

KOMUNIKIMI I SISTEMEVE TË EMERGJENCAVE DHE SINKRONIZIMI I TË DHËNAVE NË APLIKIMET E PAISJEVE MOBILE-RASTI I KOSOVËS

Doktoranti: **Zijadin KRASNIQI**

Dorëzuar
Universitetit European të Tiranës
Shkollës Doktorale

Në përmbushje të detyrimeve të programit të Doktoratës në Fakultetin Ekonomik dhe i Teknologjisë së Informacionit, me profil Menaxhim i Sistemeve të Informacionit, për marrjen e gradës shkencore “Doktor”

Udhëheqës shkencor: **Prof. dr. Adriana GJONAJ**

Numri i fjalëve: 51580

Tiranë, Qershor 2016

DEKLARATA E AUTORËSISË

Nën përgjegjësinë time personale deklaroj se ky studim është punë origjinale dhe nuk përmban plagjiaturë. Punimi është shkruar nga unë, nuk është prezantuar asnjëherë para një institucioni tjetër për vlerësim dhe nuk është botuar i tëri. Punimi nuk përmban material të shkruar nga ndonjë person tjetër, përveç rasteve të cituara dhe të referuara.

ABSTRAKTI

Menaxhimi i fatkeqësive natyrore është një domen kompleks, por edhe shumë i nevojshëm, kështu që nevojat e jetës normale kanë shumë dallim nga ato të situatave të emergjencave natyrore. Në këtë tezë, fokusi do jetë në menaxhimin e fatkeqësive natyrore, duke e eksploruar në risitë që na ofron teknologjia e informacionit për t'i lehtësuar këto situata. Agjencitë, institutet dhe organizatat që merren me menaxhimin e fatkeqësive natyrore, secila në vete, kontribuon me gjenerimin e një sasive të të dhënave, por që të gjitha nuk kanë një centralizim në ndërtimin e një DWH. Një sistem i tillë DWH-je mund të jetë efikas në organizimin dhe analizimin e një morie të informatave, e që mund të ndihmojë në krijimin e një modeli efikas me qëllim vendimmarrje në kohë. Në këtë punim do të paraqitet një studim i sistemeve datawarehouse për teknikat dhe standardet që sot për sot ekzistojnë dhe do të përshkruhet aplikimi i tyre në fushën e menaxhimit të fatkeqësive natyrore. Duke marrë për bazë problemet me të cilat po sfidohet sot Agjencia e Menaxhimit të Emergjencave (AME), si dhe mënyra si po aplikohen të dhënat operationale të shpërndara nëpër institute të tjera, nga kjo formë e këtyre të dhënave është ngritur hipoteza për ndërtimin e një data mart-i, që bën analizimin e rekordeve të detajuara. Në të prezantohet hap pas hapi modeli i këtij data mart-i, sipas modelit Kimball, që do të shërbejë për analiza të hollësishme të institutit mjekësor, të të dhënave nga instituti hidrometeorologjik, si dhe nga institute sizmologjike. Në grumbullimin e këtyre të dhënave, transformimin dhe ngarkimin e të dhënave, gjejmë aplikim metodat ETL/ELT, të cilat konsiderohen si detyrat më kritike të ndërtimit të një sistemi DWH. Përveç kësaj, do të kemi edhe mundësinë e sinkronizimit të të dhënave përmes

ndërtimit që DWH-të të komunikojnë në kohë reale, se deri më tani ka qenë një mangësi në fushën e menaxhimit të emergjencave.

Së fundi, ka lindur nevoja që, në veçanti, të shqyrtojmë mundësinë e zhvillimit të aplikacioneve mobile, ku dihet se pajisjet mobile janë një teknologji e re dhe shumë e nevojshme në menaxhim të emergjencave dhe fatkeqësive natyrore.

Zhvillimi i aplikacioneve të projektet doktorale për QOAME vjen si rezultat i ndërtimit të një arkitekture në platformën android, më pas zhvillimin e një modeli profesional në Talend open studio, që mundëson integrimin e të dhënave nga shumë burime të ndryshme të të dhënave.

Një përdorim është identifikimi i shpejtë i njerëzve që hasen në zonat e prekura nga fatkeqësitë natyrore, në identifikimin e atyre personave që nuk posedojnë ID Card apo diç tjetër për identifikim, duke përdorur identifikimin përmes shenjave të gishtërinjve apo edhe përmes numrit personal. Gjithashtu, aplikacioni mund të përdoret në sistemet e aksesit në institucionet publike, si në shkollë, biblioteka, spitale. Aplikacioni QOAME ka qasje në tërë agjencitë që kanë ndikim në menaxhim të fatkeqësive natyrore, si: ISK, IHMK, IMK, SHZK dhe ARC. Ky aplikacion zhvillohet që të kyçet edhe në web aplikacionet, si në ato që marrin informacione në kohë reale për lëkundjet sizmike dhe informacionet e motit në nivel botëror. Së fundi, duke u bazuar në studimet e literaturës, por edhe në punën konkrete me këto sisteme, prezantohen disa elemente të rëndësishme për suksesin e sistemeve DWH, data mart dhe zhvillimin e aplikacioneve mobile.

Fjalë kyçe: DWH, data mart, ETL, OLTP, OLAP, WEB mobile, sinkronizim, vendimmarrje, fatkeqësi natyrore, menaxhim i emergjencave.

ABSTRACT

Natural disasters managing is a complex domain, but also very necessary. The daily life needs differ so much from those of natural emergency situations. In this thesis, the focus will be on natural disaster management and exploring the innovations offered by the information technology in order to facilitate these situations. Agencies, institutions and organizations involved in the management of natural disasters have contributed to the generation of a large amount of data, but all of them lack the centralization in the construction of a DWH. A DWH system can be effective in organizing and analyzing a large amount of information, which can help create an effective model, which will be so useful in real time decision making. In this paper we will present a study on system data warehouse in regards to modern techniques and standards and will describe their application in the field of natural disasters management. Taking into account the problems and challenges which the Emergency Management Agency (EMA) is facing today, and the way operational data scattered across various institutions is being used, we have set an hypothesis to build a datamart, which has the capability for deep and detailed data analysis. Here we shall present the step by step modeling of this datamart, following the Kimball model, which will be useful for detailed analysis by medical, hydro-meteorological and seismic institutes. In collecting, transforming and loading the data we have applied the ETL/ELT methods, which are considered as the most critical tasks in building a DWH system. In addition to this, we will have the possibility to synchronize and build the data in a way that DWH will be able to communicate in real time, something which has not been possible until now in the field of emergency management.

We understand that there is a need to take into consideration and explore the possibility of developing mobile applications given that the mobile technology is a new and useful technology for natural disasters management.

This application developing for the doctoral project QOAME comes as result of architecture building in Android, then the development of a professional modeling in Talend Open Studio, which enabled the integration of data from many data sources. One of its uses is the quick identification of people found in areas affected by natural disasters. The application identifies the persons who do not have an ID card, or another identification document, by using identification through fingerprint or personal number. Besides this, the application can be used to grant access in various public institutions, such as schools, libraries, hospitals. QOAME Application has access to all the agencies involved in natural disasters managing, such as ISK, IHMK, IMK, SHZK and ARC. This application is developed to connect to web applications as well, such as applications which gather real time information on earthquakes and weather in the world.

The study of literature and the actual work with these systems has shown some important components of success for DWH systems, datamart and mobile application development.

Key words: DWH, datamart, ETL, OLTP, OLAP, WEB mobile, synchronization, decision making, natural disaster, emergency management

DEDIKIM

Ky punim doktrate i dedikohet nënës dhe babait tim.

