

Implementacija postupka MDM u financijskoj agenciji

Šimunović, Matea

Master's thesis / Diplomski rad

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:200:401080>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-28**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I
INFORMACIJSKIH TEHNOLOGIJA**

Sveučilišni studij

**IMPLEMENTACIJA POSTUPKA MDM U
FINANCIJSKOJ AGENCIJI**

Diplomski rad

Matea Šimunović

Osijek, 2019.

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK

Obrazac D1: Obrazac za imenovanje Povjerenstva za obranu diplomskog rada

Osijek, 21.09.2019.

Odboru za završne i diplomske ispite

Imenovanje Povjerenstva za obranu diplomskog rada

Ime i prezime studenta:	Matea Šimunović
Studij, smjer:	Diplomski sveučilišni studij Računarstvo
Mat. br. studenta, godina upisa:	D-952R, 26.09.2018.
OIB studenta:	57678488429
Mentor:	Prof.dr.sc. Goran Martinović
Sumentor:	
Sumentor iz tvrtke:	Tomislav Šimat, mag.ing.
Predsjednik Povjerenstva:	Izv. prof. dr. sc. Alfonzo Baumgartner
Član Povjerenstva:	Izv. prof. dr. sc. Krešimir Nenadić
Naslov diplomskog rada:	Implementacija postupka MDM u financijskoj agenciji
Znanstvena grana rada:	Programsko inženjerstvo (zn. polje računarstvo)
Zadatak diplomskog rada:	Svaka velika tvrtka koja se bavi raznim djelatnostima i uslugama, neizbježno nailazi na problem izrade smislenih izvješća; taja iz svih svojih poslovnih centara. Kako bi to bilo moguće napraviti, potrebno je uskladiti podatke iz različitih sustava i definirati jedinstvenu istinu na razini čitave tvrtke kako bi se podaci mogli integrirati jedni s drugima. To se postiže uvođenjem takozvanog Master Data Managementa (MDM). MDM bi trebao poslužiti za definiranje jedinstvenog pogleda na podatke, omogućiti integraciju tih podataka i smanjiti količinu smeća u podacima. Fina, kao velika tvrtka s gotovo 200 usluga, ima potrebu koristiti MDM na tri razine: na razini zajedničkih šiframnika, na razini podataka važnih za vlastito poslovanje i na razini podataka koje koriste pojedine usluge.
Prijedlog ocjene pismenog dijela ispita (diplomskog rada):	Izvrstan (5)
Kratko obrazloženje ocjene prema Kriterijima za ocjenjivanje završnih i diplomskih radova:	Primjena znanja stečenih na fakultetu: 3 bod/boda Postignuti rezultati u odnosu na složenost zadatka: 3 bod/boda Jasnoća pismenog izražavanja: 3 bod/boda Razina samostalnosti: 3 razina
Datum prijedloga ocjene mentora:	21.09.2019.
Potpis mentora za predaju konačne verzije rada u Studentsku službu pri završetku studija:	Potpis:
	Datum:

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK**IZJAVA O ORIGINALNOSTI RADA**

Osijek, 01.10.2019.

Ime i prezime studenta:

Matea Šimunović

Studij:

Diplomski sveučilišni studij Računarstvo

Mat. br. studenta, godina upisa:

D-952R, 26.09.2018.

Ephorus podudaranje [%]:

5

Ovom izjavom izjavljujem da je rad pod nazivom: **Implementacija postupka MDM u financijskoj agenciji**

izrađen pod vodstvom mentora Prof.dr.sc. Goran Martinović

i sumentora

moj vlastiti rad i prema mom najboljem znanju ne sadrži prethodno objavljene ili neobjavljene pisane materijale drugih osoba, osim onih koji su izričito priznati navođenjem literature i drugih izvora informacija. Izjavljujem da je intelektualni sadržaj navedenog rada proizvod mog vlastitog rada, osim u onom dijelu za koji mi je bila potrebna pomoć mentora, sumentora i drugih osoba, a što je izričito navedeno u radu.

Potpis studenta:

SADRŽAJ

1. UVOD	1
1.1. Zadatak diplomskog rada.....	1
2. UPRAVLJANJE MATIČNIM PODACIMA (MDM).....	3
2.1. Matični podaci	3
2.2. Postupak upravljanja matičnim podacima	4
2.2.1. Prednosti upravljanja matičnim podacima	6
2.3. Implementacija upravljanja matičnim podacima.....	8
2.3.1. Kvaliteta podataka.....	9
2.3.2. Integracija podataka	10
2.3.3. Održavanje podataka	12
2.3.4. Stilovi implementacije upravljanja matičnim podacima	13
2.3.4.1. Registarski stil.....	13
2.3.4.2. Konsolidacijski stil.....	14
2.3.4.3. Koegzistencijski stil	15
2.3.4.4. Transakcijski/Centralizirani stil	16
2.3.5. Odnosi između različitih stilova implementacije MDM-a.....	17
2.4. Primjena postupka upravljanja matičnim podacima u Fini	18
3. SKLADIŠTE PODATAKA.....	19
3.1. Definicija skladišta podataka	19
3.2. ETL proces integracije.....	21
3.3. Dimenzijsko oblikovanje skladišta podataka	22
3.4. Primjer oblikovanja skladišta podataka u Fini	25
3.4.1. Dimenzija “Klijent”	25
3.4.1.1. Punjenje dimenzije „Klijent“	29
3.4.2. Dimenzija “Organizacijski dio”	31
3.4.3. Dimenzija “Djelatnik”	32

3.4.4.	Dimenzija “Likvidator”	34
3.5.	Interakcija između skladišta podataka i MDM-a.....	34
4.	DIZAJN SREDIŠNJEG SUSTAVA ŠIFRARNIKA.....	36
4.1.	Osnovna načela MDM-a na razini zajedničkih šifrnika	37
4.2.	Specifikacija Središnjeg sustava šifrnika	39
4.3.	Komponente Središnjeg sustava šifrnika.....	40
4.3.1.	Unutarnje komponente	40
4.3.1.1.	Spremnik Središnjeg sustava šifrnika	40
4.3.1.2.	Modul za evidenciju.....	41
4.3.1.3.	Modul za obavijesti.....	41
4.3.1.4.	Sloj za podatke o korisnicima i prava pristupa	41
4.3.2.	Vanjske komponente	42
4.4.	Osnovne funkcionalnosti Središnjeg sustava šifrnika	43
4.4.1.	Održavanje Središnjeg sustava šifrnika	44
4.4.2.	Korištenje podataka iz Središnjeg sustava šifrnika	46
4.5.	Odgovornosti sudionika unutar središnjeg sustava šifrnika	47
4.5.1.	Administrator sustava.....	47
4.5.2.	Fizička osoba.....	47
4.5.3.	Mrežna usluga	47
4.6.	Analiza dizajna Središnjeg sustava šifrnika	48
4.7.	Vrednovanje dizajna Središnjeg sustava šifrnika	49
5.	ZAKLJUČAK	53
	LITERATURA.....	54
	SAŽETAK.....	56
	ABSTRACT	57
	ŽIVOTOPIS	58
	PRILOZI.....	59

1. UVOD

U prošlosti su se podaci bilježili na razne načine, no pojava digitalizacije mijenja način rada s njima i ubrzava njihovo prikupljanje, čime ujedno podaci dobivaju na važnosti. Danas se podaci smatraju vrijednom informacijskom imovinom u svakom poduzeću. Gotovo ne postoji tvrtka koja iz različitih izvora ne prikuplja velik broj podataka, pretvara ih u informacije te koristi kroz gotovo sve aspekte svoga poslovanja. Kako podaci pristižu iz različitih izvora, potrebno je uskladiti podatke i definirati jedinstvenu istinu na razini čitave tvrtke. A kako bi to bilo moguće neophodno je uvođenje postupka upravljanja matičnim podacima – MDM.

MDM (engl. *Master Data Management*) se temelji na postizanju visoko-kvalitetnih objedinjenih podataka, koji kao takvi moraju biti dostupni ostalim sustavima unutar poduzeća. Iz tog razloga, bitan aspekt MDM postupka je proces integracije, kojim se podaci iz različitih sustava konsolidiraju unutar jednog sustava. ETL (engl. *Extract, Transform, Load*) je poznati alat koji se primarno koristi za integraciju podataka prilikom skladištenja podataka (engl. *Data Warehousing*), no također može biti vrlo koristan i primjenjiv unutar MDM postupka.

U drugom poglavlju, kroz definiciju matičnih podataka te zašto je velikim organizacijama toliko bitno upravljati njima, opisuje se koncept i stilovi implementacije upravljanja matičnim podacima. Trećim poglavljem ističe se važnost skladišta podataka i ETL procesa integracije u izgradnji MDM sustava. Kako je ovaj rad nastao u suradnji sa Finom, kao primjer su navedene osnovne dimenzije Fininog skladišta podataka, koje su, zahvaljujući integraciji podataka putem ETL-a, bile ključne prilikom implementacije određenih MDM rješenja unutar Fine. U četvrtom poglavlju su opisana osnovna načela, specifikacija, komponente, osnovne funkcionalnosti te odgovornosti sudionika Središnjeg sustava šifrnika, čiji model nastaje kao produkt praktičnog dijela ovog rada, a ima ulogu specifikacije dizajna MDM sustava na razini Fininih zajedničkih šifrnika. Model Središnjeg sustava šifrnika dodatno je analiziran te na osnovu toga i vrednovan.

1.1. Zadatak diplomskog rada

Svaka velika tvrtka koja se bavi raznim djelatnostima i uslugama, neizbježno nailazi na problem izrade smislenih izvještaja iz svih svojih poslovnih centara. Kako bi to bilo moguće napraviti, potrebno je uskladiti podatke iz različitih sustava i definirati jedinstvenu istinu na razini čitave tvrtke kako bi se podaci mogli integrirati jedni s drugima. To se postiže uvođenjem takozvanog *Master Data Managementa* (MDM). MDM bi trebao poslužiti za definiranje jedinstvenog pogleda na podatke, omogućiti integraciju tih podataka i smanjiti količinu smeća u podacima. Fina, kao velika tvrtka s gotovo 200 usluga, ima potrebu koristiti MDM na tri razine: na razini zajedničkih

šifrnika, na razini podataka važnih za vlastito poslovanje i na razini podataka koje koriste pojedine usluge koje je potrebno povezati. U ovom diplomskom radu treba objasniti koncept MDM-a, prikazati metode integracije podataka na primjeru formiranja dimenzija u Fininom skladištu podataka, te dizajnirati Središnji sustav šifrnika koji bi bio jedinstveno mjesto za održavanje zajedničkih šifrnika u Fini. Dizajn središnjeg sustava šifrnika sadržavat će opis sustava i njegovih dijelova, te opis funkcionalnosti i sudionika sustava. Ostale specifikacije dizajna sustava, te scenarije koje sustav mora zadovoljiti, definirat će Fina čiji tim će ispitati i analizirati sustav.

2. UPRAVLJANJE MATIČNIM PODACIMA (MDM)

Problem vjerodostojnosti podataka dodatno se komplicira u velikim organizacijama koje se unutar svojih sustava i procesa bave raznim djelatnostima i uslugama te iz istih primaju ogromne količine podataka. Kako podaci dolaze iz različitih izvora, tako su prikazani u raznim oblicima. Kao posljedica toga javlja se veliki izazov povezivanja podataka u smislenu cjelinu, kojom će se svaki sustav poduzeća služiti u skladu sa svojim potrebama i pomoću koje će podaci stvoreni u ranijim procesima pružati kontekst podacima koji će se stvarati u kasnijim procesima. Kako bi se podaci mogli objediniti jedni s drugima, potrebno ih je iz različitih izvora uskladiti te definirati jedinstvenu verziju istine koja će vrijediti na razini čitavog poduzeća [1].

Posljedice se također pojavljuju između različitih odjela koji koriste iste podatke za različite potrebe. Primjerice, odjelima prodaje, financija i proizvodnje zajedničko je to da sva tri odjela vode računa o prodaji proizvoda, no svaki odjel ima različite potrebe i očekivanja u smislu kvalitete podatka, što rezultira nedosljednim vrijednostima podataka prikazanim u različitim formatima te u konačnici, iste te nedosljednosti imaju dramatično negativan utjecaj na cjelokupnu kvalitetu podataka.

Kao rješenje prethodno navedenih nedostataka pojavljuje se postupak poznat kao Upravljanje matičnim podacima - MDM (engl. *Master Data Management*). Glavni cilj MDM-a je, smanjivanjem količine smeća i definiranjem jedinstvenog pogleda nad podacima, postići konzistentnu i točnu razmjenu informacija između različitih sustava, procesa, odjela i ostalih jedinica unutar poduzeća.

2.1. Matični podaci

Matični podaci, kao temelj MDM-a, predstavljaju vrijednosti podataka o ključnim poslovnim subjektima koji se koriste za uspostavu konteksta podataka iz izvornih sustava. Uobičajeni matični podaci koji se pojavljuju unutar organizacije uključuju podatke o raznim ulogama pojedinaca i organizacije, unutarnjih i vanjskih čimbenika, financijskih struktura ili raznih lokacija.

Najčešći primjeri matičnih podataka su o kupcima, zaposlenicima, dobavljačima, odjelima, financijama, proizvodima, lokacijama, ugovorima, itd. Primjerice, [1] spominje sustav za upravljanje odnosima s kupcima – CRM (engl. *Customer Relationship Management*) koji predstavlja zajednički podatkovni sustav u kojem se nalaze matični podaci o korisnicima, matični podaci o zaposlenicima nalaze se u sustavima za upravljanje ljudskim resursima – HRM (engl. *Human Resource Management*), upravljanje resursima poduzeća – ERP (engl. *Enterprise*

Resource Planning) predstavlja sustav koji služi kao središte za financijske podatke, upravljanje informacijama o proizvodima – PIM (engl. *Product Information Management*) sustav je za podatke o proizvodima. Postoji još mnoštvo primjera takvih sustava koji, ovisno o konkretno definiranom subjektu, okupljaju matične podatke.

Komplikacije se događaju kada se podaci o konkretno definiranom subjektu, kao što je kupac ili proizvod, prošire na cijelo poduzeće. U tom slučaju, matični podaci se pojavljuju u više poslovnih područja unutar organizacije, pa se tako npr. isti klijent može pojaviti u prodajnom sustavu, kao i u sustavu naplate. Tada se javlja potreba za okupljanjem matičnih podataka iz više različitih sustava u jedan zajednički. U tom slučaju sustavi kao što su CRM, ERP, PIM, itd. postaju izvorni sustavi jednom zajedničkom sustavu.

Nije potrebno nužno upravljati i pohranjivati u zajednički sustav sve podatke koji odgovaraju definiciji matičnih podataka. Kako bi se izvršila selekcija podataka te kako bi se potom podaci mogli jedinstveno prikazati unutar cijelog poduzeća, moraju se definirati poslovna pravila koja pomoću meta-podataka određuju format i dopuštene raspone vrijednosti matičnih podataka. Radi lakšeg organiziranja, matični podaci se mogu hijerarhijski klasificirati [2]. Primjerice, za određeni kontekst matičnih podataka, "klijent" je glavna kategorija, koja se razlikuje u dvije potkategorije: "fizička osoba" i "pravna osoba". Potkategorije se dalje mogu klasificirati na temelju njihovih uloga, kao što su "kupac", "dobavljač", "korisnik", itd. Sve to izazovi s kojima se suočava postupak upravljanja matičnim podacima, detaljnije opisan u odlomcima 2.2. i 2.3.

Identificirane vrijednosti matičnih podataka iz različitih izvora smatraju se kao ključne vrijednosti, koje se pohranjuju u zajednički podatkovni sustav. Uz pravilno održavanje, matični podaci okupljeni u zajedničkom sustavu, mogu se za različite primjene poduzeća koristiti kao objedinjen i koherentan skup podataka, koji pruža konzistentne i pravovremene informacije [2].

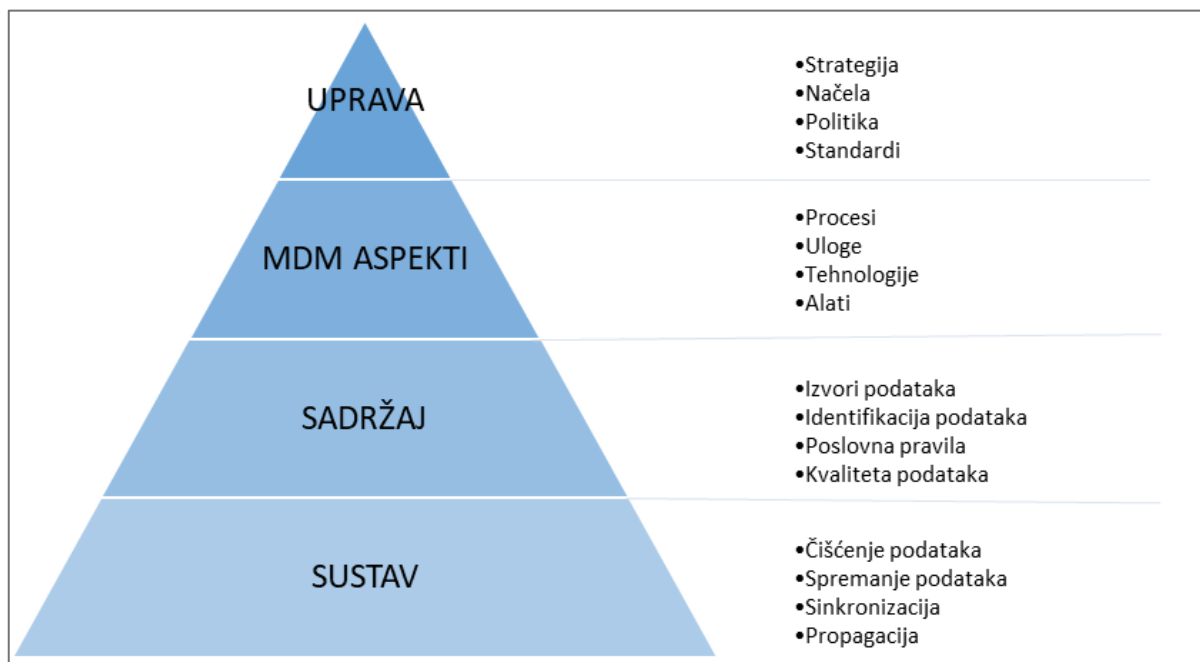
2.2. Postupak upravljanja matičnim podacima

Upravljanje matičnim podacima, ili skraćeno MDM, je skup tehnologija, alata i poslovnih procesa pomoću kojih se definira način kako će se matični podaci kreirati, integrirati, održavati i koristiti na razini cijelog poduzeća [1], [2].

MDM je vrlo zahtjevan pothvat, koji se ne smatra običnim projektom i ne može se riješiti kao dio instaliranja određene tehnologije, već je stalna inicijativa koja se dotiče svakog aspekta organizacije. Upravljanje matičnim podacima obuhvaća od vrhovne uprave (engl. *Top Management*), upravitelja podataka (engl. *Data Stewardship*), raznih procesa, pa sve do običnih

informatičkih sustava. U mnogim slučajevima biti će potrebne temeljne promjene poslovnog procesa za održavanje čistih matičnih podataka, a neki od najtežih problema MDM-a više su politički, nego tehnološki [3].

Svaka inicijativa upravljanja matičnim podacima kreće od vrhovne uprave organizacije, koja prvenstveno treba procijeniti stupanj zrelosti kojim određuje je li organizacija uopće spremna na pothvate koje upravljanje matičnim podacima zahtjeva. Nakon čega, kao što je prikazano slikom 2.1, vrhovna uprava "zrele" organizacije izgrađuje MDM strategiju u skladu sa politikom organizacije, načelima i standardima kojih bi se MDM program trebao pridržavati. Potom se definiraju potrebni procesi, tehnologije i alati te organizira potrebno osoblje koje sudjeluje u procesu MDM-a, uključujući vlasnike matičnih podataka, upravitelje podataka i vrhovnu upravu. Nakon što se definiraju izvori podataka, identificiraju se matični podaci te se definiraju poslovna pravila pomoću kojih se upravlja kvalitetom odabranih podataka. Očišćeni podaci spremaju se u sredinu koje okuplja i održava jedinstvene zapise matičnih podataka [4].



Slika 2.1. *Procedura uvođenja MDM-a*

Na plan postupka upravljanja matičnim podacima uvijek će dodatno utjecati različiti zahtjevi, prioriteti, dostupnost resursa, vremenski okvir i veličina problema. MDM je iznimno veliko područje koje obuhvaća i tehnološke i poslovno-političke discipline. Ovaj rad se fokusira na tehnološkim disciplinama koje će iz tog razloga biti detaljnije objašnjene.

Kroz tehnološke discipline, izazovi s kojima se MDM neizbježno suočava je identificirati najtočnije i najprikladnije vrijednosti podataka koje dolaze iz različitih izvora te se potom

fokusirati na te, umjesto na druge, manje točne, podatke [1]. Potom je identificirane vrijednosti podataka potrebno integrirati kao konzistentnu cjelinu koja će zadovoljavati poslovne potrebe poduzeća te u konačnici nastaviti kontinuirano upravljati kvalitetom i osiguravati integritet podataka.

