

Integralno podatkovno skladišče za slovenske banke

*Igor Terzič

This article is about data warehousing in Slovenian banking industry. It describes the nature and the design of a basic integral banking data warehouse for the Slovenian environment. It gives a definition and describes the confusion about it. It defines the distinction between banking transactional and analytical systems. Based on examples, basic elements of an integral solution are defined. The structure of the Balance Sheet, P&L, General Ledger, banking price list and market data is given. The Buy-Or-Build dilemma is discussed and some initial dilemmas described.

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

Pričujoč prispevek se dotika teme, ki je za področje slovenskega bančništva relativno nova, predvsem pa zelo skrita očem javnosti. Projekti izgradnje bančnega podatkovnega skladišča potekajo v ozadju in za banko ne prinašajo očitnih ali bolje rečeno zvenceh dosežkov. Če je (bo) kateri od bank uspešno projekt speljati do katerekoli točke, se s tem lahko hvali kvečjemu v ozkem krogu posvečenih specialistov, ki vedo, kako pomembno je za banko, da ima svoje podatke organizirano pregledno in dostopno in kako težko je do tega priti.

* Igor Terzič se ukvarja s svetovanjem na področju bančne poslovne informatike. Ima enajstletne izkušnje na področju bančništva predvsem na področju poslovne informatike, upravljanja z bančnimi tveganji in upravljanja bank. Po izobrazbi je univerzitetni diplomirani inženir elektrotehnike in magister managementa. S problematiko podatkovnega skladiščenja v banki se ukvarja že od leta 1996, sprva v obliki integracije kupljenega sistema, nato preko izgradnje lastnega. Njegove rešitve omogočajo izvajanje večjega dela integralnega poročanja in upravljanja s tveganji. Dosegljiv je na elektronskem naslovu igor.terzic@guest.arnes.si.

Ta prispevek govori predvsem o strukturi integralnega podatkovnega skladišča, prispevek, ki bo objavljen naslednjič, pa bo opisoval praktične probleme njegove izgradnje in uporabe ter dileme, ki se odpirajo pri sprejemanju odločitev.

1. Definicija podatkovnega skladišča

Največji pisec na temo podatkovnega skladiščenja Ralph Kimball nam postreže z zelo enostavno definicijo, ki pa ne bi mogla bolj zadeti bistva: "Podatkovno skladišče je kopija podatkov o transakcijah, ki je posebej strukturirana za povpraševanje in analizo." Gre torej za podatkovno strukturo, ki je v svojem namenu posvečena popolnoma drugim ciljem, kot si običajno predstavljamo o tem, kaj je naloga informacijskega sistema.

Informacijske sisteme, ki niso podatkovna skladišča in jih v bankah v glavnem poznamo pod imenom poslovne aplikacije, teorija poimenuje transakcijski sistemi. Praviloma gre za računalniške programe, ki so namenjeni zajemanju in obdelavi poslovnih transakcij. Pogosto se označujejo s kratico OLTP, kar pomeni On-line Transactional Processing oz. procesiranje transakcij "na ekranu". Podatkovna skladišča spadajo v tako imenovano skupino analitičnih sistemov ali OLAP (On-Line Analytical Processing ali analitično procesiranje "na ekranu") sistemov, čeprav nekateri pod oznako OLAP razumejo le strogo povpraševalni del celotne analitične strukture.

Transakcijski sistemi so optimizirani za obdelavo relativno malega števila računalniških zapisov na enkrat. Recimo dveh, petih ali petstotih. Sistem, ki bi odgovoril na vprašanje, kakšen je bil povprečen obrestni razmik obrestnega dela bilance v zadnjih treh mesecih, pa bi moral zajeti milijone zapisov naenkrat in dati rezultat v sekundah! In ko bi ga dobili, bi takoj želeli zvedeti, zakaj je trend v zadnjem mesecu drugačen, kot v prvih dveh. Zato bi zagnali novo povpraševanje, na primer, kakšne so v tem obdobju povprečne obrestne mere po produktih aktive in pasive in ga tako ponovno poslali žgečkati milijone in milijone zapisov z željo, da nam čimprej poteši informacijski apetit.

Pojem podatkovnega skladiščenja se v slovenskem bančnem prostoru uporablja precej svobodno, kar je razlog za marsikateri nesporazum. Poglejmo si, kako naj izgleda osnovna struktura integralnega podatkovnega skladišča.