FALËNDERIM

Falënderim dhe respekt të veçantë për udhëheqësen time shkencore, Prof. Dr. Adriana Gjonaj, e cila me maturi dhe këshilla të pakursyera më ka mbështetur dhe motivuar gjatë gjithë kohës së studimit tim doktorat. I falënderoj dhe u shprehu mirënjohje edhe profesorëve të Departamentit të Teknologjisë së Informacionit në UET.

Në fund, e falënderoj për zemërsisht edhe familjen time të dashur për mbështetjen morale, sepse pa mbështetjen e tyre ky punim nuk do ishte i mundur.

PËRMBAJTJA E LËNDËS

Deklarata e autorësisë	2
Abstrakti	4
Dedikimi	7
Falënderim	8
Përmbajtja e lëndës	9
Lista e ilustrimeve	14
Lista e tabelave ose diagrameve	17
Lista e shkurtimeve dhe e fjalorit	18

KAPITULLI I: PËRMBLEDHJE E STUDIMIT

1.1 Hyrje	20
1.2 Përmbajtja shkencore e Doktoratës - përshkrimi i problemit	21
1.3 Pyetjet kërkimore dhe hartimi i hipotezës	23
1.4 Caktimi i objektivave të studimit	25
1.5 Metodologjia e përdorimit gjatë studimit të doktoratës.....	32
1.6 Përmbledhje e literaturës kryesore	33
1.7 Struktura e punimit	43

KAPITULLI II: METODOLOGJIA DHE ARKITEKTURA E DATAWAREHOUSE (DWH).....

2.1 Hyrje	45
2.2 Koncepti i datawarehouse (DWH)	46
2.3 Historiku i sistemeve në mbështetje të vendimmarrjes (DSS)	46
2.4 Nevoja për t'i veçuar sistemet operacionale dhe informative.....	48

2.5. Metodologjia e zhvillimit të datawarehouse (DWH).....	50
2.6. Arkitektura e sistemeve të DWH-së	52
2.7. Konkluzion	56
KAPITULLI III: ANALIZA E DATA MART-ËVE PËR KRIJIMIN E QENDRËS SË EMERGJENCAVE NË KOSOVË	57
3.1 Hyrje	57
3.1.1 Arkitektura e data mart-ëve	58
3.1.2 Arkitektura e data mart-it të pavarur dhe e ruajtjes së të dhënave operacionale.....	61
3.2 Shtrimi i problemit për QOAME	66
3.3 Modelimi i të dhënave të përzgjedhura	68
3.4 Modelimi dimensional	70
3.4.1 Modelimi i formës së tretë normale (3NF)	72
3.4.2 Tabelat e fakteve	73
3.4.3 Tabelat e dimensioneve	74
3.4.4 Çelësat e tabelës së dimensioneve	75
3.4.5 Dimensionet e konfirmuara	76
3.4.6 Sjellja së bashku e fakteve dhe dimensioneve	78
3.5 Modelimi konceptual	81
3.6 Modelimi logjik	86
3.6.1 Tipi 1- SCD	90
3.6.2 Tipi 2- SCD	91
3.6.3 Tipi 3- SCD	93
3.7 Modelimi fizik	95

3.8 Përgatitja e një strategjie indeksimi	101
3.8.1 Indeksi i datawarehouse	102
3.8.2 Përmbledhje e indeksimit	103
3.8.3 Indeksi për tabelat e mëdha	104
3.8.4 Indeksi vetëm me lexim	104
3.9 Efektet e indekseve	105
3.9.1 Indeksi B-Tree	106
3.9.2 Indeksi bitmap	108
3.9.3 Treguesit bitmap	111
3.9.4 Indeksi i tabelës së fakteve	111
3.9.5 Indeksi i tabelave të dimensioneve	112
3.9.6 Rëndësia e indeksimit	113
3.10 Konkluzion	114
KAPITULLI IV: APLIKIMI I SISTEMEVE TË DWH-së NË RASTIN E EMERGJENCAVE DHE FATKEQËSIVE NATYRORE	116
4.1 Hyrje	116
4.2 Gjendja e emergjencave dhe fatkeqësive natyrore në vendin tonë	116
4.2.1 Analizimi i të dhënave në rastin e tërmeteve	117
4.2.2 Rasti i zjarreve pyjore	121
4.2.3 Analizimi i të dhënave të Institutit Hidrometeorologjik - Rasti i vërshimeve	122
4.2.4. Rasti i orteke të borës	124
4.3 Analizimi i të dhënave në rastin e emergjencave dhe fatkeqësive natyrore.....	126
4.4 Burimi i të dhënave kryesore të QOAME-së	129

4.5 Integrimi i të dhënave dhe menaxhimi i katastrofave	132
4.6 Konkluzion	134
KAPITULLI V: METODOLOGJIA E GRUMBULLIMIT TË TË DHËNAVE DHE SINKRONIZIMI NË SISTEMET ETL NË KOHË REALE	135
5.1 Hyrje	135
5.1.1 Nxjerrja e të dhënave	137
5.1.2 Transformimi.....	139
5.1.3 Ngarkimi i të dhënave.....	140
5.2 Mjetet e përdorura për ETL –Talend Open Studio for Big Data	141
5.3 Problematikat e paraqitura ne procesin e modelimit te data mart-it IM.....	162
5.3.1 Rezultatet e përfituara nga analizat per rastin e data mart-it IMK	165
5.4 Sinkronizimi i të dhënave.....	168
5.5 Arkitektura e data mart-ëve logjike dhe DWH-së në kohë reale	170
5.6 Datawarehouse në kohë reale dhe teknikat ETL.....	173
5.7 Sfidat dhe mundësitë e DWH-së në kohë reale	177
5.8 Rishikim i DWH-së në kohë reale	178
5.9 Përcaktimi i ETL-së në kohë reale	182
5.9.1 Qasjet ETL në kohë reale.....	182
5.10 Enterprise Application Integration (EAI).....	185
5.11 Ndërtimi i particioneve në kohë reale	191
5.11.1 Roli strategjik i menaxhimit të dimensioneve	191
5.11.2 Vetëm faktet apo edhe ndryshimet në dimension	193
5.11.3 Integrim të dhënash dhe integrimi i aplikacioneve.....	195

5.12 Konkluzion.....	198
KAPITULLI VI: APLIKIMET E PAJISJEVE MOBILE NË RASTIN E FATKEQËSIVE NATYRORE	201
6.1 Hyrje	201
6.2 Platforma Mobile.....	202
6.3 Kërkesat minimale për funksionim të platformës Android	203
6.4 Arkitektura Android.....	204
6.5 Rast studimor: Integrimi i databazës SQLite_QOAME e implementueshme në pajisje mobile Android	207
6.5.1 Integrimi i të dhënave në SQLite_QOAME përmes Talend Open Studio.	209
6.5.2 Projektimi i Databases SQLite _QOAME	210
6.6 Zhvillimi hyrës ne aplikacionin e QOAME.....	213
6.7 Krijimi i aktiviteteve duke përdorur intents.....	216
6.8 Databaza SQLite_QOAME-rasti i aplikacionit android për ARC.....	225
6.9 Databaza SQLite_QOAME-rasti i aplikacionit android për ISK	233
6.10 Databaza SQLite_QOAME-rasti i aplikacionit android për IHMK.....	236
6.11 Konkluzion	238
KONKLUSIONE DHE REKOMANDIME.....	239
Konkluzione	239
Rekomandime	244
LISTA E REFERENCAVE / BIBLIOGRAFIA	247