Tradicionalni procesi namijenjeni očuvanju kvalitete i integriteta među podacima, kao što je npr. skladište podataka koje se detaljnije opisuje u trećem poglavlju, usredotočeni su na statički prikaz analitičkih skupova podataka u zasebnom spremniku. Ono po čemu se MDM razlikuje je to što se njegov skup podataka ne ponaša kao zasebni podatkovni spremnik, već kao unutrašnja komponenta dijeljena za operativne, analitičke ili kombinirane primjene unutar poduzeća [2]. Upravo ta dinamička dostupnost reprezentativnog skupa podataka različitim sustavima unutar poduzeća, jedna je od osnovnih koncepata na kojima se temelji MDM.

Kako bi vrijednosti reprezentativnog skupa podataka bile konzistentne te se učinkovito dijelile kroz različite primjene u organizaciji, treba biti jasno definirano tko zahtjeva kakve informacije, koji podaci su dostupni iz različitih izvora i na koji način se isti podaci razlikuju kroz različite izvore. Poželjno je definirati izvor koji pruža najisplativije vrijednosti podataka, nakon čega se jednostavnije mogu riješiti nedosljednosti među takvim podacima.

Uspješna implementacija MDM-a rezultirat će učinkovitijom suradnjom između tehnologije i poslovanja, što će organizaciju pretočiti u snažan sustav upravljanja podacima koji će bitno utjecati na način poslovanja i pripremati organizaciju na brze promjene.

2.2.1. Prednosti upravljanja matičnim podacima

Krajnji cilj MDM-a nije samo objedinjavanje matičnih podataka, već njihovim dijeljenjem unutar organizacije, poslužiti kao sredstvo dobivanja relevantnih informacija koje pružaju niz prednosti. Kako se u poduzeću povećava broj i raznolikost odjela, informatičkih sustava, procesa, zaposlenika i njihovih uloga, tako se automatski povećava i niz prednosti upravljanja matičnim podacima [5]. Prema [2], neke od prednosti koje pruža uvođenje postupka upravljanja matičnim podacima su:

- Sveobuhvatno znanje o klijentima – zajednički podatkovni sustav, koji iz različitih izvornih sustava okuplja matične podatke o klijentima služi kao jedinstveni izvor konsolidiranih podataka o svim klijentskim aktivnostima
- Poboljšanje klijentske podrške – objedinjen prikaz matičnih podataka o klijentima također omogućuje poboljšanja prilikom ispunjavanja klijentskih očekivanja, transparentnost, dostupnost i točnost te zaštitu osobnih podataka

- Konzistentnost prilikom izvještavanja – izvještaji generirani pomoću sustava koji sadrži konsolidirane matične podatke vjerodostojniji su nego izvještaji generirani putem različitih izvornih sustava
- Poboljšanje konkurentnosti – MDM smanjuje složenost integriranja novih podataka i sustava u organizaciji, povećavajući tako agilnost i konkurentnost
- Poboljšanje upravljanja rizikom – jedinstven prikaz matičnih podataka o financijama pruža točnije i relevantnije informacije koje poboljšavaju sposobnost organizacije za upravljanje rizikom
- Poboljšanje operativne učinkovitosti i smanjenje troškova – formiranje jedinstvenog prikaza podataka smanjuje operativne troškove i dodatne napore organizacije
- Poboljšanje donošenja odluka – konzistentnost podataka koju MDM pruža smanjuje varijabilnost podataka, što povećava povjerenje u generirane izvještaje te omogućuje jasnije i brže donošenje bitnih poslovnih odluka
- Poboljšanje analize i planiranja – jedinstveni prikaz konzistentnih matičnih podataka daje jasniju sliku koja omogućuje bolje provedene analize i planiranja
- Pridržavanje propisa – konzistentnim prikazom matičnih podataka o poduzeću pojednostavljuje se revizija podataka, omogućavajući tako učinkovitu informacijsku kontrolu koja olakšava usklađenost s propisima i ostalim pravilima vezanim za industriju
- Povećanje kvalitete informacija – prikupljanje podataka sastavljenih od standardiziranih modela, vrijednosnih domena i poslovnih pravila omogućuje organizaciji učinkovitiji nadzor usklađenosti s očekivanjima kvalitete informacija
- Brži rezultati – standardizirani prikaz informacijske imovine smanjuje kašnjenje prilikom izdvajanja i transformacije podataka, što ubrzava implementaciju sustava te izradu skladišta i tržišta podataka
- Poboljšanje produktivnosti poslovanja – matični podaci pomažu organizaciji za bolje razumijevanje kako se isti objekti podataka predstavljaju, manipuliraju ili razmjenjuju kroz različite primjene unutar poduzeća te kako se odnose na tokove poslovnih procesa, što ujedno pruža priliku za istraživanjem mogućnosti poduzeća u automatizaciji vlastitih poslovnih procesa
- Pojednostavljeni razvoj aplikacija i implementacija sustava – objedinjavanjem matičnih podataka otvara se prilika i za implementacijom funkcionalnosti sustava povezanih sa životnim ciklusom podataka. Drugim riječima, ukoliko postoje različiti sustavi koji unose promjene vezane uz životni ciklus podataka, njihovim objedinjavanjem u jedan resurs

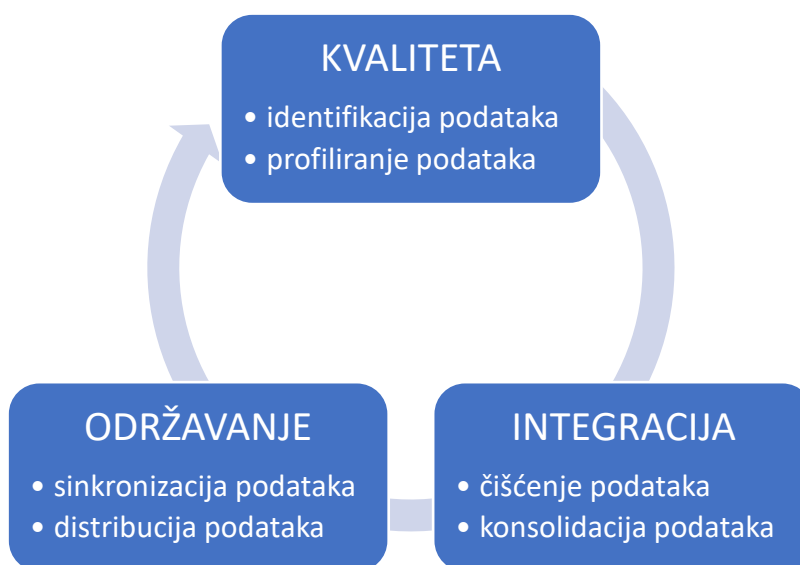
omogućuje se jedinstvena funkcionalna usluga za unos promjena, na koju se različiti sustavi unutar poduzeća mogu pretplatiti.

2.3. Implementacija upravljanja matičnim podacima

Iako ga nije jednostavno implementirati, MDM je s vremenom postao nužna potreba kako bi svaka velika organizacija u potpunosti ostvarila potencijal prihoda i profita vlastitih skupova podataka, a također može biti presudan za moderne inicijative kao što su veliki podaci (engl. *Big Data*), umjetna inteligencija (engl. *Artificial Intelligence*), digitalna transformacija (engl. *Digital Transformation*), itd.

Slikom 2.2 prikazana su tri su primarna tehnološka područja koja moraju biti ispunjena kako bi implementacija bila uspješna te kako bi se zadovoljili kriteriji MDM-a:

1. Kvaliteta podataka, koja obuhvaća proces identificiranja i prikupljanja "zanimljivih" podataka iz različitih izvornih sustava. Nakon čega se podaci profiliraju i postaju kandidati za proces integracije podataka.
2. Integracija podataka, kao proces čišćenja i konsolidacije profiliranih zapisa podataka. S čime se postiže objedinjen pogled na podatke pristigle iz različitih izvornih sustava te se nastoji riješiti problem nedosljednosti podataka.
3. Održavanje podataka, odnosi se na način kojim se podaci ažuriraju iz izvornih sustava u matično okruženje te dalje iz matičnog okruženja dijele ostalim sustavima koji ih koriste unutar poduzeća.



Slika 2.2. Primarna tehnološka područja implementacije MDM-a

Već postojeća rješenja, koja prethode MDM-u, a fokusiraju se na kvalitetu i integraciju podataka, mogu poslužiti prilikom implementacije:

- alati za čišćenje podataka (engl. *Cleaning Tools*)
- alat za integraciju skladišta podataka kao što je ETL (engl. *Extract, Transform, Load*)
- skladišta operativnih podataka ODS (engl. *Operational Data Stores*) koja služe kao manje središte matičnih podataka u svrhu operativnog izvještavanja

Također postoje i specijalizirane platforme koje nude gotova rješenja MDM-a. Neke od poznatijih platformi su: IBM InfoSphere Master Data Management [6], Oracle Product Hub [7], Talend Master Data Management [8], Orchestra Networks [9], Profisee [10], Informatica MDM [11], itd.

2.3.1. Kvaliteta podataka

Kvaliteta podataka predstavlja postupak i metodologiju prepoznavanja izvora određenih nepravilnosti u svrhu njihovog odstranjivanja, odnosno kako bi se onemogućilo daljnje pojavljivanje tih nepravilnosti [2]. Bez dovoljno fokusiranja na upravljanje kvalitetom matičnih podataka, postupak za upravljanje matičnim podacima poduzeća može se pretvoriti u samo još jedan beskorisni spremnik s hrpom nesinkroniziranih podataka [12]. Kvaliteta podataka presudna je za uspjeh i prodire u gotovo sve aspekte MDM-a. Kako bi se ciljevi kvalitete podataka MDM programa podržali nužno je prikupljanje podataka kojima će se popunjavati lista meta-podataka, odnosno podataka koji opisivanjem pružaju znanje o ostalim podacima [2].

Očekivanja kvalitete matičnih podataka procjenjuju se i mjere prema dimenzijama kvalitete podataka, koje su definirane u skladu s poslovnim procesima koje trebaju mjeriti, a prema [1] i [2] mogu se odnositi na jedinstvenost, točnost, konzistentnost, potpunost, pravovremenost, ažurnost, itd. Pomoću dimenzija kvalitete podataka vrednovan je i model Središnjeg sustava šifrnika koji se opisuje u četvrtom poglavlju.

Kako razina kvalitete matičnih podataka diktira uspješnost MDM-a, tako se i ona može poboljšati putem MDM procesa. Ovakva vrsta međuovisnosti zahtjeva određene alate i metode za procjenu, integraciju i trajno jamstvo kvalitete podataka, koji se primjenjuju prema sljedećem principu:

1. Prepoznavanje ključnih elemenata podataka pomoću liste meta-podataka
2. Određivanje elemenata koji čine matične podatke pomoću liste meta-podataka
3. Lociranje i izolacija matičnih podataka
4. Dodatni pregled i rješavanje varijacija među različito prikazanim podacima
5. Integracija jedinstvenog prikaza podataka

Prema [2], dobra procjena je mogućnost prepoznavanja elemenata podataka koji su potencijalni kandidati za integrirano matično okruženje, a ovisi o strukturiranom procesu procjenjivanja koji se oslanja na automatizirane alate za analizu kao što je profiliranje podataka (engl. *data profiling*). Prije nego što započne ikakva integracija podataka, mora se primijeniti proces profiliranja podataka kako bi se izvršilo empirijsko identificiranje potencijalnih kandidata za skup matičnih podataka te definirali primarni ključevi, strani ključevi, unutarnje relacijske strukture i ugrađena poslovna pravila.

Kako podaci pristižu iz različitih izvornih sustava, neizbježne su varijacije među matičnim podacima. Varijacije su najčešće povezane uz osobna imena, adrese, telefonske brojeve, opise proizvoda, itd. Iz tog razloga, pri integraciji je nužno korištenje alata za otklanjanje varijacija među različito predstavljenim podacima. Na temelju toga primjenjuje se statička operacija čišćenja podataka (engl. *data cleansing*), koja podržava metode za standardizaciju (engl. *standardization*), raščlanjivanje (engl. *parsing*) i podudaranje podataka (engl. *matching data*). Metode čišćenja podataka detaljnije su opisane u odlomku 2.3.2.

U svrhu što boljeg prepoznavanja potencijalnih problema kvalitete podataka te sprečavanja nastanka bilo kakvih značajnih poslovnih udaraca, čak i nakon početne integracije matičnih podataka u zajedničko okruženje postoji potreba za povećanim pregledom i kontrolom podataka.

Revizija i nadzor usklađenosti podataka prema definiranim očekivanjima kvalitete podataka, zajedno s učinkovitim praćenjem i otklanjanjem novonastalih problema, te uz čvrsto upravljanje podacima unutar modela upravljanja temeljenog na konsenzusu, osigurat će se trajno jamstvo kvalitete podataka.

2.3.2. Integracija podataka

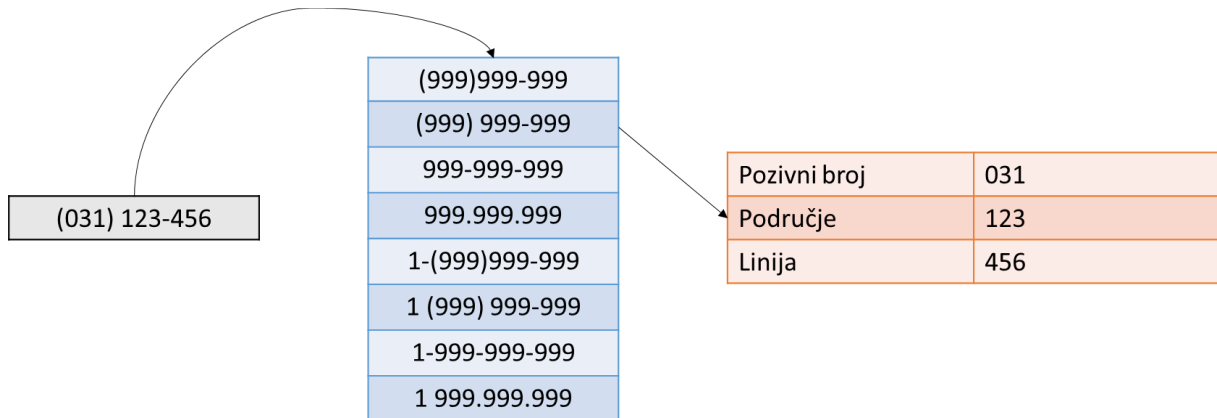
Integracija podataka je proces prikupljanja i ujedinjavanja podataka iz različitih izvora. Kroz proces integracije, identificirane vrijednosti podataka se raščlanjuju, standardiziraju i normaliziraju kroz iste domene podataka, nakon čega se klasificiraju u logičke hijerarhije [2]. Potom se analiziraju pronađeni duplikati, pretražuju ostale hijerarhijske grupe, pronalaze postojeći zapisi, ili se utvrđuje da je određeni zapis jedinstven. Kao ishod procesa integracije dobije se konsolidirani skup matičnih podataka koji treba biti dostupan na razini cijelog poduzeća.

Svaki novi podatak prilikom integracije u matično okruženje mora biti jedinstveno identificiran. U tu svrhu koriste se određene usluge identifikacije (engl. *identification services*), koje pomoću tehnike za rješavanje identiteta (engl. *identity resolution*) pronalaze odgovarajuća podudaranja među zapisima podataka. Rješavanje identiteta je tehnika koja se odnosi na sposobnost utvrđivanja

moгу li se dva ili više zapisa podataka prikazati kao jedinstveni objekt podatka. Ukoliko se određenom zapisu podatka ne može naći identično podudaranje, tada se pronalazi ono koje mu je najbližije, a uspješnost rješavanja identiteta ovisi o, već spomenutim, metodama čišćenja podataka koje su ugrađene u alate za kvalitetu podataka [13]:

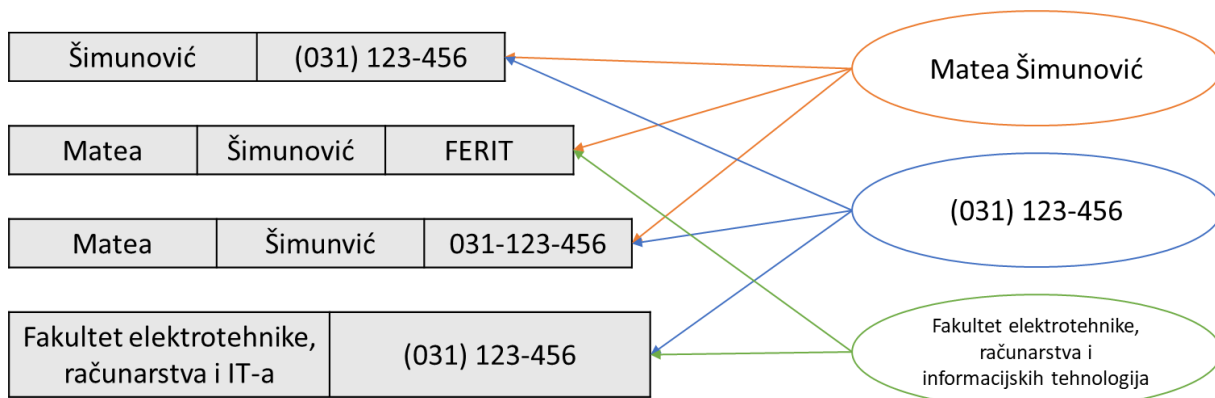
- Raščlanjivanje i standardizacija – zapisi podataka podliježu analizi uzoraka, gdje ukoliko je prema listi meta-podataka uzorak prepoznat kao relevantan, odvajanjem se raščlanjuje i stavlja u standardizirani prikaz
- Transformacija zapisa – ukoliko se prepozna uzorak koji podržava grešku, primjenjuju se pravila normalizacije podataka za izmjenu pogrešaka u prihvatljive formate
- Podudaranje zapisa – procjenjivanje sličnosti nad grupom podataka definiranih unutar liste meta-podataka kako bi se utvrdilo odnose li se na isti objekt matičnih podataka ili ne

Slikama 2.3 i 2.4, koje su nastale po uzoru na [2], prikazani su primjeri rješavanja identiteta pomoću metoda čišćenja podataka. Slikom 2.3 prikazana je metoda raščlanjivanja i standardizacije kroz primjer zapisa telefonskog broja. Ulazna vrijednost telefonskog broja je pomoću liste meta-podataka prepoznata kao relevantna, na temelju čega se raščlanjuje i stavlja u standardizirani prikaz.



Slika 2.3. Raščlanjivanje i standardizacija zapisa podataka [2]

Slikom 2.4 prikazana je metoda podudaranja zapisa, gdje se pomoću grupe zapisa definiranih u listi meta-podataka ustanovilo o kojem objektu matičnog podatka je riječ. Metodom transformacije prepoznat je pogrešno unesen zapis (Šimunvić). Također je vidljivo i kako su zapisi podataka standardizirani.



Slika 2.4. Transformacija i podudaranje zapisa podataka [2]

Integracija se odvija u dva aspekta. Prvi aspekt se odnosi na početnu migraciju podataka iz izvornih sustava u matično okruženje. Dok se drugi aspekt odnosi na ponavljajuću integraciju podataka stvorenih, modificiranih ili uklonjenih iz matičnog okruženja u svrhu očuvanja konzistentnosti matičnih podataka.

2.3.3. Održavanje podataka

MDM proces koji samo prikuplja i prikazuje matične podatke u konsolidiranom obliku, a ne podržava pristup za upotrebu i dijeljenje tih podataka, u osnovi je bezvrijedan. Jer prava vrijednost MDM-a se postiže kada objedinjeni matični podaci, kao visokokvalitetna imovina, budu podijeljeni za operativnu i analitičku upotrebu sustavima unutar poduzeća, s čime se pruža jedinstveni i sinkronizirani prikaz klijenta, proizvoda ili druge jedinice matičnog podatka. Što znači da se matični podaci moraju lako unositi u matično okruženje, a zauzvrat moraju biti lako dostupni poslovnim sustavima s kojima komuniciraju.

Održavanjem podataka definira se način kojim se matični podaci ažuriraju unutar matičnog okruženja, kao i način kojim se matični podaci distribuiraju pretplaćenih sustavima, odnosno sustavima poduzeća koji za potrebe svojih poslovnih procesa koriste matične podatke [2].

Kada se izvorni podaci prvi put apsorbiraju u matično okruženje, tada nema nikakvog postojećeg skupa podataka te takva inicijalna integracija ne zahtijeva nikakve dodatne napore za sinkronizacijom podataka. Međutim, kada matično okruženje više nije prazan skup podatka, kako bi se očuvao životni ciklus podataka, nastavlja se provoditi kontinuirana integracija i konsolidacija, gdje svaki naredni izvor podatka vjerojatno neće biti usklađen s matičnim okruženjem i zahtijevat će dodatne napore prilikom sinkronizacije podataka [2]. Količina napora sinkronizacije ovisi o samoj arhitekturi MDM-a te o dimenzijama kvalitete podataka pomoću kojih se može istražiti stupanj sinkronosti.

Distribucija se može vršiti putem repliciranja matičnih podataka u kopije dostupne pretplaćenim sustavima s pristupom samo za čitanje (engl. *read-only*) ili s potpunim pristupom čitanja i pisanja (engl. *read-write*). Ukoliko je riječ o potpunom pristupu čitanja i pisanja, tada se matični podaci unutar matičnog okruženja moraju naknadno konsolidirati ukoliko neki od pretplaćenih sustava koristi pisanje.

