč bo kvečjemu raslo.

Potreba po izgradnji področnega podatkovnega skladišča za kreditno izpostavo izhaja predvsem iz pomanjkanja enotnega podatkovnega vira, ki bi imel ustrezno strukturirane, dovolj ažurne in knjigovodsko usklajene podatke o kreditni izpostavi, ki bi jih bilo mogoče uporabiti za različne namene in dobiti vedno pravilne podatke.

Prednost podatkovnega skladišča pred predhodnim sistemom zbiranja podatkov v banki vidimo predvsem v tem, da je v največji možni meri, glede na obstoječe in nove procese evidentiranja dogodkov, ki vplivajo na izpostavo, zagotovil celovitost informacij, ki jih nudi uporabnikom, ažurnost pa zagotavlja tudi z različnimi merami kreditne izpostave, ki se uporabljajo za različne namene poročanja. S samo logiko, ki je vgrajena v delovanje podatkovnega skladišča je moč zagotoviti ustrezno strukturiranost podatkov od mikro nivoja (konta), preko nivoja pogodb, komitenta in skupine povezanih komitentov, tako na ravni izpostave banke, kot na konsolidirani ravni bančne skupine. Njegova struktura in logika omogoča tudi relativno enostavno upravljanje s podatkovnimi viri in dodajanje novih.

Bančno podatkovno skladišče v čedalje večji meri postaja enotna točka poročanja o kreditni izpostavi, kar terja in bo v prihodnosti še terjal redno vzdrževanje in nadgrajevanje sistema. Možnosti izboljšav je možna predvsem v procesih polnjenja podatkov in obveščanju uporabnikov o težavah pri polnjenju podatkov, izboljšanju frekvenc pridobivanja podatkov nekaterih segmentov izpostav, npr. odvisnih družb za konsolidirano izpostavo na nivoju bančne skupine in navezavi čim večjega števila bančnih aplikacij direktno na podatke iz bančnega podatkovnega skladišča. Predvsem pa mora podatkovno skladišče v vseh bankah

prej ali slej prevzeti vlogo glavnega vira podatkov za namene poročanja zunanjim institucijam, kar bo povzročilo in zagotovilo tudi večjo kvaliteto podatkov v celotni banki.

7. LITERATURA IN VIRI

- [1] Alur, Nagraj. 2008. IBM InfoSphere DataStage Data Flow and Job Design, First edition.
- [2] Ballard, Chuck. 1998. Data Modeling Techniques for Data Warehousing. San Jose, California. IBM Corporation, international technical support organization
- [3] Bessis, Joel. 2002. Risk management in banking. New York: John Wiley & Sons, inc.
- [4] Inmon, W.H. 1996. Building the Data Warehouse. Second Edition. New York: John Wiley & Sons, Inc.
- [5] Inmon, W.H. 2002. Building the Data Warehouse. Third Edition. New York: John Wiley & Sons, Inc.
- [6] Karpe, Primož. 1997. Upravljanje bančnih tveganj.
- [7] Kimbal, Ralph and Roos, Margy. 2002 The data warehouse toolkit. Second edition. New York: John Wiley & Sons, Inc.
- [8] Lancaster , Richard. Data Warehouse, business Intelligence and Trends in Data Analytics. Dostopno na: <http://blog.lansa.com/application-modernization/business-process-automation/data-warehouses-business-intelligence-trends-data-analytics> [29.08.2014]
- [9] Leonard, Edward M. 2011. Design and implementation of an enterprise data warehouse. Mater's theses: Marquette university.
- [10] Pirc, Dušan. 2007. Podatkovna skladišča v mednarodnem podjetju. Magistrsko delo, Univerza v ljubljani, Ekonomska fakulteta.
- [11] Ponniah, Paulraj. 2010. Data Warehousing fundametals for it prefessionals. Hoboken,New Jersey: John Wiley & Sons, Inc.
- [12] Prašnikar, Domen. 2001. Bančno podatkovno skladišče. Bančni vestnik.
- [13] SAP Business Object. Dostopno na: <http://www.clariba.com/sap-technology/business-intelligence/sap-dashboards.php> [25.7.2014]