LISTA E ILUSTRIMEVE

Figura: 2.1 Vlera e informacionit si funksion i sasisë së të dhënave	47
Figura 3.1: Arkitektura data-mart e pavarur për një DWH	58
Figura 3.2: Arkitektura trishtresore e data mart-it të pavarur.....	63
Figura 3.3: Skema e arkitekturës së përzgjedhur.....	70
Figura 3.4: Proceset e modelimit dimensional.....	72
Figura 3.5: Mostër e tabelës dimensionale	74
Figura 3.6:Tabelat Fakt-dimension	77
Figura 3.7: Tabela e fakte dhe dimensioneve në modelin dimensional.....	79
Figura 3.8: Ndarja e dimensioneve në skemën.....	82
Figura 3.9: Projektimi i tabelës fakt për trajtimin e matjeve në kohë të sakta	83
Figura 3.10: Skema e data mart-it të IMK-së.....	89
Figura 3.11: Tipi 2 - SCD historiku perfekt i particioneve.....	92
Figura 3.12:Menaxhimi i tabelës dimension ‘surrogate key’	92
Figura 3.13: Implementimi i Tipit 3-SCD në përshkrimin nr. rendor-profesioni.....	94
Figura 3.14: Struktura e një indeksi B-Tree	106
Figura 3.15: Shembull i Indeksi bitmap	109
Figura 3.16: Indeksi bitmap : rikthim të dhënave	110
Figura 4.1: Sistemi Informativ Hidrometeorologjik Çek	125
Figura 4.2: Diagrami i kalimit të informacionit drejt DWH-QOAME.....	131
Figura 5.1: Komponentët e Talend Open Studio for Big Data	142
Figura 5.2 Tregon Entity Relationship Diagram (ERD) të sistemeve target	144
Figura 5.3 Tregon një punë të thjeshtë bashkimi të dizajnuar për të kombinuar dy burime të dhënash.....	148
Figurën 5.4: Komponenti tJoin- kutia e dialogut	148

Figura 5.5. Skema e fushës tJoin për burimin e të dhënave Dim_Data	149
Figura 5.6: Skema tJoin për burimin e të dhënave Dim_Kohe.....	150
Figura 5.7: Komponenti tMap i Talend open studio	151
Figura 5.8: Konfigurimit tMap	152
Figura 5.9:Diagrama e ngarkimit të tabelës Fakt IMK	154
Figura 5.10: Job-i Talend për ngarkimin e tabelës Fakt IMK.....	155
Figura 5.11 Job-i i TOS për shtim informacioni.....	156
Figura 5.12: Job-i i TOS për shtim dhe ngarkim te informacionit.....	157
Figura 5.13: Monitorimi i ekzekutimit të Job-it në Talend Open Studio for Big Data.....	159
Figura 5.14: Setup Traces-lista përzgjedhëse	160
Figura 5.15: Setup Traces-lista jo-përzgjedhëse	160
Figura 5.16:Paraqitja e tipareve Statistics	161
Figura 5.17: Data mart-ët logjike dhe arkitektura në kohë reale e data warehouse	171
Figura 5.18: Arkitektura data warehouse per datamartet e pavarur	172
Figura 5.19: Lidhja midis formës statike dhe skemës ylli në kohë reale	180
Figura 5.20 Lidhja logjike për particionet e datamart-ëve në kohë-reale	180
Figura 5.21: Diagrami Konvencional ETL	183
Figura 5.22: Diagrami mikro-batch ETL	184
Figura 5.23: Diagrami Konvencional EAI	186
Figura 5.24: Diagrami i DWH-së në kohë reale i EAI-së	188
Figura 5.25: Integrimi i aplikacioneve point-to-point	196
Figura 5.26: Integrimi i Aplikacioneve HUB dhe SPOKE	197
Figura 6.1.Arkitektura Android	204
Figura 6.2: Diagrami Entity Relationship i databazes QOAME	208
Figura 6.3 Job-i Talend për ngarkimin e të dhënave nga tabelat ORACLE SQL ne SQLite_QOAME	210
Figura 6.4 Organizimi arkitekturor klient/server për RDBMS	211
Figura 6.5: Organizimi arkitekturor i serverit SQLite	212

Figura 6.6:Aplikacioni hyrës për QOAME	215
Figura 6.7:Aplikacioni për login ne QOAME pas ekzekutimit të emulator-it	216
Figura 6.8 : Cikli jetësor i një aktiviteti	217
Figura 6.9: Klasa Welcome	219
Figura 6.10: Emërtimi i fajllave secondactivity.xml	222
Figura 6.11:Aplikacioni bazë i QOAME	224
Figura 6.12:Kerkimi sipas numerit personal ne databazen SQLite_QOAME	228
Figura 6.13:Procesi i krahasimit sipas shenjave te gishtrineje ne databazen SQLite_QOAME	229
Figura 6.14:Kerkimi sipas shenjave te gishtrineje ne databazen SQLite_QOAME	232
Figura 6.15:Qasja ne te dhënat e institutit seizmik te Kosoves	234
Figura 6.16:Aplikacioni QOAME ,qasja ne linqet per informacione seizmike	235
Figura 6.17: Te dhënat nga databaza SQLite per rastin e IHMK	237
Figura 6.18:Aplikacioni QOAME ,qasja ne linqet tjera per informim te mëtutjeshëm ...	238
Figura: Integrimi i data mart-ëve ne SQLite_QOAME	241
Figura: Paraqitja hap pas hapi e zhvillimit të aplikacionit QOAME	243

LISTA E TABELAVE OSE DIAGRAMEVE

Tabela 2.1: Krahاسimi i sistemeve operacionale dhe i sistemeve informative	49
Tabela 2.2: Qasja sipas Inmon-it dhe Kimball-it.	52
Tabela 2.3: Elementët e arkitekturave për ndërtimin e sistemeve DWH	54
Tabela 3.1 Elementet kriter për përzgjedhjen e modelimit	69
Tabela 3.2 Matrica Dimension/fakt	88
Tabela 4.1: Të dhëna historike të rastit të tërmeteve	118
Tabela 4.2: Instituti Seizmik i Kosovës (ISK), “Buletini mujor” (Mars 2010)	118
Tabela 4.3: Instituti Seizmik i Kosovës (ISK), “Buletini mujor” (Nëntor 2013)	121
Tabela 4.4: Të dhënat e vitit nga Shërbimi i Zjarrfikësve të Kosovës (SHZK)	122
Tabela 4.5: Të dhënat e Stacionit Meteorologjik në Pejë	122
Tabela 4.6: Të dhënat e Stacionit Meteorologjik në Prishtinë	123
Tabela 4.7: Të dhënat e Stacionit Meteorologjik në Ferizaj	123
Tabela 4.8: Katastrofat natyrore në tri dekadat e fundit (1980-2014)	127
Tabela 4.9: Katastrofat natyrore – fatkeqësitë teknologjike	127
Tabela 5.1: Burimet DBMS dhe atributet	145
Tabela 5.2: Krahاسimi i ekzekutimit të job-eve	166
Tabela 6.1 . Kërkesat minimale për Android dhe Symbian OS	203
Tabela 6.2: Tabela e të dhënave agjencioni të regjistrimit civil të databazës SQLite_QOAME..	226
Grafiku 4.1 Përmbledhje e katastrofave natyrore 1980-2014	128