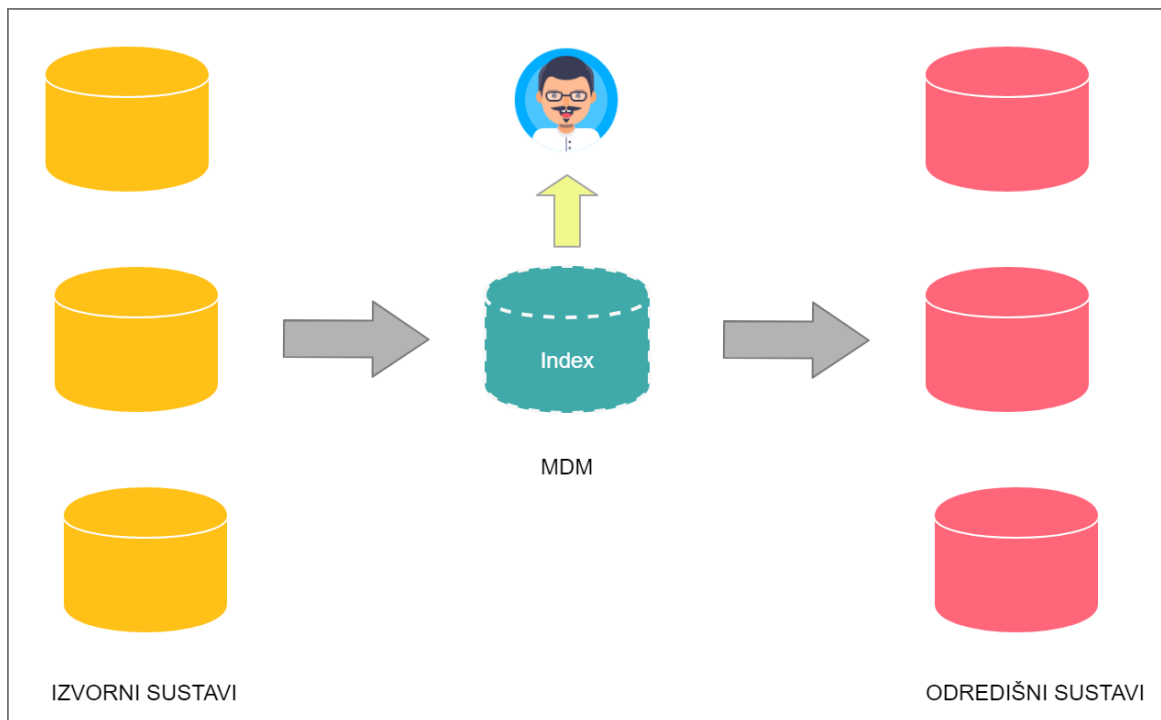
Sinkronizacija i distribucija podataka također se mogu odvijati korištenjem uslužno-orijentirane arhitekture – SOA (engl. *Service-Oriented Architecture*), čiji pristup zahtjeva implementaciju dodatnog uslužnog sloja kako bi se podržale mrežne usluge sustava s kojima matično okruženje komunicira.

2.3.4. Stilovi implementacije upravljanja matičnim podacima

Četiri su poznata stila implementacije MDM-a koje opisuju [14] i [15]: registarski, transakcijski, konsolidacijski i koegzistencijski stil.

2.3.4.1. Registarski stil

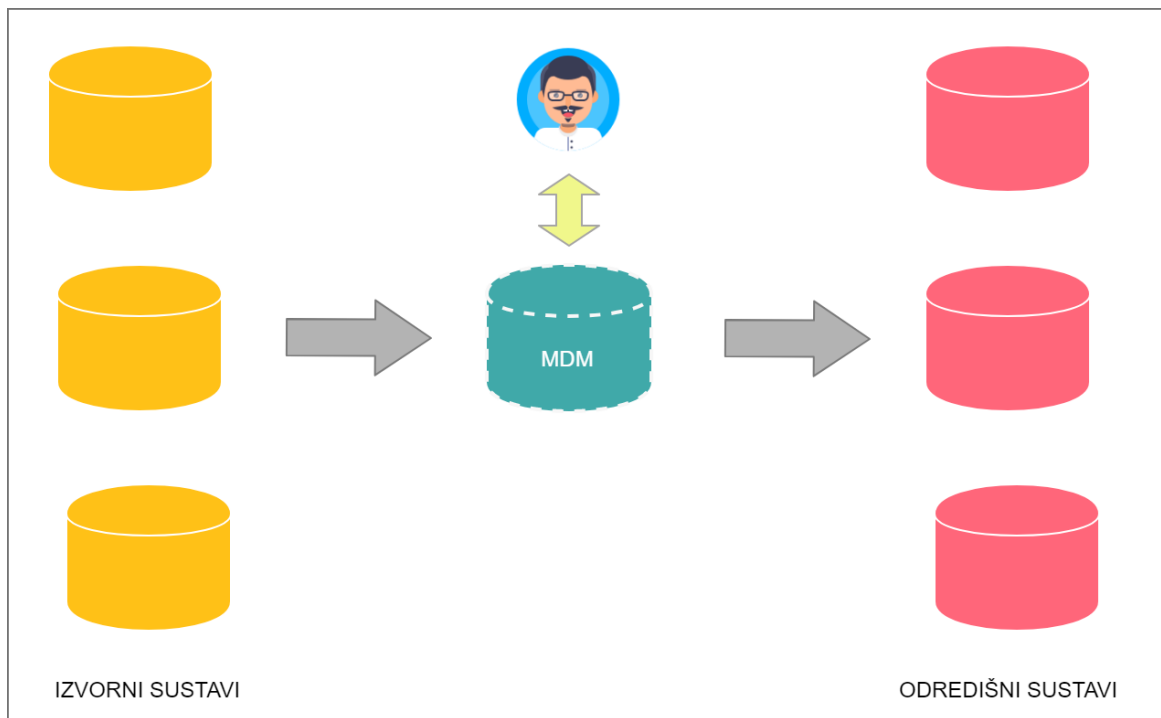
Slikom 2.5 prikazan je registarski stil (engl. *registry style*) implementacije MDM sustava. Registarski stil specifičan je po tome što podatke ne pohranjuje u fizičkom, već u virtualnom obliku. Ovaj stil je koristan kada podaci pristižu iz velikog broja izvornih sustava, od kojih je svaki izvorni sustav složen na svoj način i ima vlastita pravila upravljanja kvalitetom podataka koja je teško izmijeniti [14]. Podaci iz različitih izvora okupljaju se u MDM sustav koji je predstavljen kao registar. Registar procesom integracije ispituje izvorne sustave i pokreće algoritme za čišćenje i podudaranje. Nakon čega, u svrhu definiranja jedinstvene verzije istine, registar dupliciranim zapisima dodjeljuje jedinstvene globalne oznake (engl. *index*) koje služe kao referenca na izvorni sustav. U ovom stilu, veza između izvornih sustava i MDM sustava je jednosmjerna. Drugim riječima, registar ne šalje nikakve podatke natrag u izvorne sustave. Iz tog razloga, atributi matičnih podataka unutar izvornih sustava ostaju niske kvalitete, jer nisu konzistentni niti potpuni u odnosu na attribute ostalih izvora. Odredišni sustavi podacima iz registra mogu pristupiti samo za čitanje, dok mijenjati sadržaj registra ne mogu. Registarski stil je najjeftinija opcija rješavanja duplikata uz zadržavanje pristupa matičnim podacima. Međutim, performanse mogu biti problem u ovoj arhitekturi, jer se pozivi upućuju na svaki izvorni sustav posebno te čimbenici kašnjenja mogu brzo nastupiti [14].



Slika 2.5. *Registarski stil implementacije MDM-a [14]*

2.3.4.2. Konsolidacijski stil

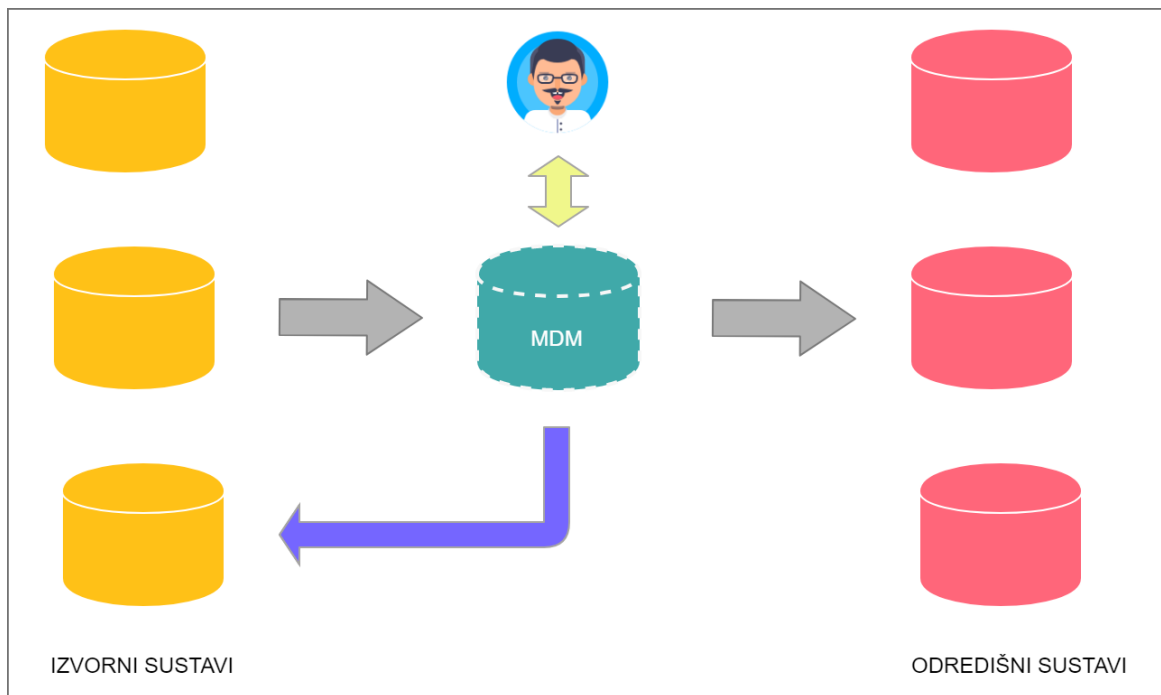
Konsolidacijski stil (engl. *consolidation style*), koji se koristi primarno za analizu i izvješća, prikazan je slikom 2.6. Ovaj stil okuplja matične podatke iz različitih izvora i objedinjuje ih kao jedinstvenu verziju istine. Česta praksa je da se prilikom implementacije ovog stila dodatno uključi i osoblje koje vodi računa da jedinstvena verzija istine bude još "istinitija" [14]. Drugim riječima, nakon što MDM pokrene algoritam čišćenja, podudaranja, povezivanja i spajanja prikupljenih podataka te uspostavi jedinstveni izvor istine, administrator ili ovlaštenu upravitelj podataka može ručno mijenjati sadržaj. Okupljene podatke unutar MDM sustava koriste odredišni sustavi za analitiku i izvještavanje. Ovaj stil omogućuje povlačenje matičnih podataka u odredišne sustav, ali se podaci ne vraćaju natrag u izvorne sustave. Drugim riječima, ne postoji sinkronizacija i veza je jednosmjerna. Konsolidirani stil jeftin je i brz način upravljanja podacima te olakšavanje kreiranja izvještaja unutar cijelog poduzeća.



Slika 2.6. *Konsolidacijski stil implementacije MDM-a [14]*

2.3.4.3. Koegzistencijski stil

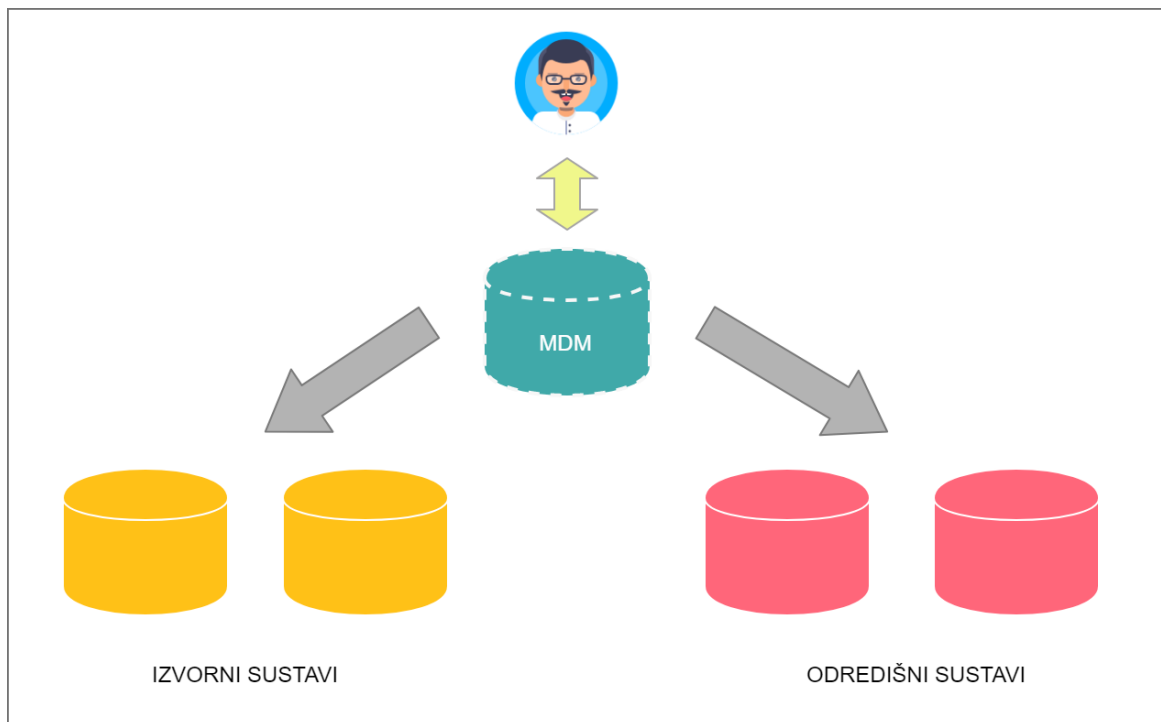
Koegzistencijski stil (engl. *coexistence style*), koji nastaje nadogradnjom konsolidacijskog stila, prikazan je slikom 2.7. Ovaj stil omogućuje uspostavljanje jedinstvene verzije istine po istom principu kao što to čini konsolidacijski stil, no razlikuje se po tome što podržava dvosmjernu vezu između MDM i izvornih sustava. Također, čišćenje i usklađivanje podataka se obavlja u izvornim sustavima, nakon čega se podaci prenose i pohranjuju u MDM sustav koji, podatke prosljeđuje ostalim vanjskim sustavima [15]. Kako bi se podržala dvosmjerna veza, mora postojati sinkronizacija između MDM i izvornih sustava, što koegzistencijski stil čini skupljim za implementaciju. Za ostvarivanje potpune konzistentnosti između svih vanjskih sustava, mora se prvo dogoditi prijenos i pohrana u MDM sustav, nakon čega se odvija distribucija u ostale sustave. To su dva koraka koja iziskuju vrijeme i tu je neizbježno kašnjenje. No, velika prednost ovog sustava je što je kvaliteta matičnih podataka značajno poboljšana i pristup je brži.



Slika 2.7. Koegzistencijski stil implementacije MDM-a [14]

2.3.4.4. Transakcijski/Centralizirani stil

Slika 2.8 prikazuje transakcijski (engl. *transaction style*), odnosno centralizirani stil implementacije MDM-a. Ovaj stil implementacije pohranjuje, održava i poboljšava attribute matičnih podataka. Nakon što se matični podaci prikupe, središte MDM-a postaje jedini izvor istine koji matične podatke nastavlja održavati kao takve. Izvorni i odredišni sustavi, koji su ovdje predstavljeni kao vanjski sustavi, nemaju mogućnost mijenjati vrijednosti matičnih podataka, već su se obavezni pretplatiti na potencijalne promjene. Ukoliko se, prilikom održavanja unutar MDM sustava, potencijalne promjene dogode, MDM sustav objavljuje vanjskim sustavima koji potom prihvaćaju ažuriranja [15]. Ključna razlika u odnosu na koegzistencijski stil je ta što se u transakcijskom stilu i ažuriranja i pohranjivanja podataka provode kroz MDM sustav i na taj način nema kašnjenja te se postiže najveća razina konzistentnosti. MDM sustav postaje dominantni izvor istine i omogućuje središnje nadgledanje kvalitete podataka. Uz ovaj stil, matični podaci su ujednačeni, točni i potpuni u svakom trenutku. Najveći nedostatak ovog stila je što iziskuje neizvjesno vrijeme za implementaciju, jer transakcijski stil zahtjeva složena pravila integracije, intenzivno usklađivanje podataka i sinkronizaciju. A također, prilikom implementacije ovog stila neizbježan je uslužno-orijentirani pristup [15].



2.8. Transakcijski stil implementacije MDM-a [14]

2.3.5. Odnosi između različitih stilova implementacije MDM-a

Stilovi implementacije upravljanja matičnim podacima očigledno se razlikuju prema složenosti te se mogu implementirati kao nadogradnja jedan na drugi. Svaki stil ima svoje karakteristike. Registarstil koristi se isključivo kao središnja referenca na izvorne sustave, konsolidacijski stil poseban je iz razloga što je njegova primarna upotreba za analitiku i kreiranje izvještaja, koegzistencijski stil služi za usklađivanje između različitih baza podataka, dok se transakcijski sustav ponaša kao sustav zapisa koji služi kao potpora transakcijskim sustavima.

Ono što stilovi imaju zajedničko je što se svaki sastoji od MDM sustava i vanjskih sustava. Vanjski sustavi razlikuju se kao izvorni i odredišni sustavi. Bitnu ulogu vanjskog sustava ima skladište podataka koje može utjecati na odabir stila implementacije MDM-a.

No ipak, kako je MDM je složeni proces koji uključuje ljude, alate, metode i politike, koji iskorištavanjem vrijednosti korporativnog informacijskog sredstva, oblikuju budućnost organizacije. Stoga i krajnji izbor rješenja implementacije MDM-a ovisan je o poslovnoj situaciji, ciljevima i strukturi organizacije.

Ovaj rad je nastao u suradnji sa Finom [16], velikom organizacijom koja svoje poslovanje temelji na skladištu podataka i MDM-u. Iz tog razloga će Fina poslužiti kao primjer pomoću kojeg se opisuje kako nastaje skladište podataka i potreba za implementacijom MDM-a te u kakvom su međusobnom odnosu ta dva pojma.

2.4. Primjena postupka upravljanja matičnim podacima u Fini

MDM, u usporedbi s njegovim prethodnicima, još je relativno mlad pojam. Standardna izvještavanja i analize, sredinom 90-ih godina zamijenila su skladišta podataka i poslovna inteligencija (engl. *Business Intelligence*), nakon čega su područja kvalitete i integracije podataka također evoluirala. Vremenom, osim za analitičke i integracijske potrebe, raste potreba za korištenjem podataka i u operativnom smislu. Tako se razvijaju mnoge MDM aktivnosti, koje postaju nužna potreba za moderne aktivnosti digitalne transformacije. Svaka organizacija koja shvaća vrijednost svoje informacijske imovine trebala bi posvetiti pažnju procesu upravljanja matičnim podacima.

Fina, kao velika organizacija, s gotovo 200 usluga, ima potrebu koristiti MDM na tri razine:

- Na razini zajedničkih šifarnika
- Na razini podataka korištenih od strane pojedinih usluga koje je potrebno povezati
- Na razini podataka važnih za vlastito poslovanje

MDM na razini zajedničkih šifarnika predstavlja potrebu za integracijom matičnog okruženja u kojem će se nalaziti svi šifarnici koje koriste Finine usluge. Takva vrsta MDM-a u Fini još ne postoji, već će se implementirati pomoću specifikacije koja je opisana kroz četvrto poglavlje rada.

Povezivanje podataka korištenih od strane pojedinih usluga podrazumijeva čišćenje i integraciju podataka iz različitih sustava, čiji ishod služi za izradu smislenih izvještaja koje također imaju na uvid i klijenti za vlastito poslovanje i ostale javne podatke. MDM na razini podataka korištenih od strane pojedinih usluga koje je potrebno povezati nije sustavno riješen, već funkcionira kroz nekoliko usluga koji prikupljaju podatke iz više različitih izvora.

Podaci koje Fina koristi za vlastito poslovanje bitni su za njezinu uspješnost i unapređivanje poslovnih procesa. Ukoliko se takvi podaci održavaju na kvalitetan i uredan način, pomoću njih se mogu dobiti vrlo bitne pravovremene informacije koje će olakšati poslovno odlučivanje, održati financijsku stabilnost, povećati konkurentnost te u konačnici dovesti do zadovoljstva klijenata. Konačan MDM na razini podataka važnih za vlastito poslovanje implementiran je kroz nekoliko prilika pojavljivanja novih sustava za Finino poslovanje, no prvi značajni pothvat u toj priči predstavlja skladište podataka. Što je to skladište podataka, koje su njegove primjene te kako je izgrađeno skladište podataka u Fini, opisuje se kroz treće poglavlje rada.

3. SKLADIŠTE PODATAKA

Čišćenjem podataka prikupljenih iz različitih izvornih sustava i njihovom integracijom nastaje skladište podataka - DW (engl. *Data Warehouse*). Postupak MDM i proces skladištenja podataka usko su povezani kada je riječ o integraciji podataka, no to su dva odvojena područja čiji se krajnji ciljevi znatno razlikuju. Skladište podataka sadrži povijesne podatke koji se smatraju neophodnim samo za analitičke potrebe, dok MDM sadrži trenutne podatke koje, osim za analitičke primjene, koristi i za operativne primjene. Iz tog razloga, podaci u skladištu podataka strukturiraju se i pohranjuju u obliku koji je pogodan za razne analize podataka i izrade izvještaja.