- [14] Terzič, Igor. Podatkovno skladiščenje v slovenskih bankah ali podatkovna zmeda po slovensko, v3, 7.4.05
- [15] Zakon o bančništvu(Zban-1), uradni list RS
- [16] Žagar, Iva. 1999. Boniteta kreditojemalca in vrste kreditnih rizikov(zbornik 5. Strokovnega posveta o bančništvu), uredil: N. Borak. Ljubljana: Zveza ekonomistov Slovenije
- [17] Bizzari, 2009. Cross Industry Standard Process for Data Mining, Dostopno na: http://en.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining [10.09.2014]
- [18] Kelly, Sean. 1997. Data warehousing in action. New York: John Wiley & Sons, Inc.
- [19] Koren, Rok. 2010. Polnjenje podatkovnih skladišč s pomočjo orodja oracle warehouse builder, Univerza v Ljubljani, Fakulteta za matematiko in fiziko
- [20] Golfarelli in Rizzi. 2009, Data warehouse design. Modern principles and Methodologies. Bologna: Seps
- [21] Kumar, Naveen, 2007, Inmon versus Kimball,Dostopno na: <http://www.scribd.com/doc/198173234/Inmon-Versus-Kimball> [10.09.2014]
- [22] IBM, 2014, Deriving Business Insight from Big Data in Banking,Dostopno na: <http://www-01.ibm.com/software/data/bigdata/industry-banking.html> [10.09.2014]
- [23] Sapperstein, Jill. 2009, What is Business Object Universe?,Dostopno na: <http://alteksolutions.com/wp/index.php/2009/03/what-is-a-business-objects-universe/> [10.09.2014]
- [24] Wikipedia, 2014, Business intelligence tools,Dostopno na: http://en.wikipedia.org/wiki/Business_intelligence_tools [23.09.2014]

INTERNI VIRI(Gradivo je zaupne narave, zato več podrobnosti ne smemo razkriti):

[25] Protokol za zagotavljanje kvalitete podatkov v podatkovnem skladišču. 2013

[26] Tehnološka specifikacija za izračun kreditne izpostavljenosti. 2012

[27] Dokumentacija podatkovnega skladišča. 2013

[28] General rules for ETL. 2011

8 . PRILOGE

V prilogi je poizvedba za poročilo (slika 15) kreirana z programom SAP BI Business Object Web intelligence za kreditno izpostavljenost ter poizvedba iz summarne tabele. Vsi podatki razvidni iz poizvedb in slik so izmišljeni, tako, da dejanska povezava z dejanskimi podatki ni smiselna.

```

SELECT /*+all_rows parallel(4)*/
BDW.AA_MSR_PRD_DMSN.FORM_DT,
BDW.AA_REL_PR_GRP_DMSN.CODE,
BDW.AA_REL_PR_GRP_DMSN.REL_PR_GRP_NM,
BDW.AA_CRDT_LMT_DMSN.TIPLIMITA,
BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.BRUTO_IZPOSTAVA,
BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.GL_TOTAL_EXPOSURE,
BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.LMT_BRUTO,
max(BDW.AA_CRDT_LMT_DMSN.ZNES_REDNI_LIMIT),
max(BDW.AA_CRDT_LMT_DMSN.ZNES_INDIVID),
max(BDW.AA_CRDT_LMT_DMSN.ZNESEK_POSOJA),
max(BDW.AA_CRDT_LMT_DMSN.ZNESEK_IZPOSOJA),
MAX(BDW.AA_CRDT_LMT_DMSN.LIMIT_NA_VOLJO)
FROM
BDW.AA_TM_PRD_DMSN,
BDW.AA_MSR_PRD_DMSN,
BDW.AA_REL_PR_GRP_DMSN,
BDW.AA_CRDT_LMT_DMSN,
BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL
WHERE
( BDW.AA_MSR_PRD_DMSN.MSR_PRD_DMSN_ID=BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.MSR_PRD_DMSN_ID )
AND ( BDW.AA_TM_PRD_DMSN.TM_PRD_DMSN_ID=BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.TM_PRD_DMSN_ID
)
AND (
BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.REL_PR_GRP_DMSN_ID=BDW.AA_REL_PR_GRP_DMSN.REL_PR_GRP_DMSN_ID )
AND (
BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.CRDT_LMT_DMSN_ID=BDW.AA_CRDT_LMT_DMSN.CRDT_LMT_DMSN_ID
)
AND
(
( BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.PARTITION_DATE IN ('23-07-2014 00:00:00') )
AND
BDW.AA_CRDT_LMT_DMSN.LVL IN ( 1 )
AND
BDW.AA_REL_PR_GRP_DMSN.CODE <> 'Ni definirano'
AND