LISTA E SHKURTIMEVE DHE E FJALORIT

AME	Agjencia për Menaxhim të Emergjencave
ARC	Agjencia e Regjistrimit Civil
AFIS	Automated Fingerprint Identification System
ART	Android Runtime
AOT	Ahead-of-time
ASCII	American Standard Code for Information Interchange
API	Application Programming Interface
AVD	Android Video Downloader
BI	Business Intelligence
CIF	Corporate Information Factory
CDC	Change Data Capture
CTF	Capture, Transform and Flow
CPU	Central Processing Unit
DWH	Datawarehouse
DSS	Decision Support System
DBMS	Database Management System
DML	Data Manipulation Language
DBA	Database Administrator
DPA	Data Processing Agreement
DP	Dimensione të Përbashkëta
ETL	Extract Transformation Load
ELT	Extract Load Transformation
EDWH	Enterprise Datawarehouse
ERP	Enterprise Resource Planning
EAI	Enterprise Application Integration
ER	Entity–Relationship

GB	GigaByte
GPRS	General packet radio service
GPS	Global Positioning System
IMK	Instituti i Mjekësisë së Kosovës
ISK	Instituti Sizmologjik i Kosovës
IHMK	Instituti Hidrometeorologjik i Kosovës
IT	Information technology
JIT	Just-in-time
KPSH	Kujdesi Parësor Shëndetësor
MZHE	Ministria e Zhvillimit Ekonomik
MARV	Metoda e Aksesit të Ruajtjes Virtuale
MASI	Metoda e Aksesit Sekuencial të Indeksuar
MGIS	Mobile Geographic Information Systems
OLTP	Online Transaction Processing
OLAP	Online Analytical Processing
ODS	Operational Data Store
OCI	Oracle Call Interface
PDA	Personal Digital Assistant
QOAME	Qendra Operative e Agjencisë së Menaxhimit Emergjent
QMF	Qendra e Mjekësisë Familjare
QKMF	Qendra Kryesore e Mjekësisë Familjare
QKUK	Qendra Klinike Universitare e Kosovës
RTDW	Real-Time Data Warehouse
SCD	Slowly Changing Dimensions
SMI	Sistemi i Menaxhimit të Informacionit
SHZK	Shërbimi i Zjarrfikësve të Kosovës
TOS	Talend Open Studio
XML	Extensible Markup Language

KAPITULLI I: PËRMBLEDHJE E STUDIMIT

1.1 Hyrje

Çdo shtet duhet t'i jep prioritet fushës së emergjencave, duke pasur parasysh se parandalimi, në të shumtën e rasteve, rezulton në zvogëlimin e fatkeqësive.

Jo rrallëherë, komuna të ndryshme të Kosovës janë përballë me situata emergjente, nga të cilat kanë humbur jetën njerëzit, si dhe janë dëmtuar pronat e patundshmëritë.

Zinxhiri i komunikimit në raste emergjencash, sipas strategjive dhe ligjeve në fuqi, jo rrallëherë ka pasur ndikimin e tij negativ. Ky zinxhir poshtë-lart i komunikimit është aq kompleks sa ndikon edhe në kohën e reagimit ndaj emergjencave.

E kjo, sidomos kur merret parasysh se Kosova deri tash, edhe pas planifikimit të hapjes së qendrave të operimit për emergjenca, nuk i disponon më këto qendra, të cilat këtë komunikim, megjithëse kompleks, do të sillnin një lehtësim të dukshëm.

Nëse situata tejkalon edhe nivelin e përgjegjësisë dhe mundësisë për veprim të rajonit të caktuar, atëherë kërkesa shkon në nivel të Agjencisë për Menaxhim të Emergjencave (Forumi për siguri, 2013).

Kur të përballemi me fatkeqësi, duhet në mënyrë të shpejtë të koordinohemi me një varg agjencish për të siguruar informacione optimale në lidhje me situatat e krijuara. Me këtë lind nevoja për një sistem integral të komunikimit për menaxhim efikas, të besueshëm dhe siguri në shkëmbimin dhe përpunimin e informacioneve (Andreas&Thomas, 2002).

Integrimi i këtyre sistemeve të ndryshëm është i rëndësishëm dhe kjo mund të realizohet (Sam & Lance, 2009). Sistemet e informacionit janë një fazë e rëndësishme në planifikimin dhe përgatitjen për fatkeqësitë (Sriganesh&Iyer, 2005).

Në një situatë emergjente informacioni përkatës në lidhje me elementet e përfshira është i nevojshëm. Për më tepër, procesi i menaxhimit të emergjencave është dinamik dhe përfshin disa hapa të caktuar (Tanasescu&Gugliotta, 2006:956).

Sistemet e komunikimit dhe informimit duhet të jenë të dizajnuara që të kenë fleksibilitet të besueshëm dhe të shkallzueshëm, në mënyrë që të funksionojnë në çdo lloj të incidentit, pa marrë parasysh shkakun, madhësinë, vendndodhjen apo kompleksitetin e tij (SIME, 2010).

1.2 Përmbajtja shkencore e Doktoratës - përshkrimi i problemit

Një nga problemet e shumta gjatë vizitave në Agjencinë e Menaxhimit të Emergjencave, është menaxhimi i të dhënave të sistemit të informimit dhe komunikimit. Kjo nuk është vetëm në aspektin e thjeshtë të gjenerimit të raporteve standarde lidhur me fatkeqësitë natyrore të ndodhura, por për të ardhur në ndihmë me këto të dhëna në kohë reale të të gjitha strukturave mbështetëse, kryesisht departamenteve:

- Agjencia e Regjistrimit Civil (ARC),
- Instituti i Mjekësisë së Kosovës (IMK),
- Instituti Sizmologjik i Kosovës (ISK),
- Shërbimi i Zjarrfikësve të Kosovës (SHZK) dhe
- Instituti Hidrometeorologjik i Kosovës (IHMK).

Në fakt, në Agjencinë e Menaxhimit të Emergjencave nuk ka asnjë metodologji të paracaktuar/vendosur për mbledhjen e informatave, ruajtjen dhe përditësimin periodik të të dhënave historike që lidhen me fatkeqësitë.

Megjithatë, mbledhja e të dhënave dhe ruajtja e tyre është bërë në mënyrë spontane, prandaj modele të ndryshme të mbledhjes së të dhënave dhe ruajtjes janë shfrytëzuar për ngjarjet e fatkeqësive.

Aktualisht, nuk ka sistem kombëtar vëzhgimi për rrezikun i cili mbledh të dhëna në një formë të centralizuar dhe prodhon një vlerësim të rëndomtë të rreziqeve të vendit – për shfrytëzim në planifikime për parandalim, zbutje dhe, madje, për masa gatishmërie.

Mungesa e kapaciteteve, sa i përket identifikimit të rreziqeve, gjithashtu, ndikon në paralajmërimin e hershëm. Nuk ka mekanizma të paralajmërimit të hershëm për të informuar autoritetet dhe publikun për trendët në rritje të rreziqeve – popullimi në rritje dhe ekspozimi i pasurisë ndaj rreziqeve të ndryshme natyrore dhe mjedisore (Armenv&Hachim&George, 2011). Përballë faktit se nuk kemi një databazë të centralizuar për menaxhim të të dhënave në kuadër të AME-së, atëherë mbetën edhe vonesa të mëdha në hedhjen e të dhënave të sakta dhe në kohë reale. Nga të gjitha këto departamente në kuadër të AME-së, kemi një vështirësi për informim të shpejtë dhe të saktë në lidhje me të dhënat e tyre që gjenerohen.

Kjo, në një mënyrë të domosdoshme, paraqet nevojën për një integrim të të dhënave nga të gjitha këto burime, duke krijuar kështu një sistem komunikimi unik për vetë AME-në.