3.1. Definicija skladišta podataka

Potreba za skladištem podataka javila se iz razloga što podaci iz transakcijskih baza podataka nisu dovoljni za razne analize i izrade izvještaja. U transakcijskim bazama bilježe se samo trenutne vrijednosti, npr. cijene određenog proizvoda ili usluge, količine na skladištu, itd. Te se tu starim podacima, koji su bitni za vremenski redoslijed zbivanja pojedinih poslovnih događaja, gubi trag. U skladištima se ne arhiviraju svi, već samo pojedini podaci iz baze podataka za koje se smatra da su važni za analize. Takvi relevantni podaci se procesom integracije periodično prebacuju iz baze u skladište podataka. Cilj skladišta podataka nije operativnost poslovanja, kao što je to slučaj kod MDM-a, nego stvoriti što bogatiji izvor informacija za različite dugoročne ili kratkoročne analize i predviđanja. Iz tog razloga, podaci u skladištu podataka su trajni i vežu se uz vrijeme nastajanja.

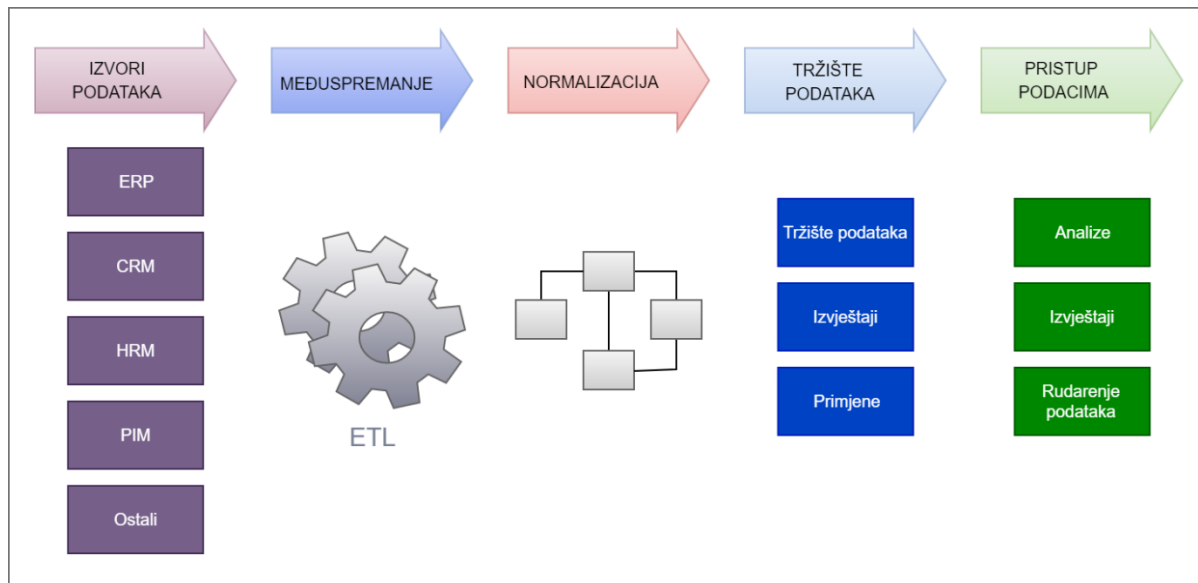
U bilo kojem kontekstu vezanom uz skladište podataka, neizbježno se pojavljuju dva imena - Bill Inmon i Ralph Kimball. To su dva pionira čiji pristupi značajno doprinose praksi oblikovanja skladišta podataka. Oba pristupa definiraju skladište podataka kao skup podataka prikupljenih iz različitih izvora koji služe kao potpora za analizu informacija diljem organizacije.

Prema Inmonu [17], skladište podataka sadrži sljedeće karakteristike:

- Subjektivno orijentiran – podaci su organizirani po poslovnim temama
- Integriran – podaci su pohranjeni kao jedinstvena cjelina
- Postojan – stari podaci u skladištu se ne brišu i ne mijenjaju, a novi se periodički dodaju
- Vremenski različit – skladište ima vremensku dimenziju koja omogućuje pregled podataka u vremenskom kontekstu

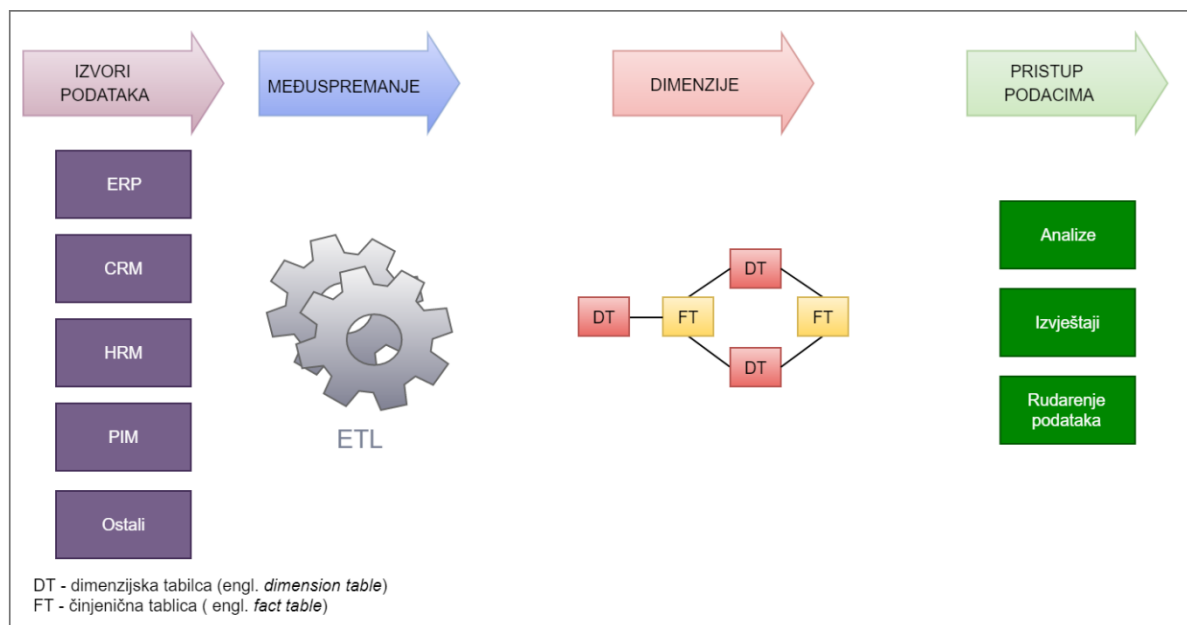
Inmonov pristup, prikazan slikom 3.2 koja je nastala prema uzoru na [18], poznat je još kao *Skladište podataka poduzeća* (engl. *Enterprise Data Warehouse*). Inmon definira skladište podataka kao centralizirani repozitorij cjelokupnog poduzeća. Kako bi podaci iz izvornih sustava,

kao što su ERP, CRM, itd., pristigli u centralni repozitorij poduzeća, prvo se odvija ETL proces integracije, koji je detaljnije objašnjen u odlomku 3.2. Unutar skladišta podataka koristi se relacijski model koji podrazumijeva normalizaciju podataka, nakon čega se posebno oblikuju dimenzijska tržišta podataka (engl. *Data Marts*). Tržišta podataka sadrže podatke specijalizirane za određene poslovne procese ili odjele, a služe kako bi grupiranjem jasnije predočili podatke iz skladišta podataka. Takvi podaci dostupni su za razne analize, izvještaje i tehnike rudarenja podataka (engl. *Data Mining*).



Slika 3.1. Inmonova arhitektura skladišta podataka [18]

Kimballov pristup, prikazan slikom 3.2 koja je nastala prema uzoru na [18], poznat je također i pod nazivom *Dimenzijsko skladište podataka* (engl. *Dimensional Data Warehouse*). Kimball definira skladište podataka kao kopiju transakcijskih podataka posebno strukturiranu za upit i analizu [18]. Nakon prikupljanja i integracije podataka iz različitih izvornih sustava, koristi se dimenzijski model oblikovanja skladišta podataka. Tržišta podataka se ne odvajaju zasebno od skladišta podataka, kao što je slučaj kod Inmonovog pristupa.



Slika 3.2. Kimballova arhitektura skladišta podataka [18]

Skladištenje podataka je skup metoda i alata koji obrađuju podatke iz skladišta podataka, a odvija se u tri koraka:

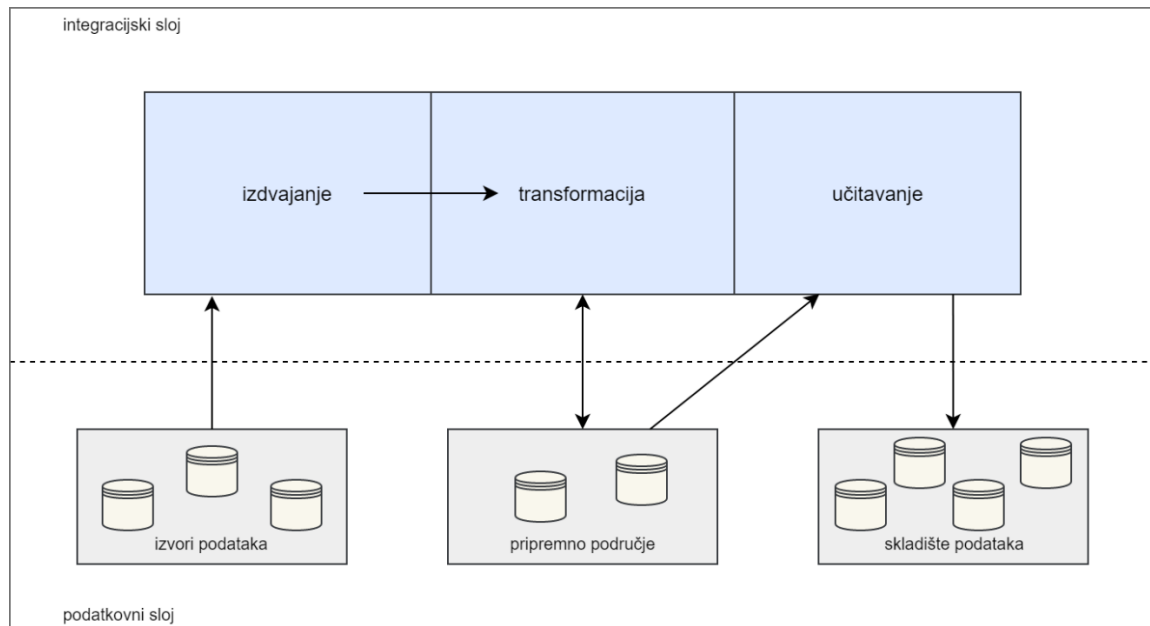
1. ETL – izdvajanje, transformacija i učitavanje podataka
2. Pohranjivanje podataka u skladište
3. Daljnja uporaba podataka iz skladišta

3.2. ETL proces integracije

ETL je temeljni i ujedno najkompliciraniji proces skladištenja podataka. To je proces integracije koji nad podatkovnim slojem (engl. *data layer*) primjenjuje metode upravljanja kvalitetom podataka [19]. Kao što je prikazano slikom 3.3, ETL započinje identifikacijom izvora podataka iz kojih se izdvajaju podaci. Izdvojeni podaci se dalje po potrebi čiste i ujedanjuju te zatim transformiraju i učitavaju u skladište podataka. Proces od izdvajanja do učitavanja poznat je kao među-spremanje podataka kroz pripremno područje – DSA (engl. *Data Staging Area*) u kojem se podaci prikupljaju i obrađuju. Moguća su dva načina obrade podataka: obrada toka podataka, gdje podaci jedan za drugim ulaze u proces transformacije, i obrada gomile podataka gdje se funkcije transformacije izvode nad svim podacima istovremeno, što traje mnogo duže [20]. Ovaj složeni proces posvećuje pozornost kvaliteti i integritetu podataka, pri čemu je potrebno definirati postupke u slučaju pogrešaka, tehnike inkrementalnog punjenja skladišta podataka i slično.

Transformirani podaci se zatim pohranjuju u bazu skladišta podataka koja mora osigurati brz i učinkovit pristup podacima, sigurnost i tajnost podataka, mogućnost povećanja kapaciteta pohranjenih podataka, te cjelokupnu raspoloživost sustava. Podaci su potom spremni za daljnju

uporabu, koja, u slučaju skladišta podataka, podrazumijeva postavljanje upita, izradu izvještaja i grafova, analizu podataka, te otkrivanje raznih zakonitosti u podacima.

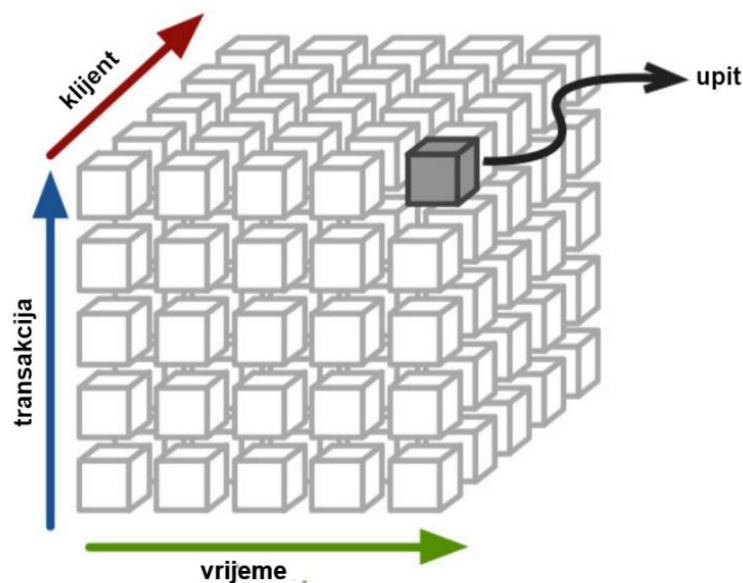


Slika 3.3. Arhitektura tradicionalnog ETL procesa

Nakon inicijalnog punjenja skladišta podataka, programi za inkrementalno punjenje pokreću se periodički, npr. dnevno, tjedno ili mjesečno te pune skladište podataka novim relevantnim podacima iz transakcijske baze podataka.

3.3. Dimenzijsko oblikovanje skladišta podataka

Kao rezultat potrebe za jednostavnošću i učinkovitošću relacijskog modela pojavio se dimenzijski oblik spremanja podataka u skladište podataka, u kojem se svako funkcionalno područje poslovanja može se promatrati kroz više dimenzija. Prikazano slikom 3.4 koja je nastala prema [21], na skladište podataka šalje se upit za željenom transakcijom koja se promatra prema klijentu i vremenu. Nakon što skladište primi upit, podaci u skladištu podataka se formiraju kao trodimenzionalna kocka koja vraća podatke o transakciji, povezanim klijentima te datum transakcije. Određena dimenzija može imati više hijerarhijskih razina. Primjerice, dimenzija datum može imati godine, mjeseci i dane kao hijerarhijske razine.



Slika 3.4. Oblikovanje upita pomoću dimenzijskog modela skladišta podataka [21]

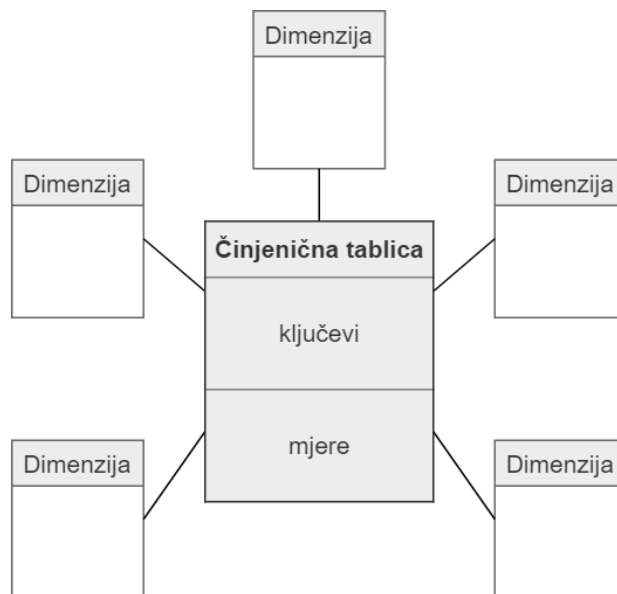
Svaki dimenzijski model sastoji se od jedne činjenične tablice (engl. *fact table*) i niza manjih dimenzijskih tablica (engl. *dimension table*) koje daju kontekst činjeničnoj tablici. Činjenična tablica sadrži kompozitne primarne ključeve koji odgovaraju ključevima dimenzijskih tablica s kojima je u spoju te numeričke atribute poznate kao mjere (engl. *measures*) koje pomoću agregatnih funkcija daju rezultat procesa koji se prati.

Još jedno važno svojstvo činjeničnih tablica je zrnatost (engl. *granularity*). Zrnatost se odnosi na razinu agregacije podataka u skladištu podataka. Što su podaci detaljnije predstavljeni, to je razina zrnatosti manja. Dok što su podaci manje detaljno predstavljeni, to je razina zrnatosti veća [22].

Spoj činjenične tablice sa dimenzijskim tablicama tvori neki od dva modela dimenzijskog oblikovanja skladišta podataka:

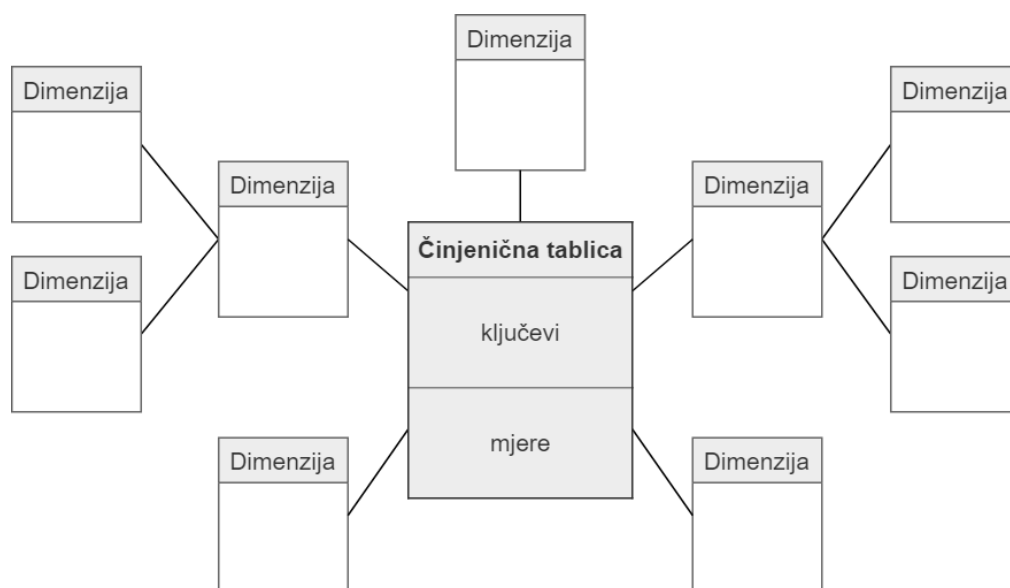
- Zvezdasti model ili zvezdasti spoj (engl. *star-join*)
- Pahuljasti model (engl. *snowflake model*)

U središtu zvezdastog modela, prikazanog na slici 3.5 koja je nastala prema uzoru na [22], nalazi se činjenična tablica koja odgovara procesu koji se prati u skladištu podataka. Određeno skladište podataka može imati N zvezdastih spojeva i N činjeničnih tablica.



Slika 3.5. Zvezdasti model [22]

Pahuljasti model nastaje kao dekompozicija jedne ili više dimenzijskih tablica. Točnije, odnosi jedan-prema-više u dimenzijskoj tablici mogu rezultirati novom tablicom, čime se formira hijerarhija koja tvori pahuljasti model. Slika 3.6 koja je nastala prema uzoru na [22], prikazuje pahuljasti model.



Slika 3.6. Pahuljasti model [22]

Iako pahuljasti model daje dobru vizualizaciju hijerarhije, on ipak ne treba biti prvi izbor za korištenje iz razloga što umanjuje lakoću korištenja i performanse pretraživanja. Prilikom obavljanja upita kod pahuljastog modela potrebno je spojiti više tablica što automatski usporava rad, model je složeniji i teže se pristupa podacima. U praksi se puno češće koristi zvezdasti model te su podijeljena mišljenja treba li uopće koristiti pahuljasti model [22].

Pomoću dimenzijskog modela oblikovano je također i skladište podataka u Fini, čijih je nekoliko ključnih dimenzija opisano u sljedećem odlomku. Finino skladište podataka izgrađeno je prije desetak godina te do danas njegovo postojanje je Fini kao organizaciji omogućilo generiranje raznih analiza i poslovnih izvještaja, implementaciju novih sustava, kao također i uvođenje MDM-a na razini podataka važnih za vlastito poslovanje.

3.4. Primjer oblikovanja skladišta podataka u Fini

U nastavku su prikazani primjeri formiranja nekoliko dimenzija Fininog skladišta podataka koje se održavaju metodama tipičnim za MDM. Iako skladište podataka ne podržava daljnju distribuciju podataka, dimenzije su, zahvaljujući integraciji podataka iz različitih izvora, poslužile kao ključne prilikom implementacije određenih MDM rješenja u Fini.