2. Integralno podatkovno skladišče banke

Podatkovno skladiščenje razlikuje predvsem med dvema glavnima tipoma podatkov: dejstva (angleško facts) in dimenzije ali atributi. Dejstva so na primer saldo, znesek transakcije in višina obrestne mere, atribut ali dimenzijski podatek pa je komitent, njegov naslov in organizacijska oblika, datum transakcije, poslovna enota, v kateri se je transakcija zgodila, valuta transakcije ali produkt, ki je predmet transakcije.

Dimenzijski podatki ali atributi se tako imenujejo, ker osvetljujejo različne, med seboj neodvisne vidike posameznega dejstva. Atribut danega salda, je komitent, ki mu saldo pripada. Atributi istega zneska so tudi njegova valuta (recimo SIT), produkt (recimo vezana vloga) in poslovna enota. Vsaka dimenzija vsebuje mnogo atributov, ki jo opisujejo. Tako lahko o komitentu nabiramo vse mogoče podatke, od naslova, do zneska prihodkov. Na enak način lahko poslovne enote opisujemo z najrazličnejšimi atributi, kot so na primer lokacija, velikost, odpiralni čas, število zaposlenih... Vsi ti atributi so namenjeni kasnejši analizi dejstev.

V integralnem podatkovnem skladišču banke naj bodo naslednji sklopi dejstev: podatki bilance stanja, podatki bilance uspeha, tržni podatki, bančna tarifa oz. cenik ter zaradi slovenske specifičnosti, ki bo pojasnjena kasneje, podatki glavne knjige.

2.1 *Bilanca stanja*

Bilanca stanja naj bo v podatkovnem skladišču razbita do največje možne mere in opremljena z atributi, ki bodo omogočali odgovarjanje tudi na taka vprašanja, ki v času začetka projekta niso bila predvidena. Predvideno strukturo si oglejmo skozi primer.

Nivo razdrobljenosti saldov, ki je vsebovan v glavni knjigi, je datum, valuta, konto in stroškovno mesto. Prva in zelo očitna stvar, ki manjka, je seveda komitent. Ko ga dodamo, se obseg bilance stanja bistveno poveča. En saldo na kontu "Kreditiranje prebivalstva" na stroškovnem mestu "Maribor" v valuti "SIT" moramo razbiti na toliko manjših zneskov, kolikor komitentov ima svoj saldo v seštevku tega konta na tem stroškovnem mestu in tej valuti. Ko to storimo po vseh kontih, stroškovnih mestih in valutah, pridemo v najboljšem in le teoretično realnem primeru do številke, ki je enaka številu komitentov, v

praksi pa so te številke lahko zelo velike. Srednje velika slovenska banka lahko računa s 100.000 zapisi dnevno, večje tudi z bistveno več. Na ta način je bilanci stanja dodana bistvena izrazna moč, saj je sedaj mogoče odgovoriti recimo na vprašanje, katere posle ima komitent odprte na določen dan in pod kakšnimi pogoji. Ali pa, kako se gibljejo volumni tolarskega kreditiranja srednje velikih gospodarskih družb. Ali kolikšna je povprečna obrestna mera najetih posojil od samostojnih podjetnikov v Mariboru v primerjavi z Ljubljano...

Kako se v izbranem obdobju spreminja struktura indeksiranja kreditov na področju kreditiranja gradbene industrije? Odgovor na to in podobna vprašanja je mogoč, če dodamo dimenzijo "vrsta obrestnega indeksa".

Analizo obrestnih mer na sklenjenih pogodbah lahko začnemo izvajati, če uvedemo nova dejstva: znesek obrestnega indeksa, znesek obrestnega dodatka, iz obeh izvedeno skupno obrestno mero ter izračunano interno stopnjo donosnosti. Slednja omogoča primerjavo rentabilnosti pogodb različnih produktov.

Pogodbena ročnost omogoča analizo obrestne krivulje in njeno spreminjanje skozi čas, hkrati pa predstavlja povezavo na bančno tarifo. Naša radovednost lahko ponovno dobi krila: kakšna je razlika v povprečni obrestni meri 60 dnevni vezanih vlog prebivalstva v EUR danes in pred enim mesecem? Ali, kakšna je naša krivulja obrestnih mer, ki izhaja iz sklenjenih pogodb na področju kreditiranja prebivalstva v primerjavi s krivuljo v istem obdobju lani?

Oznaka pogodbe omogoča vrtanje po podatkih do osnovnega nivoja, hkrati pa je potreben pogoj za izvajanje uspešnih kontrol podatkov.