Përveç integritetit të të dhënave dhe ndërtimit të sistemit DWH, nevojë tjetër është edhe modelimi i databazeve aktuale që janë ndërtuar apo duhet ndërtuar në departamentët e vet AME-së, e kjo mund të arrihet duke marrë në studim modelin data mart dhe gjithë proceset që do e bëjnë të mundur grumbullimin e informacionet nga burimet origjinale, transformimin dhe ngarkimin e tij në destinacion.

Si pjesë e trajtimit të modelimit hap pas hapi të data mart-it, do të jetë edhe procesi i ndërtimit ETL, që konsiderohet njëra nga detyrat më kritike në hapat e ndërtimit të sistemit DWH.

Me këtë, AME-ja mund të lehtësojë procesin e vendimmarrjes në situata kur është e përfshirë në mënyrë direkte në menaxhimin e fatkeqësive natyrore, e situata të tilla shumë herë duhet të monitorohen, në mënyrë të vazhdueshme për të pasur mundësi të paralajmërohen dhe të merren masa adekuate konform menaxhimit të fatkeqësive natyrore.

Ky studim përfshin analizimin e sistemeve të informacionit sipas standardeve të ndërtimit të bazës së të dhënave, të modelimit dhe aplikimit të këtyre të dhënave në pajisjet e vogla përmes aplikacioneve Mobile.

E tëra kjo analizë përmbledh sistemet:

- Sistemet DWH
- Modelimi shumëdimensional i data mart-ëve
- Sinkronizimi i të dhënave në sistemet ETL në kohë reale
- Zhvillimi i aplikacionit në platformën mobile

1.3 Pyetjet kërkimore dhe hartimi i hipotezës

Studimin cilësor, të thelluar teorik dhe përmbajtësor të kësaj teze, e ndihmon gjetja e përgjigjeve në disa nga pyetjet që mund të shtrohen në lidhje me ecurinë e punës së kësaj teze doktrate, duke bërë përmbushjen e qëllimit kryesor të këtij punimi:

Pyetja 1. Si paraqitet gjendja e trajtimit të informacionit në situatat e emergjencave dhe fatkeqësive natyrore në Kosovë?

Pyetja 2. Çfarë duhet ndërmarrë për të ndërtuar një sistem komunikimi në mes qendrave operative në agjencinë e menaxhimit të emergjencave (QOAME)?

Pyetja 3. Cilat do të ishin versionet e mundshme për të realizuar ndërtimin e një qendre të tillë?

Pyetja 4. Si do të ndikojë sinkronizimi i të dhënave në komunikimin e sistemeve në rastin e emergjencave dhe fatkeqësive natyrore në Kosovë?

Pyetja 5. Cili do jetë ndikimi i përdorimit të aplikimeve mobile në rastin e emergjencave në Kosovë?

Pra, ngritja e pyetjeve kërkimore, të shtruara më lart, na japin një këndvështrim më të qartë për gjithë ecurinë e tezës së doktoratës, duke u bazuar në nivelin e lartë të zhvillimit të teknologjisë së informacionit dhe komunikimeve në distanca të largëta, në mundësitë që ofrojnë këto zhvillime dhe në përvojën e deritanishme në këto fusha, konsideroj se ka shumë argumente që të pretendohet.

Zhvillimi i aplikacioneve për pajisjet mobile që janë në lëvizje mund të ketë këto ndikime:

- Lehtësimi i qasjes në informata nga vende të ndryshme dhe situata emergjente.
- Qasja në faqet web dhe shërbimeve web përmes pajisjeve mobile.
- Rritja e shkallës së besueshmërisë së informacionit.

Hartimi i hipotezës – Është detyrë imediate dhe shumë e rëndësishme sigurimi i një informacioni të centralizuar në kohë reale lidhur me situatat e emergjencës dhe të fatkeqësive natyrore në Republikën e Kosovës.

Ngritja e një qendre për sinkronizimin e të dhënave dhe përdorimi i aplikacioneve në pajisjet mobile do të ndihmonte në menaxhimin e emergjencave dhe të fatkeqësive natyrore, në parashikimin dhe parandalimin e mundshëm të tyre.

Rasti studimor: Ndërtimi i qendrës së emergjencave dhe të fatkeqësive natyrore në QOAME të Kosovës.

1.4 Caktimi i objektivave të studimit

Objektivë e hulumtimit (Mark, 2010) është që të shqyrtojë mënyrat për të krijuar një sistem të komunikimit që do të referohet si qendër operative. Qendra Operative e Agjencisë së Menaxhimit Emergjent (QOAME) është shumë institucionale, e cila siguron koordinimin e përgjithshëm të reagimit qendror, për incidentet me karakter kombëtar dhe zbatim të programit të menaxhimit emergjent (MPB, 2010).

Sot, më tepër se asnjëherë, është nevoja të krijojmë një sistem të koordinuar institucional, me të gjitha resurset dhe kapacitetet që i kemi në dispozicion, për të qenë të përgatitur të reagojmë në mënyrë të koordinuar dhe të planifikuar me rastin e fatkeqësive natyrore dhe incidenteve tjera (SIME, 2010).

Pra, zhvillimi i teknologjisë së informacionit dhe shfrytëzimi i tyre për qëllime të ndryshme në administrimin e QOAME-së, është një nga fushat kryesore të kërkimit të sotëm në botë

dhe moto kryesore në korporata për të analizuar mundësitë për zhvillimin e mëtejshëm të biznesit dhe përmirësimin e cilësisë së shërbimeve të tyre.

Qëllimi kryesor i këtij studimi është që të kontribuojë në fushën e implementimit të aplikacioneve të platformës mobile dhe ndërtimit të sistemeve të centralizuara të të dhënave në ndihmë të Qendrës Operative të AME-së (Dasho, 2014).

Objektivi I, i shtruar për zgjidhje

Projektimi dhe ndërtimi i sistemit DWH në kuadër të bashkërendimit dhe lehtësimit, të mundësohet integrimi i të të dhënave nga të gjitha burimet që ndërlidhen me QOAME. Ky integrim i të dhënave do bëjë të mundur marrjen e informacioneve në kohë reale.

- a. Ndërtimi i sistemit DWH
- b. Modelimi i data mart-it
- c. Metoda e grumbullimit të të dhënave ETL.

Objektivi II, i shtruar për zgjidhje

Zhvillimi i teknologjisë së informacioneve do të bëjë të mundur menaxhimin më efikas të emergjencave, fatkeqësive dhe krizave natyrore, që, kryesisht, po fokusohet drejt komunikimeve pa tela, posaçërisht rreth pajisjeve mobile, ku është rritur nevoja që aplikimet e ndryshme të jenë të pranueshme nga çdo pajisje, qoftë ai kompjuter personal apo pajisje mobile. Pra, projektimi i aplikacioneve të platformës mobile, si një platformë standarde e sistemit të teknologjisë së informacionit, duhet të merr parasysh këto objektiva:

a) Sinkronizimi i të dhënave në sistemet ETL në kohë reale

ETL është një model procesi i përdorur në Datawarehouse (DWH) për të integruar burime heterogjene të dhënash dhe për të mundësuar një analitikë uniforme të dhënash. Komponenti “Extract” i procesit ETL synon marrjen e të dhënave nga burime të ndryshme në një sërë formatesh.

Pjesa “Transform” e procesit ETL kryen operacione pastrimi dhe aplikon rregulla kodifikimi për t’i konvertuar të dhënat burimore në një formë më koherente.

Faza “Load” merr të dhënat e transformuara që i shkojnë për shtati skemës së uniformuar të të dhënave dhe i vë në dispozicion të një sistemi databazë, në të cilin, për shembull, mund të ekzekutohen vegla analizimi.