Uz dodatak iz tehničke dokumentacije koja se odnosi na punjenje dimenzije *Klijent*, dimenzije su opisane na temelju funkcionalne dokumentacije Fininog skladišta podataka. U vrijeme kada je građeno Finino skladište podataka, još nisu postojali sustavi koji su mogli poslužiti kao referentni, nego su problemi, kao što je npr. identifikacija osobe, rješavani na puno složenije načine. Iz tog razloga se u dimenzijskim tablicama pojavljuje pojam poput JMBG-a¹.

3.4.1. Dimenzija “Klijent”

Dimenzija *Klijent* je najsloženija dimenzija iz Fininog skladišta podataka. Dimenzija opisuje sve fizičke i pravne osobe koje trenutno jesu ili su u prošlosti bile korisnici usluga Fine.

Prema poslovnom značenju, atributi tablice *Klijent* mogu se podijeliti u sljedeće skupine:

- Matični atributi – osobni podaci klijenta kao što su npr. ime, matični broj, datum rođenja, itd.
- Demografski atributi – mogu biti npr. spol, adresa, itd.
- Poslovno-segmentacijski atributi – smještaju klijenta u skupinu u skladu s njegovim osobinama, statusom te vrstom poslovnog odnosa s Finom, npr. djelatnost klijenta prema NKD², veličina poslovnog subjekta, itd.
- Ponašajni (engl. *behavior*) atributi – opisuju povijest poslovnog odnosa Fine s klijentom, npr. datum prve/zadnje transakcije, datum prestanka poslovanja, internu procjenu klijenta, VPO³ klijenta, itd.

¹ JMBG - Jedinstveni matični broj građana

² NKD – Nacionalna klasifikacija djelatnosti

³ VPO – Voditelj poslovnog odnosa

Tablica 3.1 prikazuje dimenzijsku tablicu *Klijent*. U tablici su prikazani nazivi atributa koji se prate te njihov tip, opis, izvor i povijest. Tip atributa označen oznakom ID odnosi se na identifikacijski atribut, IA predstavlja informacijski atribut, D dimenziju, a QA je kvalifikacijski atribut koji klijenta svrstava kao subjekt u grupu s ostalim subjektima s kojima dijeli zajedničko svojstvo te predstavlja osnovu za segmentacijsku analizu subjekta po više kriterija.

Atributi klijenta najvećim dijelom prate informacije koje se o klijentu vode u izvornom sustavu JRIR⁴. Međutim, kako bi se upotpunila slika, koriste se i ostali izvorni sustavi kao što su PIS⁵, REGOS⁶, RGFI⁷, RKP⁸. Znatost tablice *Klijent* određena je prema matičnom broju klijenta te nema eksplicitne hijerarhije.

SCD (engl. *Slowly Changing Dimensions*) opisuje način održavanja dimenzija koje se u skladištu podataka mijenjaju postupno s vremenom. SCD podržava nekoliko tipova koji se razlikuju ovisno načinu konsolidacije povijesnih i trenutnih podataka. Mogući tipovi SCD-a su SCD0, SCD1, SCD2, SCD3, SCD4 i SCD6. Dimenzija *Klijent* praćenje povijesti provodi pomoću SCD2, odnosno čuva i trenutne i povijesne podatke u tablici. Trenutni podaci od povijesnih razlikuju se pomoću oznake, gdje su trenutni podaci označeni oznakom "1", a povijesni oznakom "0".

Tablica 3.1. Izgled tablice "Klijent"

Naziv	Tip	Opis	Izvor	Povijest
Šifra klijenta	ID	Jedinstveni broj (šifra) pod kojim se klijent prati.	Interno DW	-
Šifra u RGFI-u	ID	Ključ iz RGFI-a.	RGFI	ne
Šifra u PIS-u	ID	Šifarnik u PIS-u, preuzet iz RGFI-a.	PIS	ne
Matični broj	IA	JMBG osobe, odnosno MB poslovnog subjekta.	JRIR ,PIS, REGOS	ne
Naziv	IA	Puni naziv tvrtke za pravne osobe, odnosno Ime + Prezime za fizičke osobe.	JRIR, PIS, REGOS	ne
Adresa sjedišta	IA	Adresa (ulica i broj) prebivališta za fizičke osobe, odnosno adresa sjedišta za pravne osobe.	JRIR, RGFI	ne

⁴ JRIR – Jedinstveni registar računa

⁵ PIS – Poslovno informacijski sustav

⁶ REGOS – Središnji registar osiguranika

⁷ RGFI – Registar godišnjih financijskih izvještaja

⁸ RKP – Registar korisnika proračuna

Mjesto sjedišta	IA	Naziv naselja.	JRIR, RGFI	ne
Vrsta klijenta	QA	Vrsta klijenta prema internoj definiciji (npr. fizička osoba, tvrtka, obrt, banka).	-	da
Tip klijenta	DD	Banka (ima VBDI), korisnik proračuna (u RKP-u, SDR-u) ili fond.	-	-
Rezidentnost	QA	Rezident ili nerezident.	PIS	-
Spol	QA	Spol fizičke osobe: M, Ž, nepoznat, neprimjenjivo	Izračunato iz JMBG-a	ne
Datum rođenja	IA	Datum rođenja osobe (neprimjenjiv za pravne osobe, nepoznat za nerezidente).	Izračunato iz JMBG-a	ne
Datum prve transakcije	IA	Prva transakcija klijenta.	Transakcijski izvori	ne
Datum zadnje transakcije	IA	Zadnja transakcija klijenta.	Transakcijski izvori	ne
Datum zadnje promjene	IA	Datum zadnje promjene matičnih podataka klijenta.	Izračunato	ne
Datum osnivanja	IA	Datum od kojega posluje poslovni subjekt.	RGFI	-
Datum prestanka poslovanja	IA	Datum prestanka poslovanja poslovnog subjekta.	RGFI	-
Vezani poslovni subjekt	IA	Ključ poslovnog subjekta veze: prazan za samostalni subjekt; ključ preuzimatelja ako je pripojen; prazan za preuzimatelja; ključ novonastalog subjekta ako je spojen; prazan za subjekta nastalog spajanjem; prazan za podij.	RGFI	-
Oblik organiziranja	QA	Oblik organiziranja prema šifrniku NKD.	JRIR	da
Status klijenta	QA	Status klijenta.	RGFI	da
Interni status klijenta	QA	Interni status ili segmentacija klijenta.	-	da
Veličina poslovnog subjekta	QA	Segmentacija poslovnog subjekta prema veličini, u skladu s poslovnom definicijom banke (građanstvo, SME, velika i javna poduzeća, itd.).	RGFI	da
Radni status	QA	Umirovljenik, ostalo, nepoznato, neprimjenjivo..	REGOS	da

Aktivnost RGFI	QA	Oznaka aktivnosti: 1-aktivan; 0-neaktivan.	RGFI	-
Porijeklo kapitala	QA	Porijeklo kapitala: 1- domaći, 2- strani, 3- mješoviti.	RGFI	-
Oznaka likvidacije ili stečaja	QA	Oznaka likvidacije ili stečaja: 0-ništa, 1-likvidacija 2-stečaj.	RGFI	-
Oznaka statusa poslovnog subjekta	QA	Status poslovnog subjekta: 0-samostalni subjekt; 1-pripojen drugom subjektu; 2-preuzimatelj nekog subjekta; 3-spojen u novi subjekt; 4-nastao spajanjem subjekata; 5-podijeljen na više subjekata; 6-nastao dijeljenjem.	RGFI	-
Vrsta fonda	QA	Vrsta mirovinskog fonda: 1-obavezno, 2-dobrovoljno, Nije primjenjivo za ostale.	REGOS	ne
Glavni Račun	D	Žiro račun.	JRIR	da
Glava Korisnika proračuna	D	Glava + razdjel	RKP	--
Uloga klijenta	D	Partner (saldna konti), Klijent (nema fakture), Korisnik (faktura u PIS-u).	PIS	da
Organizacijski dio	D	Organizacijski dio banke kojem klijent pripada.	JRIR	da
VPO	D	Voditelj poslovnog odnosa.	HR Plus	da
Klijentska djelatnost	D	Obuhvaća hijerarhiju prema NKD.	JRIR	da
Geografsko područje	D	Geografsko područje prebivališta, odnosno sjedišta; obuhvaća poštanski broj, poštu, naselje, općinu, županiju, državu.	JRIR	da
Oblik organizacije	D	Oblik organizacije.	RGFI	-
Oznaka vlasništvo	D	Vlasništvo po RGFI.	RGFI	-

Tablica *Klijent* ima dvostruku ulogu, kao činjenična i kao dimenzijska tablica. Kao činjenična tablica predstavljena je svim dimenzijama s kojima je povezana pomoću kompozitnih primarnih ključeva, dok kao dimenzija služi transakcijskim tablicama kako bi se povezale sa svojim klijentima.

3.4.1.1. Punjenje dimenzije „Klijent“

Punjenje dimenzije *Klijent* opisano je putem tehničke dokumentacije Fininog skladišta podataka. Prema definiciji iz funkcionalne specifikacije, dimenzija *Klijent* generira se iz nekoliko različitih izvornih sustava. Prvi izvor za generiranje podataka o klijentu je JRIR, jer kako bi klijent mogao poslovati u Fini, on prvenstveno mora otvoriti klijentski račun. Daljnje informacije o klijentu nadopunjuju se iz sustava RGFI i REGOS. Korisnicima državnog proračuna kojima se podaci o glavi i razdjelu ne nalaze u sustavu RGFI, ta će se informacija morati nadopuniti iz sustava RKP. Podaci o adresi će se naknadno aktualizirati podacima iz sustava PIS iz razloga što su u tom sustavu oni najaktualniji.

Zbog složenosti ove dimenzije, vrlo je bitan redoslijed izvođenja njezinog punjenja. Redoslijed se, uz pretpostavku da prethodno su napunjene sve izvorne tablice, sastoji od niza koraka. Kada se podaci prikupe iz izvornih tablica, tada se podaci čiste i ujedinjuju u normaliziranoj tablici unutar pripremnog područja (engl. *stage area*). Uz to se pokreće procedura specijalizirana za punjenje "umjetnih" klijenata iz sustava JRIR, nakon čega se pomoću određene vezne tablice klijenti povezuju s njihovim otvorenim računima. Takvi podaci se potom pohranjuju u dimenziju *Klijent* koja unutar skladišta podataka prati povijest izmjena nad SCD2 atributima.

Kao posljedica integracije dimenzije *Klijent* iz više različitih izvora, pojavljuje se više poslovnih ključeva BK (engl. *Business Key*). Inicijalna ideja povezivanja podataka bila je da matični broj (MB/JMBG) bude šifra za povezivanje, no koncept je promijenjen spoznajom da u pojedinom izvornom sustavu, kao što je npr. sustav PIS u tablici 3.3, postoji više zapisa s istim matičnim brojem koji se kao takvi moraju očuvati u dimenziji jer se određene transakcije vežu konkretno na šifru iz izvornog sustava. Stoga je odlučeno da se spajaju samo oni klijenti kod kojih postoji samo jedan redak s istim matičnim brojem u izvornom sustavu.

Punjenje dimenzije *Klijent* prikazano je kroz nekoliko stvarnih primjera opisanih tablicama 3.2-3.5 čije oznake predstavljaju:

- BSKEY – identifikator logičkog reda koji, ovisno o sustavu iz kojega je klijent inicijalno stigao, dobiva prefiks s pripadajućom šifrom
- LTRIM_MBR – matični broj s uklonjenim vodećim nulama
- GLAVNI_MBR – oznaka za prepoznavanje glavnog ("1") od naknadno dodanog retka ("0") koja vrijedi za isti LTRIM_MBR
- SIFRA_KLIJENTA_RGFI – BK za klijenta iz sustava RGFI
- SIFRA_KLIJENTA_PIS – BK za klijenta iz sustava PIS

U tablici 3.2 prikazan je primjer spajanja podataka gdje je klijent inicijalno stigao iz sustava RGFI. U sustavu PIS postoji samo jedan klijent s istim matičnim brojem, pa su oba zapisa spojeni u jedan. Ovo je najjednostavniji te ujedno i najčešći slučaj povezivanja.

Tablica 3.2. Spajanje podataka o klijentu iz sustava RGFI i sustava PIS

BSKEY	LTRIM_MB R	GLAVNI_MB R	SIFRA_KLIJENTA _RGFI	SIFRA_KLIJENTA _PIS	NAZIV
R-142882	3934187	1	142882	000011	TRGOVINA d.o.o.

Tablica 3.3 prikazuje primjer gdje je klijent inicijalno stigao iz sustava RGFI, no u sustavu PIS postoji više klijenata s istim matičnim brojem. Zadržan je redak iz sustava RGFI te su dodani svi reci iz sustava PIS. Nije provedeno spajanje retka s oznakom „glavni_mbr“ jer ne postoji jedinstveni identifikator glavnog (hijerarhijski najvišeg) retka u sustavu PIS.

Tablica 3.3. Dodavanje podataka o klijentu iz sustava RGFI i sustava PIS

BSKEY	LTRIM_MB R	GLAVNI_MB R	SIFRA_KLIJENTA _RGFI	SIFRA_KLIJENTA _PIS	NAZIV
R-104667	3276147	1	104667		OSIGURAVAJUĆE DRUŠTVO d.d.
P-000784	3276147	0		000784	OSIGURAVAJUĆE DRUŠTVO d.d. Filijala Sisak
P-000786	3276147	0		000786	OSIGURAVAJUĆE DRUŠTVO d.d. Filijala Split
P-000787	3276147	0		000787	OSIGURAVAJUĆE DRUŠTVO d.d. Filijala Šibenik
P-019828	3276147	0		019828	OSIGURAVAJUĆE DRUŠTVO d.d. Osig.prijevoza i kredita

Tablica 3.4 prikazuje primjer gdje u sustavu računa (JRIR) postoji matični broj koji ne postoji niti u jednom izvornom sustavu za klijenta te je iz tog razloga klijent u tablici umjetno stvoren.

Tablica 3.4. Umjetno stvaranje klijenta ukoliko se matični broj prvi put pojavljuje

BSKEY	LTRIM_MB R	GLAVNI_MB R	SIFRA_KLIJENTA _RGFI	SIFRA_KLIJENTA _PIS	NAZIV
D-1055769	983373	1			KOPIRAONA d.o.o.

Tablica 3.5 prikazuje primjer gdje u sustavu računa (JRIR) ne postoji matični broj, odnosno matični broj je 0 ili null, te je iz tog razloga klijent u tablici umjetno stvoren.

Tablica 3.5. Umjetno stvaranje klijenta ukoliko je matični broj u potpunosti nepoznat

BSKEY	LTRIM_MB R	GLAVNI_MB R	SIFRA_KLIJENTA _RGFI	SIFRA_KLIJENTA _PIS	NAZIV
D-1054394	DWH- 1054394	1			ZGRADA STJEPANA RADIĆA 10b

3.4.2. Dimenzija “Organizacijski dio”

Dimenzija *Organizacijski dio* definira osnovnu organizacijsko-poslovnu jedinicu Fine i hijerarhijsku strukturu. Znatost tablice određena je prema jedinstvenoj šifri organizacijske ili poslovne jedinice. Povijest se prati pomoću SCD2 praćenja promjene u hijerarhiji, a postoje dvije paralelne hijerarhije:

1. Sektor – Centar – Podružnica – Poslovnica
2. Sektor – Centar – Odjel – Odsjek

Prva hijerarhija opisuje poslovnu mrežu Fine koja služi za raspoređivanje pripadnosti djelatnika organizacijskim jedinicama. Dok, druga hijerarhija opisuje Finine potrebe određivanja mjesta pružanja usluga Centara proizvoda. Ove dvije hijerarhije preklapaju se samo unutar Sektora poslovne mreže. Tablica 3.6 opisuje dimenziju *Organizacijski dio*.

Tablica 3.6. Izgled tablice „Organizacijski dio“

Naziv	Tip	Opis	Izvor
Šifra organizacijskog dijela	ID	Jedinstvena šifra organizacijskog dijela.	Tarifa
Šifra organizacijskog dijela PIS	ID	Jedinstvena šifra koja se koristi u PIS-u.	HR Plus
Središnji ured	QA	Početna razina hijerarhije.	HR Plus
Šifra sektora	ID	Jedinstvena šifra sektora.	HR Plus
Sektor	QA	Sektor Fine.	HR Plus

Šifra centra	ID	Jedinstvena šifra centra.	HR Plus
Centar	QA	Centar Fine (regionalni centar za poslovnu mrežu).	HR Plus
Šifra podružnice	ID	Jedinstvena šifra podružnice.	HR Plus
Podružnica	QA	Podružnica Fine.	HR Plus
Šifra poslovnice	ID	Jedinstvena šifra poslovnice.	HR Plus
Poslovnica	QA	Poslovnica Fine.	HR Plus
Šifra odjela	ID	Jedinstvena šifra odjela.	HR Plus
Odjel	QA	Odjel	HR Plus
Šifra odsjeka	ID	Jedinstvena šifra odsjeka.	HR Plus
Odsjek	QA	Odsjek.	HR Plus

3.4.3. Dimenzija “Djelatnik”

Dimenzija *Djelatnik* definira osobe koje su zaposlenici Fine. Znatost tablice određena je prema jedinstvenoj šifri djelatnika Fine. Ne postoji eksplicitna hijerarhija, a praćenje povijesti provodi se pomoću SCD2. Tablica 3.7 opisuje dimenziju *Djelatnik*.

Ova dimenzija ima više uloga, koje su u skladu sa zadacima koje djelatnik Fine može imati u poslovnom procesu:

1. Klijent – opisuje voditelja poslovnog odnosa (VPO), ako djelatnik ima ovu ulogu tada se povezuje s dimenzijom „Klijent“
2. Ugovor – opisuje djelatnika Fine koji je kreirao ugovor
3. Transakcija – opisuje koji djelatnik je odradio transakcije u izvornim sustavima

Dimenzija *Djelatnik* ima ulogu činjenične tablice u vezi sa dimenzijom *Organizacijski dio*.

Tablica 3.7. Izgled tablice „Djelatnik“

Naziv	Tip	Opis	Izvor	Povijest
Šifra zaposlenika	ID	Jedinstveni broj (šifra) pod kojim se djelatnik vodi u HR aplikaciji.	HR Plus	ne
Organizacijski dio	D	Organizacijska jedinica u kojoj djelatnik radi.	HR Plus	da
Odjel	QA	Odjel.	HR Plus	da
Služba	QA	Služba u kojoj djelatnik radi.	HR Plus	da
Neposredni rukovoditelj	D	ID neposrednog menadžera (hijerarhija).	HR Plus	da
Ime	IA	Ime djelatnika.	HR Plus	ne
Prezime	IA	Prezime djelatnika.	HR Plus	ne
Datum rođenja	IA	Datum rođenja djelatnika.	HR Plus	ne
Datum zapošljavanja FINA	IA	Datum zaposlenja u Fini.	HR Plus	ne
Datum prvog zaposlenja	IA	Datum prvog zaposlenja.	HR Plus	ne
Šifra radnog mjesta	ID	Jedinstvena šifra radnog mjesta.	-	da
Naziv radnog mjesta	QA	Pozicija unutar hijerarhije.	HR Plus	da
Stručna sprema	QA	Stručna sprema zaposlenika.	HR Plus	da
Grupa vještina	D	Veza na dimenziju s više vrijednosti.	HR plus	-
Adresa	IA	Adresa djelatnika.	HR Plus	-
Geografsko područje	D	Mjesto i poštanski broj.	HR Plus	da
Razlog promjene	DD	Razlog promjene SCD2 atributa.	-	-

3.4.4. Dimenzija “Likvidator”

Dimenzija *Likvidator* definira osobe koje izvršavaju transakcije unutar Fine. Znatost tablice određena je prema jedinstvenoj šifri likvidatora te nema eksplicitne hijerarhije. Praćenje povijesti provodi se pomoću SCD1, odnosno održavaju se samo trenutni podaci, a povijesni se brišu. Tablica 3.8 opisuje dimenziju *Likvidator*.

Tablica 3.8. Izgled tablice „*Likvidator*“

Naziv	Tip	Opis	Izvor	Povijest
Šifra likvidatora	ID	Jedinstveni broj (šifra) pod kojim se djelatnik vodi.	PP, NKS, ...	ne
Šifra sustava	QA	Šifra sustava u kojem je likvidator definiran.	PP, NKS, ...	ne
Korisničko ime	IA	Korisničko ime djelatnika u produkcijskom sustavu.	PP, NKS, ...	ne
JMBG	IA	Matični broj.	PP, NKS, ...	ne
Vrsta likvidatora	QA	Automatski, Sistemski, Ljudski.	-	ne

Dimenzija *Likvidator* uparuje s sa dimenzijom *Djelatnik* pomoću posebne vezne tablice (engl. *mapping table*). Na takav način usluga (engl. *service*) prepoznaje korisničko ime djelatnika koji je likvidator. Dugo vremena se ta vezna tablica između djelatnika i likvidatora mogla se održavati samo ručno, no uvođenjem elektroničke prijave u Finine usluge riješen je i taj problem.