```

```

BDW.AA_CRDT_LMT_DMSN.TIPLIMITA = 'BRUTO'

AND

BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.SOLO_SKUPINA = 'SKUPINA'

)

GROUP BY

BDW.AA_MSR_PRD_DMSN.FORM_DT,

BDW.AA_REL_PR_GRP_DMSN.CODE,

BDW.AA_REL_PR_GRP_DMSN.REL_PR_GRP_NM,

BDW.AA_CRDT_LMT_DMSN.TIPLIMITA,

BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.BRUTO_IZPOSTAVA,

BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.GL_TOTAL_EXPOSURE,

BDW.BDWV_AA_DM_CRDT_EXPSR_LMT_REL.LMT_BRUTO

HAVING

(

max(BDW.AA_CRDT_LMT_DMSN.ZNES_INDIVID) > 0

OR

max(BDW.AA_CRDT_LMT_DMSN.ZNES_REDNI_LIMIT) > 0

)

```

Priloga 1: poizvedba za kreditno izpostavljenost skupine do skupine povezanih oseb z bruto izpostavo nad 8% konsolidiranega regulativnega kapitala

/* Formatted on 13/08/2014 19:33:06 (QP5 v5.240.12305.39446) */

CREATE OR REPLACE FORCE VIEW BDWV_AA_EXPSR_AR_ESR_SMY

```
(
MSR_PRD_ID,
IP_ID,
AR_ID,
FNC_MKT_IMT_ID,
FNC_MKT_IMT_TP_ID,
OU_ID,
BOOK_TP_ID,
PD_ID,
FNC_AST_ACG_CL_TP_ID,
TM_TO_MAT_ID,
CURRENT_EXPOSURE,
GL_CURRENT_EXPOSURE,
IMPRMNT_AND_RSRVES,
ACCRUALS_AMOUNT,
POTENTIAL_EXPOSURE,
DDCN_CPTL_F,
PNP_DUE_AMT,
PNP_ALWC,
INT_RCVB_LBY_AMT,
INT_RCVB_DUE,
INT_ALWC,
ALLWC_RCVB_ORRE_FEE,
ALLWC_RCVB_OTHR_FEE,
ODPIS_NETRAJNI,
BRUTO_TRZNA,
BRUTO_KOMITENT
)
AS
SELECT /*+ALL_ROWS*/
MSR_PRD_ID,
SMY.IP_ID,
SMY.AR_ID,
CAST (NULL AS INTEGER) AS FNC_MKT_IMT_ID,
AR.FININSTRUMENT_PD_ID AS FNC_MKT_IMT_TP_ID,
OU_ID,
BOOK_TP_ID,
PD_ID,
FNC_AST_ACG_CL_TP_ID,
CAST (0 AS INTEGER) AS TM_TO_MAT_ID,
-- BRUTO TRŽNA IZPOSTAVA
CAST (
CASE
WHEN SMY.SRC_STM_ID = 65
THEN
FAIR_VAL_AMT
- ODPIS_NETRAJNI
+ INT_ALWC
+ ALLWC_RCVB_ORRE_FEE
+ ALLWC_RCVB_OTHR_FEE
+ PNP_DUE_AMT
+ INT_RCVB_DUE
ELSE
FAIR_VAL_AMT
- ODPIS_NETRAJNI
+ INT_ALWC
+ ALLWC_RCVB_ORRE_FEE
+ ALLWC_RCVB_OTHR_FEE
```

```

END AS NUMBER (17, 2))
AS CURRENT_EXPOSURE,
-- BRUTO IZPOSTAVA
CAST (
  CLS_TOT_CR_ESR_AMT
- ODPIS_NETRAJNI
+ INT_ALWC
+ ALLWC_RCVB_ORRE_FEE
+ ALLWC_RCVB_OTHR_FEE AS NUMBER (17, 2))
AS GL_CURRENT_EXPOSURE,
-- OSLABITVE IN POPRAVKI TER REZERVACIJE
CAST (- (PNP_ALWC + RSRVES_975) AS NUMBER (17, 2))
AS IMPRMNT_AND_RSRVES,
ACR_AMT AS ACCRUALS_AMOUNT,
CAST (0 AS NUMBER (17, 2)) AS POTENTIAL_EXPOSURE,
CAST (0 AS INTEGER) AS DDCN_CPTL_F,
CAST (PNP_DUE_AMT AS NUMBER (17, 2)) AS PNP_DUE_AMT,
CAST (PNP_ALWC AS NUMBER (17, 2)) AS PNP_ALWC,
CAST (INT_RCVB_LBY_AMT AS NUMBER (17, 2)) AS INT_RCVB_LBY_AMT, --
CAST (INT_RCVB_DUE AS NUMBER (17, 2)) AS INT_RCVB_DUE, --
CAST (INT_ALWC AS NUMBER (17, 2)) AS INT_ALWC,
CAST (ALLWC_RCVB_ORRE_FEE AS NUMBER (17, 2)) AS ALLWC_RCVB_ORRE_FEE,
CAST (ALLWC_RCVB_OTHR_FEE AS NUMBER (17, 2)) AS ALLWC_RCVB_OTHR_FEE,
CAST (ODPIS_NETRAJNI AS NUMBER (17, 2)) AS ODPIS_NETRAJNI, --
CAST (FAIR_VAL_AMT - ODPIS_NETRAJNI AS NUMBER (17, 2))
AS BRUTO_TRZNA,
CAST ((FAIR_VAL_AMT - ACR_AMT) AS NUMBER (17, 2))
AS BRUTO_KOMITENT
FROM (SELECT MSR_PRD_ID,
  IP_ID,
  AR_ID,
  OU_ID,
  BOOK_TP_ID,
  PD_ID,
  FNC_AST_ACG_CL_TP_ID,
  SRC_STM_ID,
  FAIR_VAL_AMT,
  ODPIS_NETRAJNI,
  INT_ALWC,
  ALLWC_RCVB_ORRE_FEE,
  ALLWC_RCVB_OTHR_FEE,
  PNP_DUE_AMT,
  INT_RCVB_DUE,
  CLS_TOT_CR_ESR_AMT,
  PNP_ALWC,
  RSRVES_975,
  ACR_AMT,
  INT_RCVB_LBY_AMT,
  PARTITION_DATE,
  AR_ESR_SMY_TP_ID
FROM AR_ESR_SMY
WHERE PARTITION_DATE =
  (SELECT DATUM
  FROM VLD_VALID_DATE_TO_PROCESS
  WHERE TP_SRC IN (1, 2)
  AND TP_NAMEN = -2001
  AND SYSDATE BETWEEN EFF_DT AND END_DT
  AND UTLZ_F = 0)
UNION ALL
SELECT b.MSR_PRD_ID,