Në tregun e punës së Enterprise Database apo të kompanive që merren me databaza, vegla softuerike, si Oracle Warehouse Builder, DB2 Warehouse Edition dhe Microsoft SQL Server Integration Services, implementojnë dhe automatizojnë modelin ETL. Këto vegla i kushtojnë rëndësi nevojave të bizneseve të kësaj fushe në lidhje me performancën dhe shkallëzueshmërinë, si dhe kërkesat në lidhje me migrimin e të dhënave dhe auditimin.

Paralelet me kriteret për menaxhimin e të dhënave e të kërkimeve shkencore janë të qarta: të dhënat kërkimore, edhe brenda të njëjtës disipline shkencore, manifestohen në një varietet formatesh heterogjene dhe ekziston gjithnjë nevoja për “harmonizim të dhënash”. Shkalla dhe shpërndarja e të dhënave kërkimore, do të thotë, duhen aplikuar modelet dhe metodat e përdorura në magazinimin e të dhënave të zakonshme. (Daniel & Andre, 2011).

- Talend open studio – big data

Talend është ofruesi i parë i softuerit të integrimit të të dhënave me kod të hapur. Produkti i saj kryesor është Talend Open Studio.

Pas tre vjetësh kërkim dhe investim në këtë projekt, versioni i parë i softuerit u lansua në vitin 2006. Është një projekt me kod të hapur për integrim të dhënash bazuar në Eclipse RCP, e që, kryesisht, përkrahë implementime të orientuara drejt ETL dhe që ofrohet si për on-premises deployment, si dhe për modelin software-as-a-service (SaaS).

Talend Open Studio përdoret, kryesisht, për integrim mes sistemeve operative, si dhe për ETL (Extract, Transform, Load) për Business Intelligence dhe Datawarehouse, por edhe për migrim. Talend ofron një vizion tërësisht të ri, që reflektohet në mënyrën se si përdor teknologjinë, si dhe në modelin e saj të biznesit.

Kompania ka refuzuar modelin tradicional të pronësisë mbi aplikacionet dhe ka zgjedhur të ofrojë zgjidhje të hapura, inovative dhe të fuqishme softuerike me fleksibilitetin për t'i përballuar të gjitha kërkesat e integrimit të të dhënave që kërkojnë të gjitha organizatat. (Talend, 2015).

b) Zhvillimi i aplikacioneve mobile. Për arritjen e kësaj objektive nevojiten këto mjete:

- Platforma ANDROID STUDIO

Android është një sistem operativ i bazuar në Linux, me një ndërfaqe të programimit Java. Android është krijuar nga Aleanca Celular e Hapur, e cila është e udhëhequr nga Google. Android përdor një makinë të posaçme virtuale Machine Dalvik Virtual. Dalvik përdor

bytecode të veçantë, prandaj ne nuk mund të ekzekutojmë bytecode standarde Java në Android. Android ofron një mjet "dx", që lejon për të kthyer fajllat e klasave Java në fajlla "dex" (Dalvik executable).

Aplikacionet Android janë të paketuara në një fajllë apk (Android Package). Për të thjeshtuar zhvillimin e programeve në android, Google ofron Android Development Tools (ADT) për Eclipse. ADT kryen automatikisht konvertimin nga klasa në fajllë dex dhe krijon apk gjatë zhvillimit. Android suporton grafikën 2-D dhe 3-D, duke përdorur libraritë OpenGL dhe suporton magazinimin e të dhënave në një bazë të dhënash SQLite. Çdo aplikacion Android ekzekutohet në procesin e vet dhe sipas përdoruesit të tij. Prandaj, aplikimi është i izoluar nga aplikacionet e tjera dhe një aplikim tjetër nuk mund të dëmtojë lehtë aplikacionet e tjera Android.

Si një diskutim i gjatë në një fushë të gjerë, nuk është vështirë të kuptohet pse ka pasur një konfuzion në lidhje me çfarë është me saktësi platforma Android. Android nuk është një implementim i Java ME për aplikacionet Android që janë të shkruara në gjuhën Java. Aplikacionet nuk ekzekutohen në një makinë Java ME virtuale dhe kompilimi i klasave të shkruara në Java nuk ekzekutohet vetvetiu në platformën Android.

Kjo platformë nuk është pjesë e forumeve standarde të Linux Phone (LiPS) ose të Open Mobile Alliance (OMA). Android ekzekutohet në një bërthamë Linux, ndërsa qëllimet e tyre janë të ngjashme. Staku i plotë softuere i Android-it shkon më tej me një fokus më të gjerë të këtyre standardeve në vizionin e organizatave të ndryshme.

Android është i përbërë nga pjesë të ndryshme duke përfshirë:

- Një sistem operativ me bërthamë Linux që siguron, të paktën, një ndërfaqe minimale me harduere, menaxhimin e kujtesës, kontrollin e procesit, të gjithë të optimizuar për pajisjet celulare.
- Një kohë ekzekutimi për aplikacionet Android, duke përfshirë makinën virtuale Dalvik dhe libraritë kryesore që ofrojnë funksionet specifike Android. Koha e ekzekutimit është projektuar për të qenë e vogël dhe efikase në përdorim në pajisjet e lëvizshme.
- Një sërë librarish me burim të hapur për zhvillimin e aplikacionit, duke përfshirë SQLite, WebKit, OpenGL dhe një menaxher media. Një ndërfaqe me përdoruesit që duhet për të lansuar aplikimet. (Aplikime mobile, 2015)

JAVA

Java është gjuhë object-oriented. Objekti në object-oriented është gjithçka dhe aplikacionet punojnë në bazë të saj. Objekt sipas object-oriented është gjithçka që mund të shihet, të preket etj. Gjithashtu, çdo pjesë e vogël e programit është objekt. Objektet komunikojnë në mes veti duke i dërguar mesazhe njëri-tjetrit.

Në këtë formë aplikacioni në object-oriented funksionon. Java ka shumë klasa mbi të cilat mund të ndërtohen aplikacionet, të cilat përdorin rrjetin, protokollet e tij etj. Me anë të këtyre klasave është e mundur që aplikacionet të ndërtohen, madje edhe client/server aplikacionet. Java është gjuhë kompajluese, edhe gjuhë interpretuese.

Java, përveç që ka veglën (ang. tool) java (Java Compiler) ka edhe JVM (Java Virtual Machine), e cila interpreton kodin e Java-s, i cili është i kompajluar. JVM është një vegël që

bytecode-in (kodi i kompajluar) e kthen në machine code (kodin binar). Prandaj, Java është një gjuhë e fuqishme.

Vetë fuqia e saj object-oriented dhe karakteristikat e programimit OO, bëjnë që Java të jetë një ndër gjuhët më të fuqishme. Ajo punon me klasat bazike që ka, veglat për gabime (ang. exceptions), si dhe karakteristikat e tjera të OOP. Java është një gjuhë e sigurt. Java mundëson aplikacionet që t'i krijosh si paketa (ang. packages). Gjithashtu, Java përdor Collection. Java, po ashtu, e ka platformën e vet. Java është e dizajnuar për qëllime të përgjithshme. Ajo sot më së shumti përdoret në aplikacionet e biznesit, por kjo nuk do të thotë se Java nuk mund të realizojë aplikacione të tjera, si, për shembull, të shkencës, të teknologjive të tjera.

Java është gjuha e vetme deri më tani e cila e ka aftësinë për të ekzekutuar të njëjtin kod në të gjitha platformat apo pajisjet.