3.5. Interakcija između skladišta podataka i MDM-a

Ono što skladište podataka i postupak MDM imaju zajedničko je potreba za očuvanjem kvalitete podataka, kojom se upravlja pomoću procesa integracije. ETL proces integracije podataka, koji se koristi prilikom izgradnje skladišta podatka, može također podržati protok podataka iz izvornih sustava u matično okruženje, te na takav način skladište podataka može ujedno poslužiti kao dobar temelj za implementaciju MDM postupka. Upravo to je slučaj navedenih primjera formiranja dimenzija Fininog skladišta podataka. Dimenzije su integrirane procesom koji je specifičan za MDM okruženje, ali kako je skladište podataka namijenjeno analitičkim poslovnim potrebama, a ne transakcijskim zahtjevima, tako je nesposobnost upravljanja matičnim podacima unutar arhitekture samog skladišta podataka, kao i nemogućnost daljnje distribucije podataka, dovelo do potrebe za uvođenjem MDM procesa.

Dodatni razlog zbog kojeg skladište podataka može poslužiti kao "okidač" za uvođenjem MDM-a je što pravilno implementiran MDM može postati sustav koji osigurava točne i pravovremene informacije unutar skladišta podataka kao odredišnog sustava. Na taj način se drastično poboljšava cjelokupna kvaliteta podataka te se smanjuje veličina, opseg i složenost arhitekture skladišta podataka. Također, svaki podatak iz transakcijske baze podataka bude već cjelovit i točan prije pohrane u skladište podataka, time se umanjuje napor koji je unutar skladišta podataka potreban za integriranje transakcijskih podataka [23]. Na takav način se skladište podataka i MDM međusobno podupiru.

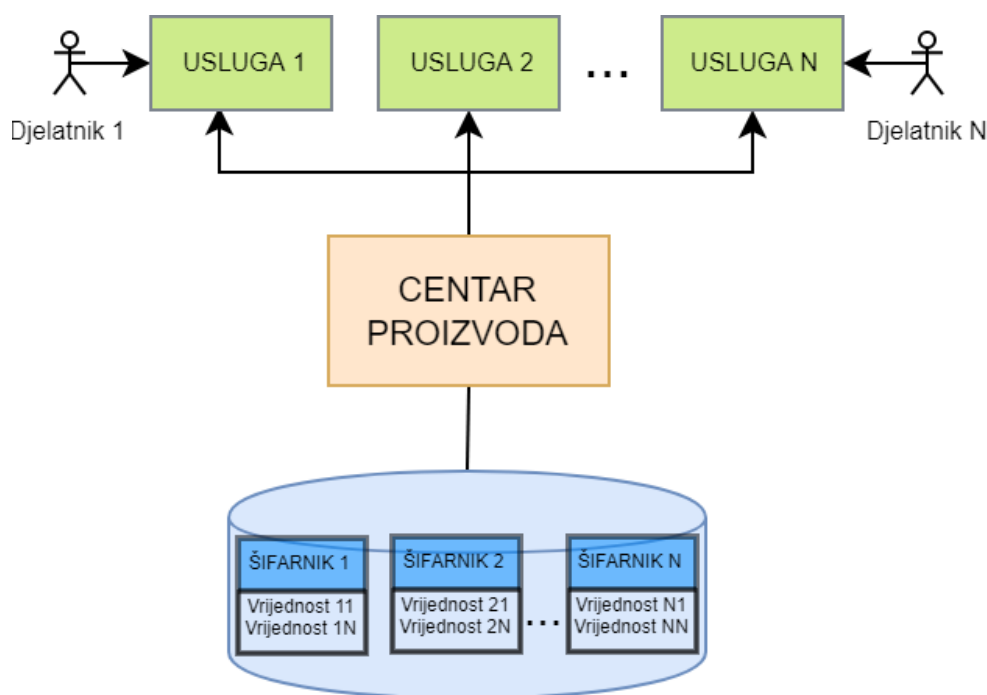
Implementacija upravljanja matičnim podacima je vrlo mukotrpan pothvat koji zahtjeva puno vremena te se svaka migracija vezana uz MDM proces mora odvijati modularno. To je razlog zbog kojeg ga većina organizacija ne uspije do kraja pravilno implementirati. U Fini se implementacija MDM-a odvijala kroz nekoliko pokušaja. Neki pokušaji su imali pozitivnog ishoda, a neki su ostali samo "pokušaj". Bez obzira na navedene činjenice, Fina, kao i svaka velika organizacija koja želi iskoristiti prednosti većih performansi te agilnije i jeftinije arhitekture skladišta podataka, trebala bi shvaćati važnost MDM postupka i pravilno se pridržavati protokola prilikom njegovog uvođenja.

S obzirom da sve Finine transakcije koriste podatke koji se nalaze unutar različitih šifrnika, nepostojeći šifrnici može biti vrlo poguban za izvršenje određene transakcije. Da bi se to izbjeglo, svaki sustav unutar Fina mora imati jedinstvenu i najaktualniju verziju šifrnika kojeg koristi za potrebe svojih poslovnih procesa. Iz tog razloga se javila potreba za implementacijom MDM-a na razini Fininih zajedničkih šifrnika. Taj problem se kroz sljedeće poglavlje rada opisuje pomoću specifikacije tzv. Središnjeg sustava šifrnika.

4. DIZAJN SREDIŠNJEG SUSTAVA ŠIFRARNIKA

Svaka usluga koju Fina pruža pripada određenom Centru proizvoda koji je bitan element Finine organizacijske strukture. Osim što je svaki Centar proizvoda vlasnik određenog broja usluga, on također ima svoje djelatnike, posjeduje vlastitu bazu podataka te funkcionira kao zaseban sustav unutar kojeg vodi računa o kopijama šifrnika koji su mu potrebni za izvršavanje različitih transakcija i unapređivanje vlastitog poslovanja.

Na slici 4.1 prikazani su elementi Centra proizvoda. Centar proizvoda posjeduje određen broj usluga i vlastitu bazu podataka (engl. *database*) u kojoj se nalaze kopije potrebnih šifrnika. Pojedine usluge Centra proizvoda funkcioniraju pomoću određenih mrežnih usluga (engl. *web services*), dok za ostale usluge Centra proizvoda zaduženi su njegovi djelatnici.



Slika 4.1. Prikaz elemenata Centra proizvoda

Šifarnici su tablice u kojima se nalazi popis svih mogućih vrijednosti određenih stavki, od kojih je svaka vrijednost definirana kroz jedinstvenu šifru i naziv. Primjerice, šifrnici *Država* sadrži popis država, od kojih svaka država ima vlastitu šifru koja joj služi kao jedinstveni identifikator u procesima vezanim uz taj šifrnici. Važnost šifrnika očituje se prema tome ukoliko je određena vrijednost unutar šifrnika neispravna ili zastarjela, svaka transakcija koja se veže za taj šifrnici neće uspjeti i morati će biti odbijena. U tom slučaju stvara se lanac problema, koji počinju od toga da Finini djelatnici moraju samostalno istraživati gdje je nastao problem i koja je trenutna vrijednost šifrnika, do toga da Fina gubi novac, a u konačnici i klijente.

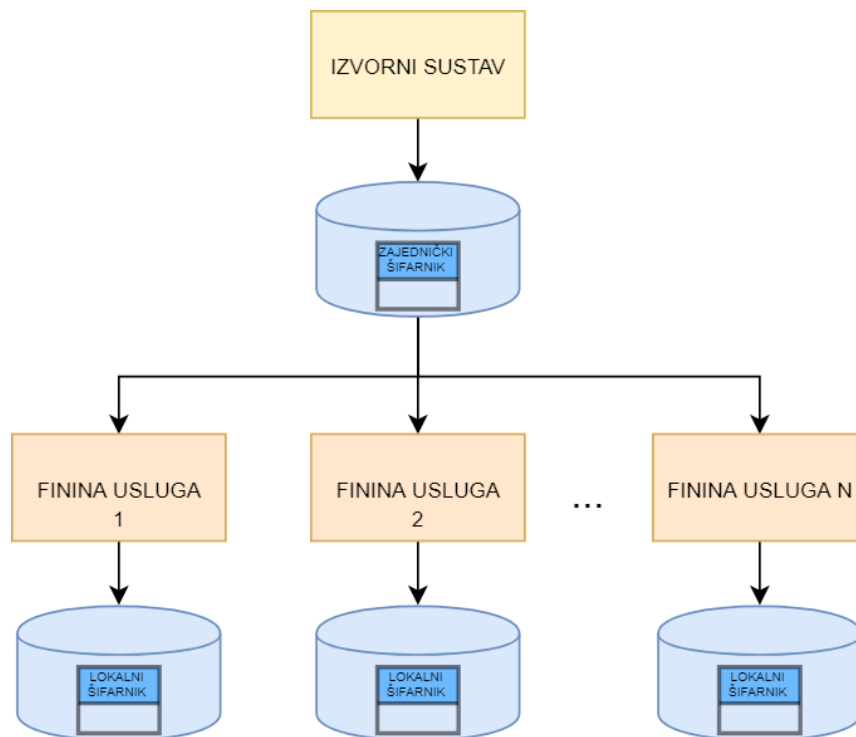
4.1. Osnovna načela MDM-a na razini zajedničkih šifrnika

Unutar Fine postoje dvije vrste šifrnika:

- Zajednički šifrnik – šifrnik koji je zajednički svim Centrima proizvoda čije usluge koriste njegovu kopiju u svrhe svojih poslovnih procesa
- Lokalni šifrnik – kopija globalnog šifrnika unutar određenog Centra proizvoda čije ga usluge koriste u svrhe svojih poslovnih procesa

Po uzoru na zajednički šifrnik stvara se kopija lokalnog šifrnika, pa bi se moglo reći kako su, hijerarhijski gledano, zajednički i lokalni šifrnik u odnosu roditelj-dijete. Ovisno o vrsti, razlikuje se i način održavanja šifrnika koji se također odvija hijerarhijski. Zajednički šifrnici održavaju se pomoću svojih izvornih sustava, dok lokalne šifrnike, po uzoru na zajedničke, održavaju usluge Centra proizvoda koji ih koriste u svrhe svojih poslovnih procesa.

Na slici 4.2 prikazan je hijerarhijski način održavanja šifrnika. Određeni izvorni sustav održava zajednički šifrnik čiju kopiju sve Finine usluge koriste i održavaju kao lokalni šifrnik.



Slika 4.2. Hijerarhijsko održavanje šifrnika

Održavanje zajedničkih šifrnika bitno je kako Finine usluge ne bi ovisile o dostupnosti izvornog sustava, odnosno kako bi Finine usluge kroz zajedničke šifrnike u svakom trenutku imali dostupne podatke koji su se posljednji ažurirali iz izvornog sustava. Zajednički šifrnici održavaju se ovisno o tome podržava li izvorni sustav određene uslužno-orijentirane usluge.

Načini održavanja zajedničkih šifrnika mogu biti:

- Ručno održavanje – izvorni sustav ne podržava mrežne usluge te korisnik koji ima ovlasti administratora ručno ažurira promjene
- Održavanje putem mrežnih usluga – izvorni sustav podržava mrežne usluge:
 - Automatski – ažuriranje kroz određeni period vremena
 - Na zahtjev – ažuriranje slanjem određenih mrežnih zahtjeva

Prethodno spomenut šifrnici *Država* primjer je zajedničkog šifrnika koji se održava ručnim unošenjem u sustav, dok primjer Fininog zajedničkog šifrnika koji se održava putem mrežnih usluga je primjer OIB-a⁹, odnosno tzv. Centralni OIB – COIB. Izvorni sustav je registar OIB-a koji održava tvrtka APIS IT. COIB se automatski, kroz određeni period vremena, spaja na mrežne usluge koje APIS IT kao izvorni sustav ima otvorene prema van, te odande ažurira sve izmjene u odnosu na zadnje spajanje.

COIB sam po sebi izvrsno funkcionira te njegova implementacija unutar Fine uvelike popravlja kvalitetu te "uvodi više reda" među podacima. Primjerice, prilikom ispunjavanja određenih obrazaca gdje se unosi OIB. Ukoliko je unesena vrijednost OIB-a ispravna i nalazi se u zajedničkom šifrniku, na taj način će se i ostali podaci povezani uz uneseni OIB, kao što su npr. ime i prezime za fizičke osobe, naziv tvrtke za pravne osobe, adresa, itd., također "povući" te automatski popuniti u obrazac. Rezultat toga je izbjegavanje mogućih grešaka prilikom unosa te u konačnici pohranjivanje i daljnja obrada točnih i istinitih podataka unutar skladišta. COIB je primjer kako bi trebali i ostali šifrnici funkcionirati.

Lokalni šifrnici održavaju se ovisno o tome podržavaju li Finine usluge koje ih koriste određene uslužno-orijentirane usluge. Načini održavanja lokalnih šifrnika mogu biti:

- Ručno održavanje – Finine usluge ne podržavaju mrežne usluge te ovlašteni djelatnik Centra proizvoda kojem usluga pripada ručno ažurira promjene
- Održavanje putem mrežnih usluga – Finine usluge podržavaju mrežne usluge

Još neki primjeri šifrnika koji se koriste u Fini su *Valuta, Nacionalnost, Jezik, Nacionalna klasifikacija djelatnosti, Obrtni registar, Registar poslovnih jedinica, Razdjel i glava*, itd. Ukupan broj šifrnika koje koriste Finine usluge je iznimno velik, a u budućnosti će ih se pojavljivati još.

⁹ OIB – Osobni identifikacijski broj, predstavlja identifikacijsku oznaku svake fizičke i pravne osobe sa sjedištem u RH

4.2. Specifikacija Središnjeg sustava šifrnika

Svaki šifrnika koji već postoji, kao i onaj koji će se tek pojaviti, specifičan je na svoj način. Pojavljuje se od najjednostavnijih šifrnika koji se sastoje od gotovo jedne tablice, hijerarhijskih šifrnika koji se sastoje od više tablica povezanih u hijerarhiju, pa sve do složenih šifrnika koji sadržavaju složene strukture službenih registara.

Iako svaki šifrnika nosi zasebnu težinu odgovornosti te "probavlja" određene podatke, zajedničko im je to da je svaki od njih podjednako bitan za Finine usluge koje ih koriste za potrebe svojih poslovnih procesa. Pregledniji pristup i brža interakcija sa promjenama u šifrniciima znači bolji rad Fininih usluga, što u konačnici također predstavlja i bolji rad cjelokupnog Fininog poslovnog sustava.

Važnost šifrnika prepoznata je među Fininim zaposlenicima te se iz tog razloga javila potreba za idejom oko implementacije tzv. Središnjeg sustava šifrnika. Takav sustav šifrnika služio bi kao MDM, odnosno kao integrirano mjesto koje pruža jedinstvenu verziju istine nad svim Fininim zajedničkim šifrniciima i održava ih u skladu s pristupom na koji su njihovi izvorni sustavi orijentirani. Ideja Središnjeg sustava šifrnika je da svi zajednički šifrnici, bez obzira jesu li oni jednostavni, hijerarhijski ili složeni te održavaju li se ručno ili putem mrežnih usluga, budu implementirani na jedinstvenom mjestu kojemu će različiti korisnici, ovisno o razini prava, pristupati i pomoću njega održavati svoje lokalne šifrnike. Te koje će u konačnici osigurati izvršenje transakcija i pojednostaviti ETL proces Fininog skladišta podataka.

Središnji sustav šifrnika ponaša se kao matično okruženje koje upravlja kvalitetom matičnih podataka prikupljenih iz različitih šifrnika. Kako bi se različiti šifrnici uspješno integrirali kao konzistentna cjelina, nužno je pravilno identificirati karakteristike pojedinog šifrnika te pomoću definiranih pravila svaki od njih klasificirati i ispravno uklopiti u Središnji sustav šifrnika:

1. Klasifikacija šifrnika prema složenosti:
 - Jednostavni šifrnici
 - Složeni šifrnici
 - Hijerarhijski šifrnici
2. Svaki šifrnika ima vlastiti način održavanja:
 - Ručno održavani šifrnici imaju zajedničku programsku podršku pomoću koje se pojedinom šifrniku pristupa te ga se ručno ažurira
 - Svaki mrežno održavan šifrnika podržava zasebne mrežne usluge specijalizirane za svoj slučaj šifrnika

3. Određivanje prava pristupa:

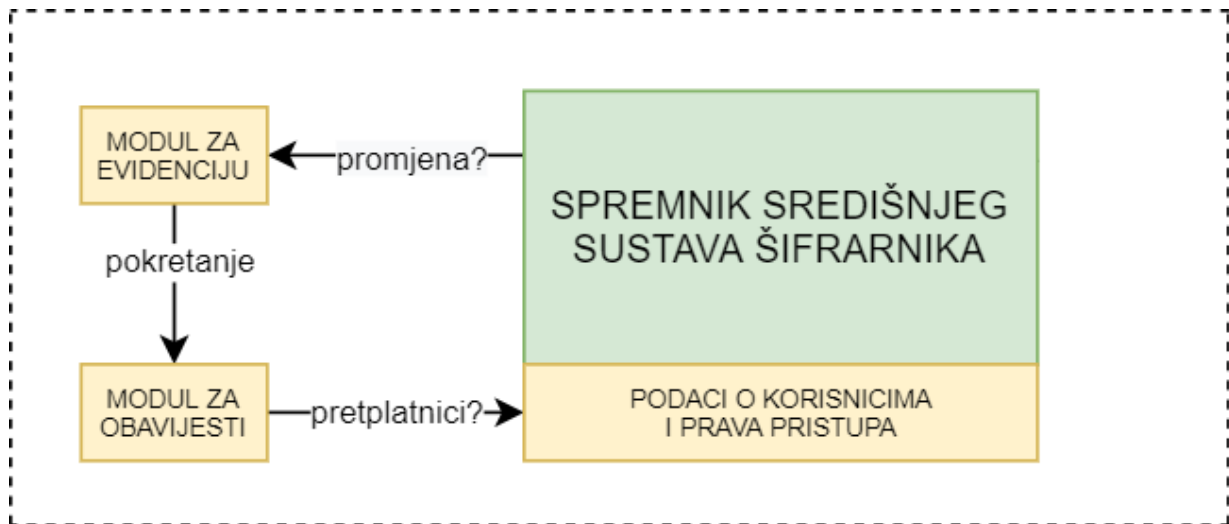
- Tko i do koje razine ima pravo pristupa određenom šifrniku

4.3. Komponente Središnjeg sustava šifrnika

Komponente koje je potrebno implementirati kako bi se Središnji sustav šifrnika ponašao kao jedinstveno mjesto i funkcionirao kao cjelina dijele se na unutarnje i vanjske.

4.3.1.1. Unutarnje komponente

Unutarnje komponente, prikazane slikom 4.3, odgovorne su za glavne funkcionalnosti koje Središnji sustav šifrnika treba pružati te za svu komunikaciju koju uspostavlja prema van. Svaka unutarnja komponenta je pasivna dok ju ne pokrene određena mrežna usluga ili zahtjev, korisnik ručno, neka od preostalih komponenti ili pak automatski brojač vremena (engl. *timer*).



Slika 4.3. Unutarnje komponente Središnjeg sustava šifrnika

Unutarnje komponente Središnjeg sustava šifrnika čine spremnik, sloj odgovoran za podatke o korisnicima i prava pristupa, modul za evidenciju i modul za obavijesti.

4.3.1.2. Spremnik Središnjeg sustava šifrnika

Spremnik je ključna unutarnja komponenta Središnjeg sustava šifrnika koja služi kao mjesto na kojem se integriraju i konsolidiraju svi zajednički šifrnici. Prilikom inicijalne integracije potrebno je definirati početno stanje koje će se napraviti unosom šifrnika iz izvornih sustava. Nakon čega se spremnik kroz procese sinkronizacije i distribucije održava. Svaka novonastala promjena identificirana na izvoru, sinkronizira se unutar registra te dalje distribuira svim Fininim uslugama pretplaćenim na šifrnika iz registra u kojem su se dogodile promjene. Sinkronizacija i distribucija se odvijaju ručno ili putem uslužno-orijentiranog pristupa, više o tome opisano je u odlomku 4.3.

4.3.1.3. Modul za evidenciju

Svaki događaj unutar Središnjeg sustava šifrnika, bilježi se pomoću modula za evidenciju prema sljedećim pravilima:

- Prilikom promjene šifrnika unutar registra, neophodno je zabilježiti u kojem šifrniku se promjena dogodila
- Prilikom promjene šifrnika unutar registra, poželjno je zabilježiti što se točno promijenilo
- Svaki događaj povezan sa određenom mrežnom uslugom s kojom Središnji sustav šifrnika komunicira neophodno je zabilježiti, posebno ukoliko su se dogodile potencijalne greške u radu ili je usluga nedostupna
- Svaki događaj izazvan ručnom radnjom administratora ili nekog drugog korisnika unutar Središnjeg sustava šifrnika neophodno je zabilježiti
- Ukoliko određeni događaj zahtjeva slanje obavijesti, inicira se modul za obavijesti

4.3.1.4. Modul za obavijesti

Ukoliko modul za evidenciju zabilježi događaj koji inicira modul za obavijesti, tada modul za obavijesti provjerava u sloju za korisničke podatke i prava pristupa koji su korisnici pretplaćeni na šifrnici unutar registra čiji je događaj inicirao obavijest. Potom svi pretplaćeni korisnici putem elektroničke pošte dobivaju obavijest o događaju koji se zbija unutar Središnjeg sustava šifrnika. Pretplaćeni korisnici su svi oni koji lokalne kopije svojih šifrnika održavaju putem spremnika Središnjeg sustava šifrnika, a slanje obavijesti ovisi o konfiguraciji šifrnika.