Z uvedbo produktne dimenzije, dobi banka možnost oblikovanja povpraševanj in poročanja po produktih, ki ustrezajo njeni lastni ideji o tem, kaj produkti so in ni več vezana na zakonsko predpisani kontni plan. Produktni šifrant mora biti podrobno razčlenjen, saj to omogoča uvedbo različnih agregatnih nivojev, ki lahko podpirajo tudi nezdružljive poročevalske zahteve. Primer: transakcijski račun podjetja je moč za potrebe poročanja seštevati v agregat vpoglednih sredstev gospodarstva, za potrebe upravljanja z likvidnostjo pa v agregat visoko volatilnih virov.

Število zapisov v tako razdrobljeni bilanci stanja srednje velike banke lahko sedaj naraste že na 200.000 na dan, bistveno pa se poveča šele z uvedbo denarnih tokov in planskih kategorij.

Uvedba atributa "datum zapadlosti denarnega toka" dimenzije "preostala dospelost" sedaj že dodobra razdrobljene salde še dodatno razbije, saj je zaradi tega potrebno na primer en saldo enega anuitetnega kredita razbiti na vse posamezne razdolžnine mesečnih anuitet. Različni hierarhični nivoji te dimenzije združujejo bodoče datume zapadlosti v časovne košarice (1M, 3M, 6M, 1L ali 30D, 60D, 90D, 1L) in tako omogočajo elegantno analizo vseh bodočih stanj bilance. Glavni vidik je seveda napoved likvidnosti.

Analiza izpostavljenosti neto obrestnega rezultata tveganjem spremembe obrestnih mer ni mogoča brez uvedbe preostale dospelosti obrestnih mer. Ta podatek pove, kako dolgo še velja trenutno veljavna obrestna mera vsake pogodbe in posledično, kako dolgo je iz trenutne bilance možno tržiti neto obrestni rezultat, ki iz nje izhaja.

Sedanja vrednost banke je seštevke vseh diskontiranih denarnih tokov banke, ki jo lahko začnemo izračunavati, če uvedemo dimenziji planskih denarnih tokov in bilančnih objektov. Planski tokovi se nanašajo predvsem na vse bodoče denarne tokove iz obrestnih mer (pri anuitetnem kreditu tako ne prenašamo le razdolžnin, temveč ločeno tudi vsakokratne pripadajoče obresti), ki pa jih moramo v podatkovnem skladišču ločiti od denarnih tokov, ki v svojem seštevku že nastopajo v računovodski bilanci stanja. Temu je namenjena dimenzija bilančnih objektov, ki uvaja kategorizacijo na glavnice, obresti (redne, razmejene, planske), provizije, popravke vrednosti ipd..

Preostala dospelost obrestnih mer in planski denarni tokovi so potreben predpogoj za izvajanje modernih pristopov upravljanja bilance, ki temeljijo na iskanju optimalne uravnoteženosti izpostavljenosti neto obrestnega rezultata in neto sedanje vrednosti banke nihanjem tržnih obrestnih mer. Gledano s stališča informatizacije žel predstavljajo veliko dodatno kompleksnost in občutno povečanje števila zapisov.

2.1.1 Stanja in prometi

Bilanca stanja ima še svoj prometni del, ki ga predstavljajo vse transakcije, ki tvorijo posamezne salde. Saldi bilance stanja se za razliko od večine glavnih knjig v podatkovnem skladišču zaradi performančnih razlogov shranjujejo že izračunani, tako da prometov samih po sebi ne bilo potrebno prenašati. So pa prometi zanimivi iz dveh razlogov: omogočajo samostojno analizo dogajanj na posameznih računih in omogočajo popraviljanje shranjenih stanj z valuto za nazaj. Delta algoritem, ki na osnovi datiranja prometov popravlja vse že shranjene salde, ali pa popravljene salde zapisuje v nova polja, je kar se tiče realizacije kljub svoji navidezni enostavnosti eden najkompleksnejših postopkov podatkovnega skladiščenja.

2.2 Bilanca uspeha in kontroling

Tudi bilanca uspeha oz. njeni prometi naj bodo opremljeni z bistvenimi dimenzijskimi atributi, kot je to opisano v bilanci stanja. Torej, vsak prihodek ali strošek naj bo opremljen z datumom, oznako komitenta, produkta, valute, profitnega centra, kontom, oznako pogodbe, odveč pa ne bi bila niti pogodbeni ročnost in vrsta obrestne mere. Datumi zapadlosti tu niso relevantni.