Java aplikacionin, të cilin ne e përdorim në sistemin operativ Windows, të njëjtin mund ta përdorim edhe në Linux, Solaris, Unix, Mac etj. Java ka një performancë të lartë si gjuhë programuese. Java si çdo gjuhë tjetër OO mundëson reuseability (shqip: ripërdorimin e kodit). Java është krijuar për të bërë programe të vogla dhe efikase për pajisje (dhe platforma) të ndryshme. Edhe pse Java nuk kishte një sukses të madh në komercializëm, ajo sot përdoret në shumë pajisje të ndryshme, siç janë: mobilët (cell phones), PDA (Personal Digital Assistant). (Mesime Java).

SQLite

SQLite është një Open Source Database, e cila është e ngulitur në Android. SQLite përkrahë tipare standarde që lidhen me databazat si sintaksë SQL, transaksione dhe stejtmënt të përgatitura. Përveç kësaj, ajo kërkon shumë pak memorie për të punuar (afërsisht 250 KByte). SQLite është një librari softueri që implementon një vegël databaze SQL të vetëmbajtur, pa serverë, me zerokonfigurim dhe transaksional. SQLite është vegla e databazës SQL më i përdorur në botë. SQLite përkrahë tipat e të dhënave TEXT (ngjashëm me String në Java), INTEGER (ngjashëm me long në Java) dhe REAL (ngjashëm me double në Java). Të gjithë tipat tjerë duhet të konvertohen në njërën nga këto fusha për t'u ruajtur pastaj në databazë. SQLite vetë nuk arrin të vlerësojë nëse të dhënat janë shkruar sipas tipit të definuar, për shembull, të shkruhet një integer në një kolonë varg (string column) dhe anasjelltas. SQLite është e disponueshme në çdo pajisje Android. Të përdorësh një databazë SQLite në Android nuk kërkon ndonjë ndërtim ose administrim databaze. Stejtmëntat SQL për krijimin dhe përditësimin e databazave vetëm se duhet të jenë të definuara dhe databaza menaxhohet automatikisht për përdoruesit nga platforma Android. (Aditya, 2015)

1.5 Metodologjia e përdorimit gjatë studimit të doktoratës

Teknologjia informative sot është bërë një bazë për zhvillimin e shumë sistemeve të informacionit. Në shumicën e fushave jetësore është bërë shumë e vështirë të realizohen komunikime, ndërlidhje institucionale apo edhe komunikime ndërkombëtare pa përdorim të gjerë të teknologjisë së informacionit.

Teknologjia informative merr për bazë komunikimin e pajisjeve mobile, që është e një rëndësie të veçantë. Dhe, duke ditur se numri i tyre është shumë më i madh se numri i kompjuterëve, është e arsyeshme që edhe përmes pajisjeve mobile të mund të shfrytëzohen resurset e ndryshme.

Si pikë e rëndësishme në teknologjinë e informacionit është përdorimi, adaptimi dhe ndërtimi i aplikacionit mobil. Po ashtu, edhe standardet dhe zhvillimi i teknologjive të ndryshme në fushën e pajisjeve mobile janë pjesë e rëndësishme e studimit.

Shërbimet në hapësirat mobile po bëhen për arsye se janë duke u optimizuar protokollet dhe janë duke u bërë adekuate për pajisje mobile.

Studimet e reja të teknologjisë së informacionit dhe të komunikimit kanë hapur mundësi për përmirësime të rëndësishme në integrimin dhe koordinimin e shërbimeve emergjente. Aplikimet e teknologjive të informacionit ofrojnë mundësi të reja për integrimin dhe përmirësimin e shkëmbimit të informacionit ndërmjet institucioneve që menaxhojnë me shërbimet emergjente dhe fatkeqësive natyrore.

Me këtë rast, ekzistojnë një numër metodash për qasje të ndryshme, si për analizim të sistemit, zhvillim të aplikacioneve, implementimin e tij dhe shfrytëzimin, e më pas edhe mirëmbajtjen e sistemit.

Nëse marrim për bazë objektivin I, metodologjia mbi të cilën ndërtohet sistemi DWH është specifikuar sipas hapave vijues.

1.6 Përmbledhja e literaturës kryesore

Studimin cilësor, të thelluar teorik dhe përmbajtjesor të këtij punimi, e bën të ndarë në disa grupe kryesore të literaturës:

- Ndërtimi i sistemeve për menaxhimin e fatkeqësive natyrore dhe emergjencave
- Sistemi datawarehouse
- ETL, OLTP
- Sinkronizimi i të dhënave në sistemet ETL në kohë reale
- Aplikimet mobile

Ndërtimi i sistemeve për menaxhimin e fatkeqësive natyrore dhe emergjencave

Në studimin e tanishëm të menaxhimit të fatkeqësisë natyrore, sistemet ende nuk janë në gjendje të mbledhin informacion online për aktivitete pas ndodhjes së fatkeqësisë natyrore, si, për shembull, shpëtimi i njerëzve nga vendet e prekura në mënyrën e duhur.

Për t'i menaxhuar këto probleme nevojitet një sistem efikas që mund t'i mbledhë informacionet nga lokacionet e dëshiruara dhe që mund t'i vendosë ato informacione në kohë reale. (Muhammad, 2012).

Në vazhdim do paraqesim disa sisteme të ndërtuara për menaxhimin e fatkeqësive natyrore dhe emergjencave:

Sipas Karnatak në vitin 2011 prezantuan përdorimin e aplikacionit të pajisjeve mobile për menaxhimin e përmbytjeve në Asam, Indi. Aplikacioni mobil i zhvilluar në këtë rast përdori Windows, si platformë për zhvillimin e aplikacionit. Të dhënat nga terreni në zyrën e kontrollit u transferuan përmes mesazhit XML, që përmban të dhëna të ndryshme që përfshijnë imazhe dhe gjeolokacion.

Transmetimi i të dhënave me bazë XML, nga aplikacioni mobil në server, rezultoi të ishte një mënyrë efikase komunikimi. Të dhënat e marra nga pajisja mobile u integruan në një

aplikacion Web-GIS në një ambient në kohë reale dhe të dhënat e tjera nga përmbytjet e marra nga GIS ishte e mundur të aksesoheshin nga serverë në distancë.

Doner dhe Yomralioglu (2008) kanë bërë një vlerësim të MGIS (Mobile Geographic Information Systems), duke e përdorë si vegël për mbledhjen e gjeo të dhënave.

Ata përshkruan dhe diskutuan mbi MGIS-in, i zhvilluar për mbledhje të dhënash në terren, që gjithashtu përdor GPS. Rezultati ishte me përfitime të dukshme dhe efikasitet, saktësi dhe ulje të kostove, duke kryer çdo gjë në kohë të duhur sipas projektit. Ata e quajtën MGIS-in një vegël të nevojshme për të mbledhë të dhëna në terren, duke treguar kështu efikasitetin dhe saktësinë e pritshme nga një aplikacion për menaxhim fatkeqësish natyrore. (TANYA, 2014).

Interes të veçantë duhet t'i kushtohet edhe projektit të **Fondacionit Sahana**, i cili synon të ofrojë një komplet aplikacionesh për menaxhim fatkeqësish natyrore të bazuara në ueb.