Događaji prilikom kojih se šalje obavijest:

- Ukoliko se dogodila promjena unutar šifrnika za koji je u konfiguraciji postavljeno da se šalje obavijest
- Mrežna usluga pomoću koje se održava šifrnici trajno je ispala iz rada ili je zamijenjena drugom uslugom
- Administrator je ažurirao prava pristupa koja se odnose na pretplaćenog korisnika

4.3.1.5. Sloj za podatke o korisnicima i prava pristupa

Kako bi se povelu računa o tome tko i do koje razine može pristupiti pojedinom šifrniku, nužno je unutar Središnjeg sustava šifrnika održavati podatke o korisnicima. Prije samog pristupa spremniku, vrši se provjera na temelju sloja odgovornog za korisničke podatke i prava pristupa. Ukoliko je provjera o pravima pristupa uspješna, tek tada se može do određene razine pristupiti

odobrenom sadržaju spremnika. Podatke o korisnicima i prava pristupa administrator sustava unosi i evidentira.

Razina prava pristupa korisnika se razlikuju prema sljedećem:

- Korisnik je ovlašten samo korištenje određenih šifrnika iz spremnika
- Korisnik je ovlašten i za korištenje i za održavanje određenih šifrnika iz spremnika
- Administrator ima sve ovlasti i najveću razinu prava pristupa
- Korisnik nema prava pristupa i ne može pristupiti Središnjem sustavu šifrnika

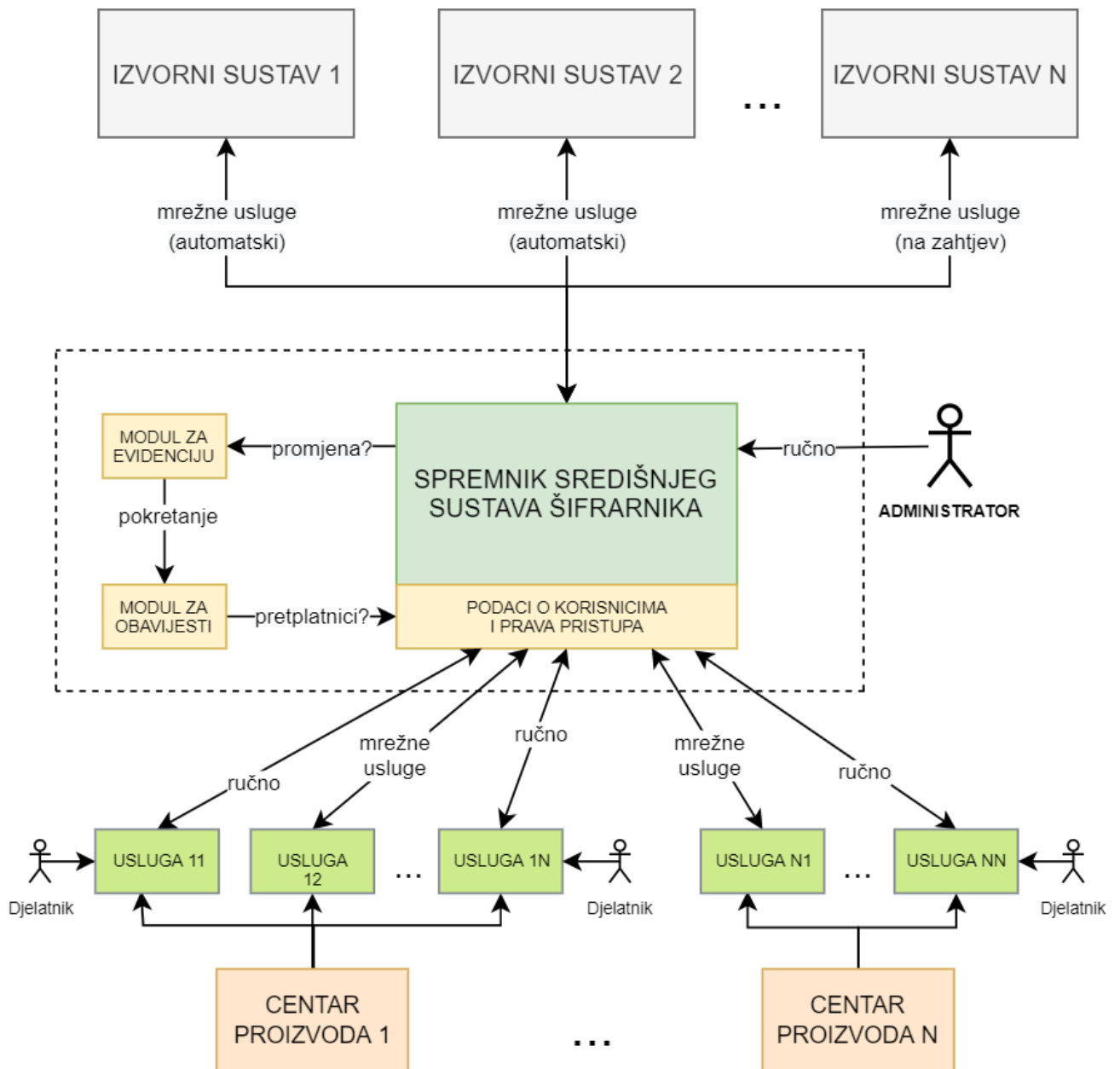
4.3.2. Vanjske komponente

Kako bi cijeli sustav bio upotpunjen te kako bi unutarnje komponente Središnjeg sustava šifrnika uspješno uspostavljale komunikaciju prema van, u toj priči bitnu ulogu imaju vanjske komponente sustava.

Vanjske komponente razlikuju se ovisno o razini i načinu održavanja šifrnika:

- Izvorni sustavi:
 - Izvorni sustavi šifrnika
 - Djelatnici s administratorskim pravima
- Odredišni sustavi:
 - Samostalne usluge Centara proizvoda
 - Djelatnici Centara proizvoda

Kompletan dizajn sa svim unutarnjim i vanjskim komponentama te načinom na koji komponente međusobno komuniciraju prikazan je slikom 4.4. Šifrnici iz spremnika Središnjeg sustava šifrnika sinkronizirani su sa svojim izvornim sustavima ovisno o tome podržava li izvorni sustavi određene mrežne usluge. Šifrnici čiji izvorni sustavi ne podržavaju mrežne usluge, njih ručno održava administrator sustava. Samostalne usluge i djelatnici Centara proizvoda pristupaju spremniku Središnjeg sustava šifrnika i koriste njegove podatke ukoliko im je, na osnovu podataka o korisnicima koji se održavaju u Središnjem sustavu šifrnika, pristup odobren.



Slika 4.4. Model Središnjeg sustava šifrnika

4.4. Osnovne funkcionalnosti Središnjeg sustava šifrnika

U prethodnom odlomku prikazane su unutarnje i vanjske komponente Središnjeg sustava šifrnika koje su ključne u implementaciji samog sustava. Kako bi te komponente imale svrhu te kako bi sustav u cjelini bio uporabljiv, neophodno je osvrnuti se i na njihove funkcionalnosti. Dvije osnovne funkcionalnosti koje sustav mora podržavati su održavanje Središnjeg sustava šifrnika i korištenje podataka iz Središnjeg sustava šifrnika. To su dva odvojena procesa koja se u sustavu odvijaju asinkrono, odnosno potpuno neovisno jedan o drugome.

4.4.1. Održavanje Središnjeg sustava šifrnika

Novonastale promjene nad određenim šifrnikom koje je potrebno održavati u spremniku Središnjeg sustava šifrnika ažuriraju se ovisno o tome podržava li izvorni sustav određene mrežne usluge otvorene prema Središnjem sustavu šifrnika.

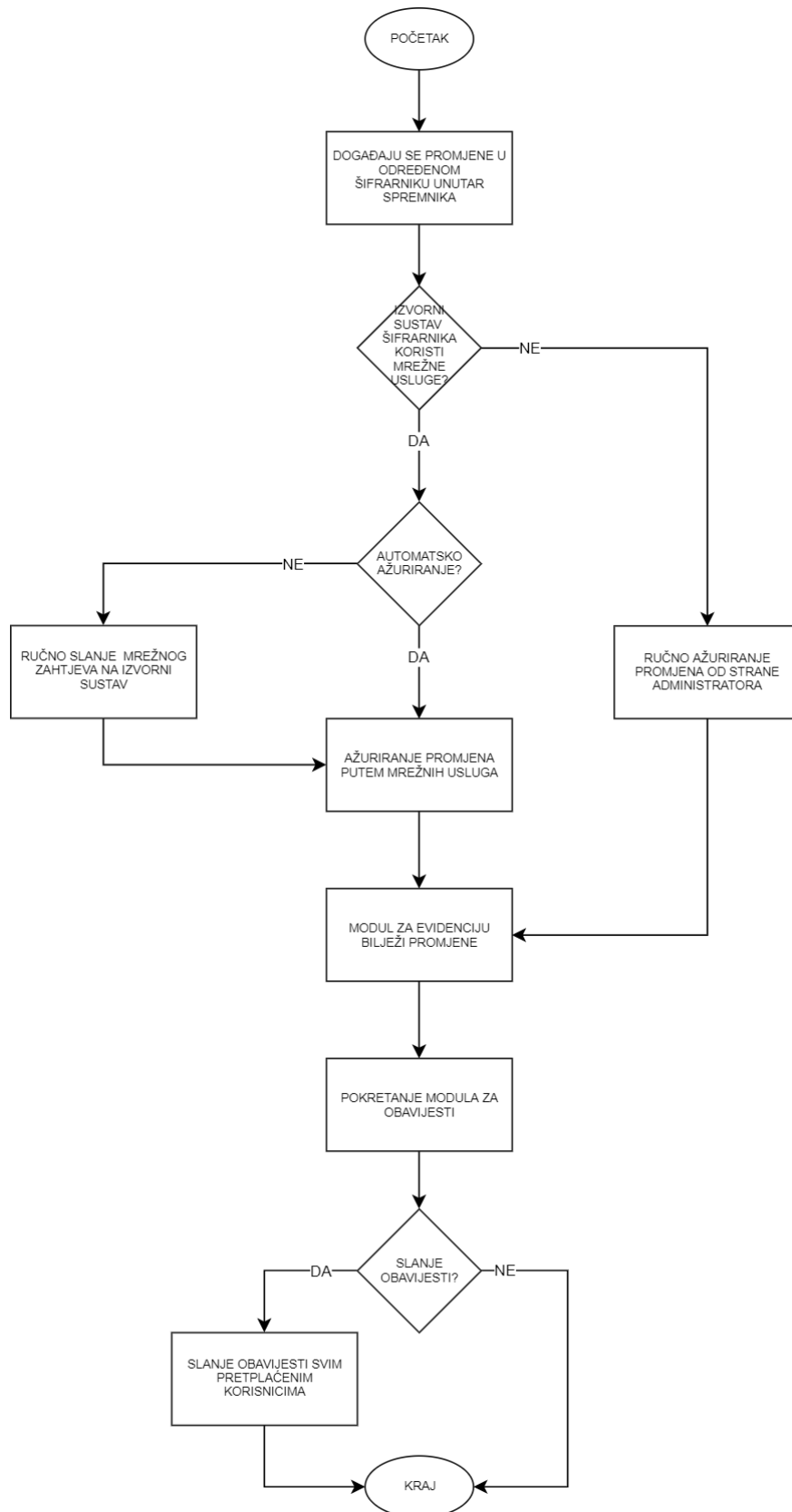
Pomoću mrežnih usluga šifrnici se mogu održavati:

- Automatski, na način da se kroz osvježavanja u određenim vremenskim intervalima u spremnik ažuriraju najsvježiji podaci iz izvornih sustava
- Slanjem određenih mrežnih zahtjeva na izvorne sustave koji, ukoliko su spremni, šalju odgovore na Središnji sustav šifrnika

Šifrnici iz spremnika Središnjeg sustava šifrnika čiji izvorni sustavi ne podržavaju određene mrežne usluge održavaju se ručno od strane administratora.

Modul za evidenciju bilježi svaku promjenu koja se dogodi u spremniku Središnjeg sustava šifrnika te potom pokreće modul za obavijesti koji, ovisno o konfiguraciji šifrnika, obavještava sve pretplaćene korisnike. U konfiguraciji šifrnika može se postaviti šalju li se obavijesti ili ne, no neovisno o tome, svaka promjena i dalje mora biti zabilježena te vidljiva unutar modula za evidenciju.

Proces koji se odnosi na održavanje Središnjeg sustava šifrnika prikazan je kroz dijagram toka na slici 4.5.

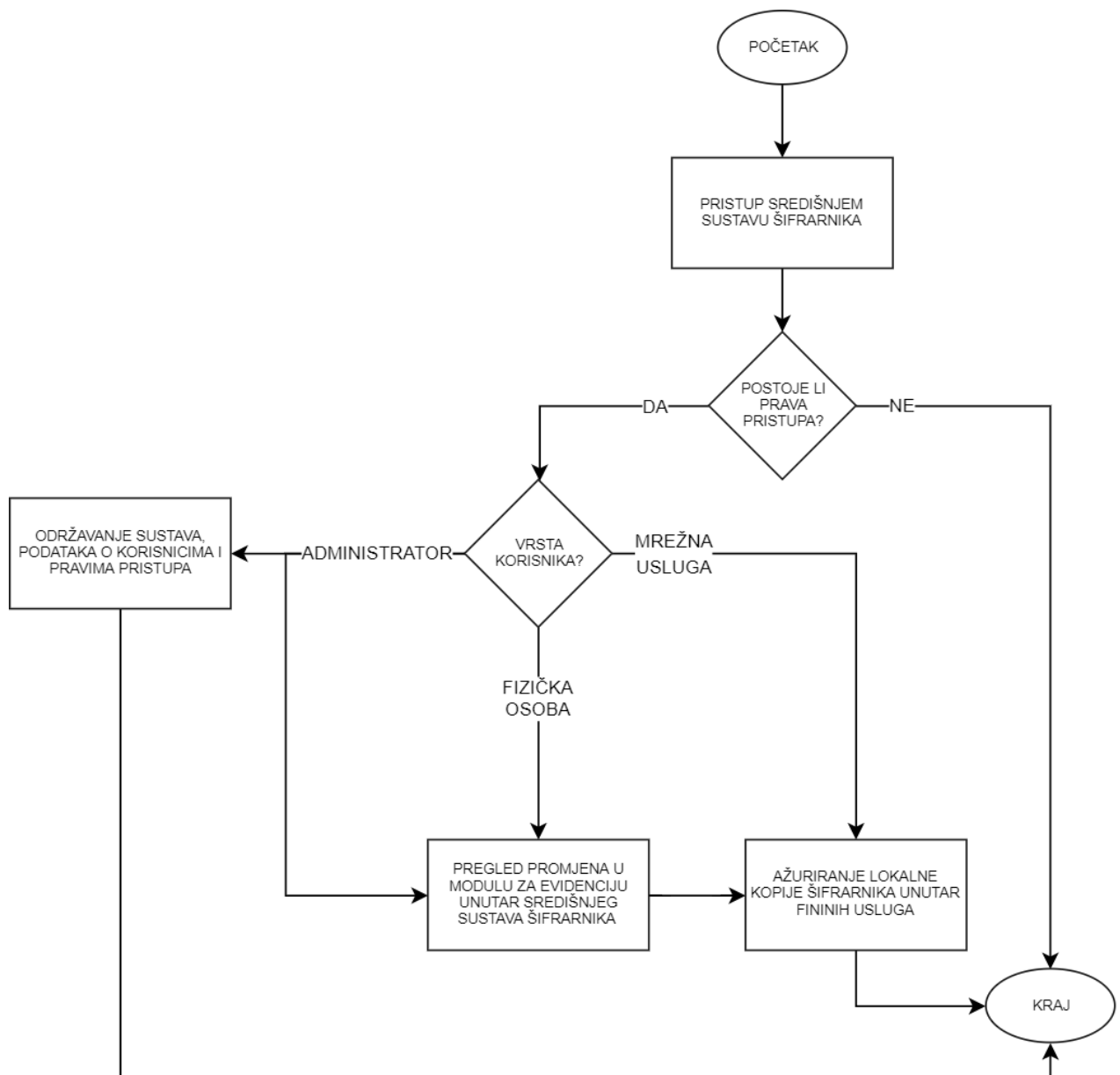


Slika 4.5. Proces održavanja Središnjeg sustava šifrnika

4.4.2. Korištenje podataka iz Središnjeg sustava šifrnika

Središnji sustav šifrnika posjeduje podatke o korisnicima i pravima pristupa. Nakon što određeni korisnik pokuša pristupiti Središnjem sustavu šifrnika, vrši se provjera prava pristupa. Ukoliko je pristup dozvoljen, korisnik pristupa Središnjem sustavu šifrnika te, ovisno o razini prava, koristi podatke iz sustava i u skladu s njima ažurira svoje lokalne šifrnike.

Proces koji se odnosi na korištenje podataka iz Središnjeg sustava šifrnika prikazan je kroz dijagram toka na slici 4.6.



Slika 4.6. Proces korištenja podataka iz Središnjeg sustava šifrnika

4.5. Odgovornosti sudionika unutar središnjeg sustava šifrnika

Postoje tri vrste korisnika Središnjeg sustava šifrnika:

- Administrator sustava
- Fizička osoba
- Mrežna usluga (engl. *web service*)

4.5.1. Administrator sustava

Administrator sustava ima najveća prava pristupa te uz to i najveću razinu odgovornosti prema Središnjem sustavu šifrnika. On održava šifrnike iz spremnika čiji izvorni sustavi ne koriste mrežne usluge, na način da ručno unosi promjene u spremnik Središnjeg sustava šifrnika. Kako bi se izbjegao problem točnosti podataka, administrator je dužan izvršiti dodatnu provjeru je li vrijednost ručno održavanih šifrnika u skladu sa stvarnom vrijednosti. Administrator ima ovlasti nad kompletnim sustavom i odgovoran je za održavanje podataka o korisnicima kojima može mijenjati prava pristupa. Ima uvid u modul za evidenciju te može primati obavijesti, no za razliku od ostalih korisnika, administrator sustava je također i odgovoran za ispravan rad modula za evidenciju i modula za obavijesti.

4.5.2. Fizička osoba

Pod fizičkim osobama smatraju se djelatnici Fininih Centara proizvoda koji imaju potrebu koristiti Središnji sustav šifrnika. Finini djelatnici Središnjem sustavu šifrnika pristupaju ovisno o pravima pristupa koja su definirana za Centar proizvoda u kojemu su zaposleni. Odnosno, ako određeni Centar proizvoda nema pravo pristupa na određeni šifrnika u spremniku Središnjeg sustava šifrnika, također ga ne može imati niti jedan djelatnik tog Centra proizvoda, dok na razini Centra proizvoda djelatnici mogu biti ovlašteni za pristup svim ili samo nekim lokalnim šifrnicima.

Djelatnik kao fizička osoba može biti istovremeno i administrator sustava. To je poseban slučaj u kojem su jednom korisniku dodijeljene dvije različite uloge.

4.5.3. Mrežna usluga

Mrežna usluga u kontekstu korisnika odnosi se na Finine usluge koje posjeduju određene mrežne usluge pomoću kojih se spajaju na spremnik Središnjeg sustava šifrnika i na takav način održavaju svoje lokalne šifrnike. Ovakva vrsta korisnika Središnjeg sustava šifrnika je poželjna te je zamišljeno da u budućnosti sve Finine usluge rade po takvom principu.

4.6. Analiza dizajna Središnjeg sustava šifrnika

S obzirom na važnost šifrnika unutar Fininih poslovnih procesa, kroz posljednjih nekoliko godina često se spominjala tema koja se odnosi na potrebu implementacije MDM sustava na razini zajedničkih šifrnika. U tom periodu neophodno je bilo odrediti Finin stupanj zrelosti te potom izgraditi MDM strategiju u skladu sa politikom organizacije, načelima i standardima kojih bi se MDM program trebao pridržavati.

Kao sljedeći korak u tom procesu, proizlazi model Središnjeg sustava šifrnika sa svim njegovim bitnim karakteristikama, a postignut je kao produkt praktičnog dijela ovog rada. Model je dizajniran po uzoru na koegzistencijski stil implementacije MDM-a. Podaci se čiste i usklađuju unutar izvornih sustava, nakon čega se integriraju i konsolidiraju u spremnik Središnjeg sustava šifrnika. Integrirani podaci se unutar spremnika održavaju sinkronizacijom sa izvornim sustavima te potom distribuiraju svim Fininim uslugama koje koriste lokalnu kopiju šifrnika. Svaka novonastala promjena unutar šifrnika bilježi se te ovisno o konfiguracijskim postavkama šalje se obavijest o promjeni. Centri proizvoda "povlače" najsvježije podatke unutar svojih lokalnih kopija šifrnika ovisno o tome podržavaju li određene mrežne usluge, u suprotnom se promjene unutar lokalnih kopija šifrnika unose ručno. Središnji sustav šifrnika također posjeduje administratora koji ima ovlasti upravitelja cjelokupnog sustava.

U budućnosti je planirano da sve Finine usluge podržavaju uslužno-orijentirani pristup. Time će se ujedno postići i ciljno stanje modela u kojem će, one usluge koje se danas održavaju ručno, također biti automatizirane.

Na temelju modela, arhitekti sustava (engl. *system architect*) određuju alate i tehnologije koji će se koristiti u procesu izgradnje sustava. Nakon toga potrebno je raspisati specifikaciju mogućih slučajevi ponašanja sustava (engl. *use case*) kroz funkcionalne cjeline, što čine analitičari. Dizajneri (engl. *database designer*) osmišljavaju detaljan dizajn baze podataka, dok se o dizajnu određenih mrežnih usluga brinu arhitekti sustava. Kada su svi prethodni koraci ispunjeni i poprilično jasni, slijedi rješavanje programskog dijela implementacije, koji je naknadno potrebno ispitati.