Problem bilance uspeha se izpostavi v navezi s sistemom kontrolinga in njegovimi transfernimi cenami, ki običajno nista dovolj podrobno in dosledno vpeljana. Določanje in beleženje transfernih cen mora biti vpeljana do nivoja posameznih pogodb, saj bilanca uspeha sama po sebi nima dovolj uporabne izrazne vrednosti za učinkovito upravljanje banke. Le dovolj dodelan sistem transfernih cen generira smiselne interne transakcije, ki šele po prenosu v podatkovno skladišče in v navezi z rednimi transakcijami omogočajo hitro in učinkovito analizo komitentov, produktov in poslovnih enot.

2.3 Glavna knjiga

Glavna knjiga uživa velik ugled, kar se tiče točnosti podatkov, ki jih vsebuje, saj ga glede tega podpira razvejan knjigovodsko/računovodsko/revizorski sistem kontrol. V integralnem podatkovnem skladišču jo želimo zato, da lahko izdelamo avtomatizirana kontrolna poročila, ki z njenimi saldi primerjajo stanja, ki se nahajajo v razdrobljenih bilanci stanja in uspeha in so tja prispela direktno

iz transakcijskih sistemov. Takšna kontrolna poročila identificirajo razlike na nivoju konta in stroškovnega mesta ter jih poročajo agregirane v najvišjih produktnih hierarhijah. OLAP funkcionalnost omogoča vrtanje nazaj v globino in identifikacijo posameznih vzrokov za odstopanja.

2.4 *Tarifa bančnih storitev*

Avtomatiziran sistem poročanja, ki identificira vse pogodbe, katerih sklenjeni pogoji odstopajo od v času sklepanja pogodbe veljavne tarife, mora izdelati poročilo v agregirani obliki po najvišjih produktnih nivojih in/ali segmentacijah komitentov.

Dimenzionalnost tarife je sestavljena iz njenega časovnega obdobja veljave, produkta ali skupine produktov, komitenta ali segmenta komitentov, bonitete pogodbe ali komitenta ter valute, mnoge tarife pa diferencirajo svoje pogoje tudi glede na zneske. Vsi atributi morajo biti poenoteni čez celotno banko, prav tako pa mora vsaka pogodba vsebovati podatek, po kateri tarifni postavki je bila ovrednotena.

Ustrezno tarifiranje in modifikacija izvornih sistemov, da bi lahko ob pogodbah shranjevali tudi tarifne informacije, sta projekta, ki ga bodo morale mnoge banke šele izpeljati. To je vsekakor potreben predpogoj, če bi naj imel prenos tarife v podatkovno skladišče svoj smisel.

2.5 *Tržni podatki*

Nujno je pripeljati vsaj srednje tečaje Banke Slovenije, saj so potrebni za preračunavanje zneskov v skupno valuto SIT in/ali evro. Ti preračunani zneski se morajo že shraniti v podatkovnem skladišču in se ne smejo preračunavati ob vsakem povpraševanju.

Tržnih podatkov, ki so zanimivi za analizo, je mnogo, za tukaj opisano skladišče pa so pomembne predvsem obrestne mere, ki jih potrebujemo pri izračunu neto sedanje vrednosti bilance stanja.

3. *Kje začeti?*

Jasno je, da vsega tukaj opisanega ni mogoče realizirati čez noč. Zato se mora banka odločiti, kaj je pomembno sedaj in kaj lahko ali pa bo moralo počakati.

Do takšne odločitve mora priti management in sicer bistveno globlje, kot le na deklarativnem nivoju. Zgolj ponavljanje, kako je potrebno imeti in kako je nujno narediti, projekta namreč ne pospeši, niti ne približa rešitve, če za temi izjavami ne stojita volja za reševanje problemov, ki jih takšen projekt neizogibno izpostavi in volja ter možnost bodočo bilanco uspeha vsaj začasno obremeniti z dodatnimi stroški.