```

```

IP_ID,
AR_ID,
OU_ID,
BOOK_TP_ID,
PD_ID,
FNC_AST_ACG_CL_TP_ID,
SRC_STM_ID,
FAIR_VAL_AMT,
ODPIŠ_NETRAJNI,
INT_ALWC,
ALLWC_RCVB_ORRE_FEE,
ALLWC_RCVB_OTHR_FEE,
PNP_DUE_AMT,
INT_RCVB_DUE,
CLS_TOT_CR_ESR_AMT,
PNP_ALWC,
RSRVES_975,
ACR_AMT,
INT_RCVB_LBY_AMT,
PARTITION_DATE,
AR_ESR_SMY_TP_ID
FROM AR_ESR_SMY a,
(SELECT MSR_PRD_ID, DATUM
FROM VLD_VALID_DATE_TO_PROCESS
WHERE TP_SRC IN (1, 2)
AND TP_NAMEN = -2001
AND SYSDATE BETWEEN EFF_DT AND END_DT
AND UTLZ_F = 0) b
WHERE PARTITION_DATE IN (SELECT MAX (RPT_DT)
FROM FBA_HEADER_DT
WHERE SRC_STM_ID = 82)
AND SRC_STM_ID IN (117, 118)) SMY
LEFT JOIN AR ON (SMY.AR_ID = AR.AR_ID)
WHERE AR_ESR_SMY_TP_ID = 1
AND (FAIR_VAL_AMT <> 0 OR CLS_TOT_CR_ESR_AMT <> 0);

```

Priloga 2: poizvedba iz sumarne tabele pogodb za polnjenje dejstva za kreditno izpostavljenost