Kroz inicijalno punjenje Središnjeg sustava šifrnika će se, po uzoru na šifrnike iz izvornih sustava, unijeti svi šifrnici koje Fina koristi te po potrebi dopuniti određenim povijesnim podacima koji su korišteni su kroz Finine usluge. XML (engl. *EXtensible Markup Language*) je format za razmjenu podataka pomoću kojeg će se podaci unutar MDM okruženja prenositi.

Takav model bi trebao biti na razini učinkovitosti pri kojoj će služiti kao prevencija zastoja u radu, odnosno kao sigurnost da će se sve Finine transakcije pravilno izvršiti. Što automatski poboljšava Finine poslovne procese, konkurentnost i odnose s klijentima te smanjuje rizik od financijskog gubitka. Dodatno, integracijom šifrnika kroz MDM sustav, trebao bi se pojednostaviti ETL proces prilikom skladištenja podataka, što će omogućiti brže upite na skladište podataka, učinkovitiju analizu, izradu smislenijih izvještaja, mogućnost jednostavnijeg upravljanja rizikom i donošenja važnih strateških odluka.

4.7. Vrednovanje dizajna Središnjeg sustava šifrnika

Da bi se utvrdila očekivana razina učinkovitosti Središnjeg sustava šifrnika, izvršeno je vrednovanje dizajniranog modela. Kako je model još uvijek u fazi planiranja te ne postoje nikakve numeričke vrijednosti pomoću kojih je moguće izračunati točnu vrijednost, vrednovanje modela izvršeno je na temelju osobne procjene. Tablica 4.1 prikazuje rezultate vrednovanja izvršenih usporedbom procijenjenog trenutnog i stanja nakon implementacije dizajniranog modela Središnjeg sustava šifrnika, a temelji se na ključnim dimenzijama kvalitete podataka opisanim u [1] i [2].

Između zajedničkih Fininih šifrnika trenutno nema integriteta, već bi se on trebao ostvariti implementacijom Središnjeg sustava šifrnika. Iz tog razloga, vrijednosti između trenutnog i očekivanog stanja u tablici 4.1 maksimalno osciliraju. Problem konzistentnosti jedan je od najvećih razloga zbog kojih se javlja potreba za MDM-om na razini zajedničkih šifrnika. Čest je slučaj kako se vrijednosti istih šifrnika razlikuju kroz lokalne kopije šifrnika, što dovodi do blokiranja Fininih transakcija koje koriste zastarjele vrijednosti unutar lokalnih šifrnika. Središnji sustav šifrnika rješava taj problem i omogućuje veliku razinu konzistentnosti, no kako se određene Finine usluge još uvijek održavaju ručno, može se dogoditi da se ljudskom pogreškom određeni lokalni šifrnici ne ažuriraju na vrijeme. Taj problem će se riješiti kada se sve Finine usluge budu održavale putem mrežnih usluga.

Pravovremenost predstavlja maksimalno vrijeme koje je potrebno da se postigne konzistentnost između svih sustava. U trenutnom stanju pravovremenost je iznimno loša, jer se određeni lokalni šifrnici ažuriraju tek kada nastane problem, odnosno kada dođe do pogreške i otkrije se da su određene vrijednosti unutar zajedničkih šifrnika promijenjene. Pravovremenost bi uvelike poboljšala implementacija Središnjeg sustava šifrnika. No, kako je potrebno vrijeme između održavanja šifrnika unutar izvornih sustava, pohranjivanja u spremnik Središnjeg sustava šifrnika i dijeljenje Fininim Centrima proizvoda, tu se također mogu pojaviti manja kašnjenja.

Ažurnost odgovara stupnju "svježine" matičnih podataka. Finini zajednički šifrnici, čiji izvorni sustavi podržavaju mrežne usluge, učestalo se ažuriraju u skladu s najsvježijim podacima. Iznimno, ukoliko je u određenom trenutku izvorni sustav nedostupan, zajednički šifrnici će podržavati stariju inačicu podataka, što bi moglo umanjiti stupanj "svježine" šifrnika. Trenutno se veći problem ažurnosti javlja između zajedničkih i lokalnih šifrnika, no implementacijom modela Središnjeg sustava šifrnika to bi se trebalo poboljšati.

Točnost podataka mjeri se u odnosu na nepravilno unesene vrijednosti podataka. Problem točnosti podataka najčešće je uzrokovan ljudskom pogreškom unutar sustava koji se održavaju ručno. Pomoću dodatne provjere administratora Središnjeg sustava šifrnika, povećao bi se nadzor nad ručno unesenim vrijednostima određenih šifrnika. Latentnost se odnosi na vrijeme između slanja zahtjeva i odgovora prilikom sinkronizacije Središnjeg sustava šifrnika s izvornim sustavom, koja ovisi o mrežnim uslugama i dostupnosti izvornog sustava. Potpunost je učestalost pojavljivanja praznih vrijednosti podataka, što ovdje nije slučaj, jer unutar šifrnika svaka jedinstvena šifra ima svoj naziv.

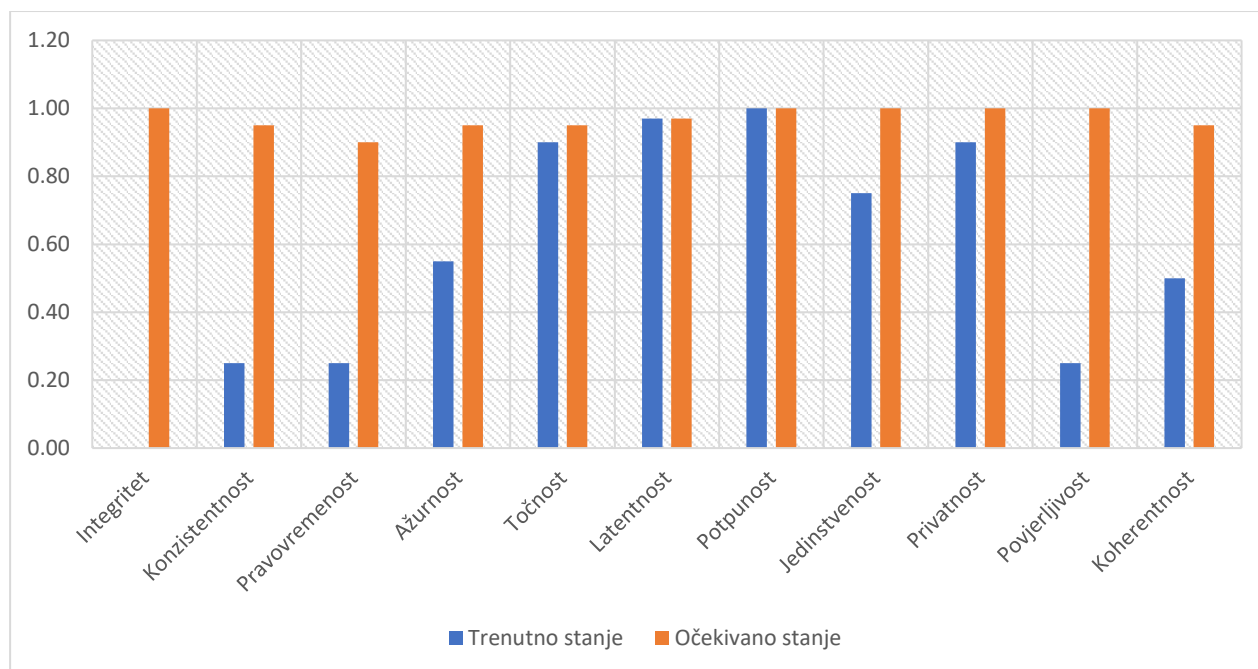
Jedinstvenost definira da se vrijednost šifrnika pojavljuje samo jednom unutar istog skupa podataka. Svaki šifrnici kao zaseban skup podataka već je relativno jedinstven, no implementacijom Središnjeg sustava šifrnika, planira se postići apsolutna jedinstvenost između svih zajedničkih šifrnika. Privatnost se očituje kod šifrnika koji zahtijevaju ograničenja upotrebe ili pristupa. Unutar Središnjeg sustava šifrnika, to se poboljšava putem sloja koji evidentira podatke o korisnicima i razine prava pristupa.

Povjerljivost u podatke procjenjuje se na osnovu mogućnosti praćenja promjena vrijednosti. Model Središnjeg sustava šifrnika definira modul za evidenciju, čija zadaća je bilježiti promjene i pratiti njihovu povijest. Dok trenutno, povijest i mjesto promjene djelatnik Centra proizvoda može jedino pretpostaviti na osnovu pogreške koja se dogodila. Koherentnost opisuje povezanost sustava. Trenutno je svaki Finin zajednički šifrnici dobro povezan sa svojim izvornim sustavom. Objedinjavanjem svih Fininih šifrnika u središnji sustav, omogućit će se dobra povezanost i sa lokalnim šifrniciima koji Finine usluge koriste.

Tablica 4.1. Procijenjene vrijednosti trenutnog i očekivanog stanja

rbr.	Dimenzija	Trenutno stanje	Očekivano stanje
1	Integritet	0.00	1.00
2	Konzistentnost	0.25	0.95
3	Pravovremenost	0.25	0.90
4	Ažurnost	0.55	0.95
5	Točnost	0.90	0.95
6	Latentnost	0.97	0.97
7	Potpunost	1.00	1.00
8	Jedinstvenost	0.75	1.00
9	Privatnost	0.90	1.00
10	Povjerljivost	0.25	1.00
11	Koherentnost	0.50	0.95

Na temelju podataka iz tablice 4.1 dobiven je graf koji je prikazan slikom 4.7. Procijenjene vrijednosti očekivanog stanja zadovoljavaju sve dimenzije kvalitete podataka u vrijednostima većim od 0.9. Najveći utjecaj modela Središnjeg sustava šifrnika, u odnosu na trenutno stanje, očituje se kod integriteta, konzistentnosti, pravovremenosti, ažurnosti, povjerljivosti i koherentnosti, koje i jesu ključne dimenzije kako bi se ispunila razina učinkovitosti pri kojoj će model služiti kao prevencija zastoja u radu. Vidljivo je i poboljšanje točnosti koja je bitna prilikom daljnjih analiza i kreiranja poslovnih izvještaja unutar skladišta podataka. Kod latentnosti i potpunosti promjene nisu uočljive.



Slika 4.7. Grafički prikaz procijenjenih vrijednosti trenutnog i očekivanog stanja

Ovim načinom vrednovanja, pokušava se procijeniti razina učinkovitosti modela Središnjeg sustava šifrnika, a on može poslužiti i za procjenu očekivane kvalitete podataka. Nakon implementacije Središnjeg sustava šifrnika obaviti će se mjerenje dimenzija kvalitete podataka na stvarnim vrijednostima, pomoću kojih će se ustanoviti je li kvaliteta podataka Središnjeg sustava šifrnika u skladu s očekivanom.

5. ZAKLJUČAK

Svakoj velikoj organizaciji koja je predvođena velikom količinom podataka i usmjerena prema digitalnoj transformaciji, potreban je fleksibilan i sveobuhvatan pristup koji započinje s podacima, a završava poslovnim rezultatima. U tu svrhu ključno je uvođenje postupka upravljanja matičnim podacima (MDM).

Upravljanje matičnim podacima je vrlo zahtjevan i mukotrpan postupak koji uključuje od vrhovne uprave, preko arhitekata i upravitelja podataka, pa sve do raznih sustava, procesa, tehnologija i alata. Nakon što se u vrhu razriješe političko-poslovne discipline upravljanja matičnim podacima, tada je potrebno uključiti tehnološke discipline. U odnosu na zahtjeve, mogućnosti, funkcionalnosti i već postojeće sustave unutar organizacije odabire se odgovarajući stil implementacije MDM sustava, pomoću kojeg se definira kako se podaci čiste, ujedanjuju i dijele između MDM sustava i različitih vanjskih sustava.

Ovim radom dizajniran je model Središnjeg sustava šifrnika čiji cilj je kontrolom kvalitete i integracijom zajedničkih šifrnika osigurati pravilno izvođenje Fininih poslovnih procesa. Model je grubom procjenom, na temelju dimenzija kvalitete podataka, vrednovan te se time pokazao kao učinkovit. Vrednovanjem je također postavljena očekivana kvaliteta podataka, pomoću koje će se nakon implementacije Središnjeg sustava šifrnika utvrditi je li njegova kvaliteta podataka u skladu sa očekivanom kvalitetom. Planirane nadogradnje, kojima će se ujedno povećati učinkovitost modela, omogućit će se uvođenjem uslužno-orijentiranog pristupa unutar svih Fininih usluga.

LITERATURA

- [1] M. Mosley, M. Brackett, S. Earley, D. Henderson, *DAMA, The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOOK Guide)*, 1st Edition, Technics Publications, New Jersey, 2009.
- [2] D. Loshin, *Master Data Management*, Morgan Kaufmann, USA, 2009.
- [3] Profisee, *Master Data Management – What, why, how & who*, dostupno na: <https://profisee.com/master-data-management-what-why-how-who/> (02.04.2019.)
- [4] R. Vilminko-Heikkinen, S. Pekkola, *Changes in Roles, Responsibilities and Ownership in Organizing Master Data Management*, 47, 76-87, *International Journal of Information Management*, Tampere, 2019.
- [5] Search Data Management, *Master Data Management (MDM)*, dostupno na: <https://searchdatamanagement.techtarget.com/definition/master-data-management/> (02.04.2019.)
- [6] IBM, *IBM InfoSphere Master Data Management*, dostupno na: <https://www.ibm.com/us-en/marketplace/ibm-infosphere-master-data-management/> (31.03.2019.)
- [7] Oracle, *Oracle Product Hub*, dostupno na: <https://www.oracle.com/applications/master-data-management/product-hub/> (31.03.2019.)
- [8] Talend, *Master Data Management*, dostupno na: <https://www.talend.com/products/mdm/> (31.03.2019.)
- [9] TIBCO Orchestra Networks, *Master Data Management*, dostupno na: <https://www.orchestranetworks.com/mdm/> (31.03.2019.)
- [10] Profisee, *Get Started with MDM*, dostupno na: <https://profisee.com/i-want-to/get-started-with-mdm/> (31.03.2019.)
- [11] Informatica, *Master Data Management*, dostupno na: <https://www.informatica.com/products/master-data-management/> (31.03.2019.)
- [12] F. Haneem, R. Ali, N. Kama, S. Basri, *Resolving Data Duplication Inaccuracy and Inconsistency Issues Using Master Data Management*, *International Conference on*

- Research and Innovation in Information Systems (ICRIIS), IEEE, Langkawi, Malaysia, 2017.
- [13] S. Suram, R. Muppala, *Master Data Management – CDI*, TENCON IEEE Region 10 Conference, IEEE, Hyderabad, India, 2008.
- [14] Towards Data Science, *Understanding Various MDM Implementation Styles*, dostupno na: <https://towardsdatascience.com/understanding-various-mdm-implementation-styles-5b4c8fcbbecf/> (04.08.2019.)
- [15] Innovit, *MDM Implementation Styles: An Overview*, dostupno na: <https://www.innovit.com/2018/08/01/mdm-implementation-styles/> (04.08.2019.)
- [16] FINA, *Financijska agencija*, dostupno na: <https://www.fina.hr/> (10.08.2019.)
- [17] W.H.Inmon, *Building the Data Warehouse*, 3rd Edition, John Wiley, New York, 2002.
- [18] Computer weekly, *Inmon or Kimball: Which Approach is Suitable for your Data Warehouse*, dostupno na: <https://www.computerweekly.com/tip/Inmon-or-Kimball-Which-approach-is-suitable-for-your-data-warehouse/> (16.08.2019.)
- [19] T. Jörg, S. Deßloch, *Towards Generating ETL Processes for Incremental Loading*, Proceedings of the 2008 international symposium on Database engineering & applications (IDEAS), ACM, Coimbra, Portugal, 2008.
- [20] G. Martinović, A. Baumgartner, *Računarstvo usluga i analiza podataka*, Sveučilište J.J. Strossmayera u Osijeku, Fakultet elektrotehnike računarstva i informacijskih tehnologija, Osijek, 2016.
- [21] S. Rizzi, *Conceptual Modeling Solutions for the Data Warehouse*, DEIS-University of Bologna, Italy, 2008.
- [22] I. Mekterović, Lj. Brkić, *Skladišta podataka i poslovna inteligencija*, Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva, Zagreb, 2017.
- [23] Industry week, *Integrating MDM into the Data Warehouse*, dostupno na: <https://www.industryweek.com/information-technology/integrating-mdm-data-warehouse/> (30.08.2019.)

SAŽETAK

Postupak upravljanja matičnim podacima (MDM) služi za jedinstveni prikaz različito predstavljenih podataka. Svaka velika organizacija koja iz različitih sustava prima velike količine podataka trebala bi shvaćati njegovu važnost. MDM je vrlo zahtjevan postupak koji, osim tehnoloških, uključuje i poslovno-političke discipline. Prilikom implementacije MDM-a ključna su tri procesa: kvaliteta podataka, integracija podataka i održavanje podataka. Bolja kvaliteta podataka postiže se pravilnom integracijom podataka. Za integraciju podataka može se primijeniti ETL proces, koji je karakterističan prilikom pohrane podataka unutar skladišta podataka. Iz tog razloga skladište podataka može poslužiti kao dobar temelj za implementaciju MDM sustava. Ono što MDM čini jedinstvenim je održavanje podataka koje se odvija sinkronizacijom sa izvornim sustavima te daljnjom distribucijom odredišnim sustavima. Fina je primjer jedne velike organizacije koja je prepoznala važnost upravljanja matičnim podacima te potencijal svoga skladišta podataka iskoristila za implementaciju MDM-a na razini podataka važnih za vlastito poslovanje. Kako bi svi sustavi u Fini imali ažurne podatke o šifranicima, javila se potreba i za implementacijom MDM-a na razini zajedničkih šifrnika. Mjesto na kojemu će se integrirati i održavati svi Finini šifrnici definirano je kao Središnji sustav šifrnika, a razlikuje se prema unutarnjim i vanjskim komponentama. Održavanje Središnjeg sustava šifrnika i korištenje njegovih podataka dva su odvojena procesa koja se unutar sustava odvijaju asinkrono. Postoji nekoliko različitih korisnika Središnjeg sustava šifrnika, od kojih administrator ima najveću razinu ovlasti, dok ovlasti samostalnih usluga i djelatnika ovise o Fininom Centru proizvoda kojem pripadaju.

Ključne riječi: integracija, konsolidacija, kvaliteta podataka, skladište podataka, šifrnici, upravljanje matičnim podacima.

ABSTRACT

MDM method implementation in financial agency

Master data management (MDM) is used to uniquely view differently presented data. Any large organization that receives big amounts of data from different systems should understand its importance. MDM is a very demanding process which involving technological and business-political disciplines. These three areas are crucial when implementing MDM: data quality, data integration, and data maintenance. Better data quality is achieved by proper data integration. An ETL process, which is characteristic for storing data within a data warehouse, can be used for data integration. For that reason, the data warehouse can serve as a good foundation for deploying MDM systems. What makes MDM unique is the maintenance of data that is synchronized with the source systems and further distributed to the destination systems. Fina is an example of a large organization that recognized the importance of master data management and used the potential of its data warehouse to implement MDM at the level of data relevant to its own business. Also, for all systems in Fina to have up-to-date codebooks data, there is needed to implement MDM at the level of common codebooks. The location where all of Fina's codebooks will be integrated and maintained is defined as the Central System of Codebooks and differs according to internal and external components. Maintaining the Central System of Codebooks and using its data are two separate processes that occur asynchronously within the system. There are several different users of the Central System of Codebooks, of which administrator has the highest level of authority, while the powers of Fina services and employees depend on the Center of the product to which they belong.

Key words: integration, consolidation, data quality, data warehouse, codebook, master data management.

ŽIVOTOPIS

Matea Šimunović rođena je 11. lipnja 1994. godine u Osijeku. Osnovnu školu upisuje i završava u Tenji, prigradskom naselju kod Osijeka u kojem ujedno i odrasta. Školovanje nastavlja u II. (jezičnoj) gimnaziji Osijek. A svoje daljnje obrazovanje nastavlja upisavši Fakultet elektrotehnike, računarstva i IT-a u Osijeku. Stečena znanja na fakultetu upotpunjuje praktičnim znanjima pohađajući neobavezne prakse unutar nekoliko programerskih tvrtki te se u tom periodu također počinje intenzivnije baviti područjem Internet programiranja. Od 2017. godine članica je i dopredsjednica osječkog IEEE studentskog ogranka, gdje aktivnim volontiranjem sudjeluje u mnogim projektima čime ujedno usavršava vlastite organizacijske i komunikacijske vještine, a također ima priliku posjetiti nekoliko velikih konferencija na svjetskoj i regionalnoj razini. 2018. godine postaje Finin stipendist i budući zaposlenik. Kako posljednjih nekoliko godina raste popularnost iskorištavanja podataka kao moćnog resursa, intrigira se za to područje te kao temu svog diplomskog rada odabire rješavanje problematike koja se pojavljuje unutar Fininog odijela za upravljanje podacima (engl. *Data Management*).

PRILOZI

Prilog 1: Elektronička verzija diplomskog rada (.doc)

Prilog 2: Elektronička verzija diplomskog rada (.pdf)