Predvsem se mora management odločiti, kaj je za banko najpomembnejše in posledično, na katerem področju se bo začelo z izgradnjo skladišča. Ali gre za CRM (Customer Relation Management), ali so to tveganja in če da, katera, ali je to upravljanje z dobičkom banke, ali je ALM (Asset Liability Management)? Želimo ugotoviti neto vrednost banke in preučevati njeno občutljivost na različne vpliva trga ali pa smo zadovoljni le s preprostejšo analizo izpostavljenosti obrestnega rezultata? Ali se želimo iti kontrolinga in do katere mere nas zanimajo podrobnosti iz bilance uspeha? Vsako od teh področij je samo zase ogromno in zahteva zelo dosti naporov, da se v ozadju izdela vsaj približno uporabna analitična struktura. Še pomembnejše pa je, da gre za med seboj skoraj neodvisne dele, ki bi jih načeloma lahko gradili vzporedno, če bi imeli dovolj virov. Ker pa so ti, tako finančni kot človeški, vedno omejeni, se je potrebno odločiti, s čim začeti.

Ključnega pomena je tudi ocena, v koliki meri so ti sistemi že razviti v svojem transakcijskem delu. Ali je kontroling postavljen na nivoju podatkov glavne knjige, ali pa banka že leta dolgo postavlja transferne cene na vsako sklenjeno pogodbo? Ali se za upravljanje z obrestnim tveganjem uporablja letno, mesečno ali dnevno poročanje? Ali se uporabljajo moderni in napredni pristopi, ali pa zgolj ocene čez palec? Stopnja razvitosti določenih segmentov upravljanja je odločilen faktor pri tem, kaj želimo peljati v podatkovno skladišče. Slednje teh sistemov namreč ne bo vzpostavilo, temveč kvečjemu pohitrilo in omogočilo. Povedano drugače, če kontrolinga še ni, ga tudi ne bo zaradi izgradnje podatkovnega skladišča. Lahko se začne z izgradnjo kontrolinga hkrati z izgradnjo podatkovnega skladišča, vendar ta dva projekta ostajata ločena in nepovezana.

4. Kupiti ali izdelati sam?

Opisana struktura integralnega podatkovnega skladišča je zelo zahtevna za realizacijo. Postavlja se vprašanje, ali jo je smiselno graditi v lastni režiji ali pa

naj se poišče ustrezen produkt na trgu oz. v kakšni meri iskati pomoč izven banke.

Podatkovno skladišče je namenjeno analitičnemu povpraševanju po integralnih podatkih, ki omogočajo optimizacijo poročanja ter upravljanja s tveganji, z bilanco in z banko. Žal ne obstaja konsenz, kaj od tega je najpomembnejše in katera izmed metod je za posamezno področje najprimernejša. Iz tega sledi, da je za vsako banko pomembno nekaj drugega, ko se govori o tem, kaj je ključnega pomena in kaj ne.

Velja ocena, da 90% dela ni v izdelavi samega podatkovnega skladišča, čemur bi se z nakupom izognili, temveč na primer v tem, da ga je potrebno vpeti v vsakokratni bančni sistem, da se mora banka ustrezno (re)organizirati, vpeljati ostale ustrezne sisteme, spremeniti mišljenje ljudi in managementa ter sestaviti ustrezno ekipo¹. Morebitni nakup rešuje torej le 10% naloge, velika verjetnost pa je, da se z njim nevede sprejme neka filozofija, ki banki ni lastna, ki ne izhaja iz njenega načina razmišljanja o tem, kako upravljati, kaj je pomembno in čemu dati prioriteto. Pri tem nakup tujega produkta še dodatno oteži situacijo, saj je ob izdelavi že tako zahtevnih vmesnikov, ki pridobivajo podatke, potrebno še prilagajati slovenski bančno-podatkovni model na tujega.

Slovenske banke naj same gradijo svoje podatkovno skladišče. Na trgu naj iščejo pomoč v obliki strokovnjakov, ki so na tem področju že nabrali določene izkušnje in poznajo specifične slovenskih bančnih rešitev. Na trgu naj najemajo programerje, projektne vodje in analitike. Torej, kupujejo naj ljudi, ne produkta. Izdelajo naj rešitev, ki bo v glavnem plod lastnega poslovnega znanja in ki bo za seboj zapustila team visoko usposobljenih strokovnjakov, sposobnih samostojnega nadaljnjega razvoja podatkovnega skladišča.

Viri:

Ralph Kimball, 1996, The Data Warehouse Toolkit

Ralph Kimball, 1998, The Data Warehouse Lifecycle Toolkit

¹ Več o tem bo govora v prispevku v naslednji številki.

Ralph Kimball, 2004, The Data Warehouse ETL Toolkit

Christopher Adamason, 1998, Data Warehouse Design Solutions

Henner Schierenbeck, 1997, Ertragsorientiertes Bankmanagement

Henner Schiernebeck, 1994, Handbuch Bankcontrolling