Aplikacionet e zhvilluara si pjesë e Sahana-s përfshijnë Regjistrin e personave të zhdukur, Regjistrin e organizatave, Sistemin e menaxhimit të kërkesës dhe ofertës, Regjistrin e strehimoreve, Menaxhimin e inventarit, Vetëdijesimin mbi situatën dhe koordinimin e vullnetarizmit. Sahana, gjithashtu, ka edhe mjete për sinkronizim ndërmjet shumë instancave, për shembull, u mundëson përgjegjësve ose zyra të qarqeve që të sigurojnë të dhëna në lidhje me viktimat në terren dhe të shkëmbejnë të dhëna me të tjera zyra në terren, shtabe ose emergjenca civile. Sahana bazohet në ueb dhe përballet me problemin e përcimit të informacionit te serveri qendror, kështu që nuk mund të përdorë pajisje të lidhura me internet, duke krijuar ndërprerje në operatione. Për më shumë, përdoruesit nuk kanë akses në të dhëna në sistemin Sahana kur janë offline.

Ne kemi vënë re se korniza e sinkronizimit në Sahana mbështetet plotësisht në modifikimin e vulave kohore dhe si e tillë përballlet me problemin e orëve të pa sinkronizuara në pajisjet mobile, duke sjell edhe konflikt, shkaku i mosnjohjes së tyre. Nga ana tjetër, aplikacionet në Sahana nuk mund të ndryshohen nga përdoruesit, kështu që nuk është e lehtë të përshtaten me nevojat e përdoruesve në raste fatkeqësish. (Nicholas, 2012).

Fillimisht, u konceptua si Sahana Software, u zhvillua nga komuniteti IT në Sri Lankë, që të ndihmonin në gjetjen e personave të humbur pas cunamit të Oqeanit Indian në vitin 2004. Kjo paketë softuerike evoluoi derisa u njohë si Sahana Eden. Ajo u përkrahë nga një komunitet vullnetarësh, që përfshin vullnetarë të menaxhimit të situatave emergjente, akademikë, studentë dhe kompani. Komunitetet promovuan kostumizimin, modifikimin e softuerit, duke e bërë atë softuer me kod të hapur, me standardet përkatëse.

Projekti ishte implementim i suksesshëm i zgjidhjes së menaxhimit të emergjencave natyrore, duke ndihmuar organizatat e ndihmës në raste emergjence, njerëzit e përfshirë, asetat dhe inventarët. Projekti përdori harta për të bërë vlerësime dhe për të vetëdijesuar njerëzit rreth situatës.

Sipas Asif, i cili gjatë vitit 2012 ka prezantuar një program nëpërmjet pajisjeve mobile për menaxhimin e fatkeqësive natyrore për t'u përballur me lehtësimin e situatave pas fatkeqësive natyrore.

Prandaj, për të mundësuar mbledhjen e informacionit dhe aksesimin e tij nga publiku dhe vendimmarrësit, u mendua të bëhej një aplikacion për pajisje mobile. Kjo do të ndihmonte në procesin e rehabilitimit dhe shpëtimit gjatë këtyre fatkeqësive natyrore. Ky program mundësoi mbledhjen e të dhënave nga zonat e goditura nga fatkeqësitë natyrore edhe në modalitetin offline, për shembull, kur nuk ka mbulim me pajisje mobile. Aplikacioni që në

gjendje të transferonte të dhëna në një databazë qendrore përmes aplikacionit të uebit. Këto të dhëna u vendosën në një databazë qendrore, për t'u përdorur për më tepër analizim.

Të gjitha këto raste sjellin shembuj perfekt dhe motivimin për të menduar një zgjidhje të bazuar në teknologji, e cila do të ndihmonte në mënyrën se si përballemi me fatkeqësitë natyrore në vendin tonë. (TANYA, 2014).

Sistemi i propozuar pritet të jetë një zgjidhje e mirë për problemet e sipërpërmendura. (Muhammad, 2012).

Në dekadën e fundit, në vende të ndryshme të botës, fatkeqësitë natyrore janë rritur si nga shpeshtësia ashtu edhe nga intensiteti. Tërmeti në Haiti, uragani “Katrina”, cunami në Oqeanin Indian dhe cikloni “Nargis”, kanë qenë fatkeqësi natyrore të mëdha. Këto ngjarje e kanë shtyrë njerëzimin të mendojë për mënyrat se si të përballlet me to. Numri në rritje i njerëzve të prekur dhe hapësira e madhe e prekur nga shkatërrimet, si dhe humbjet e mëdha ekonomike nga fatkeqësitë natyrore, janë një arsye e madhe për shqetësim.

Rritja ekonomike e vendeve të prekura pëson ulje dhe kërkon dekada për t'u rimëkëmbur plotësisht. Sidomos, vendet në zhvillim dhe ato nën zhvillim duan më shumë kohë për të marrë veten nga këto humbje... Kjo bën thirrje për metoda dhe teknika të përforcuara në lidhje me menaxhimin e fatkeqësive natyrore. Çdo vend duhet ta kuptojë këtë nevojë dhe të mendojë për mënyra sa më të mira për t'i menaxhuar këto fatkeqësi.

Është e vërtetë, fatkeqësitë natyrore shkaktojnë shumë dëme dhe humbje jete, e që vështirë mund t'u shmangemi. Edhe pse shkencëtarët dhe hulumtuesit punojnë pandërprerë për të parashikuar motin, ndonjëherë dështojnë të parashikojnë impaktin e një fatkeqësie të caktuar. Disa herë ata zihen në befasi nga fatkeqësitë e tilla natyrore.

Por, menaxhimi i operacioneve të shpëtimit dhe i ardhjes në ndihmë mund të planifikohen para se të ndodhin fatkeqësitë natyrore dhe të vihen në zbatim në momentin që ndodhin.

Teknikat e menaxhimit të fatkeqësive natyrore, edhe nëse ekzistojnë, duhet të rishikohen dhe përditësohen vazhdimisht, sepse tashmë ka teknologji të reja dhe duhet të përdoren për menaxhimin e fatkeqësive që të sjellin rezultate fitimprurëse. (TANYA, 2014).

Mungesa e informacionit të saktë, preciz, mbi lëvizjen dhe komunikimin e popullsisë gjatë fatkeqësive natyrore, mund të kufizojë efikasitetin e reagimit me qëllime humanitare.

Sidoqoftë, prezenca në rritje e pajisjeve mobile ka shpalosur mundësi të reja për akses në informacione të tilla.

Të dhënat në kohë reale të pajisjeve mobile mund të na ofrojnë vëzhgime me vlerë mbi sjelljen e popullsisë së prekur gjatë një fatkeqësie. Duke ekzaminuar të dhënat e aktivitetit të pajisjes mobile përpara dhe pas një fatkeqësie natyrore, mund të krijohet një pikënisje e të kuptuarit të sjelljes në momente emergjence dhe mund të krijohet kapaciteti për të matur shkallën e rimëkëmbjes nga fatkeqësia. (WFP, 2014).

Sistemi datawarehouse

Kjo pjesë e punimit paraqet anën teorike të ndërtimit të DWH-së, që është shumë e mirëpritur për proceset vendimmarrëse të fushave të biznesit, telekomunikimit, shëndetësisë. Por, ndikim të madh ka edhe te menaxhimi i fatkeqësive natyrore, e që, në fakt, kjo lëmi ka nevojë për të dhëna në kohë reale.

Sistemet DWH nga këndvështrimi i zhvillimit janë shumë evolutive dhe dinamike dhe nuk ndalen së transformuari konform kërkesave. Sistemi DWH është sistem që grumbullon informacione nga burime të ndryshme heterogjene (Kimball&Becker, 2008). Megjithëse një datawarehouse, fillimisht mund të jetë e vogël dhe e thjeshtë, të ofrosh të dhënat e duhura ka

një rëndësi shumë të madhe në vendimmarrje të ndryshme, të cilat do të mbështeten në këto të dhëna dhe për këtë qëllim shpeshherë kërkohen të dhëna historike, të shtrira në kohë dhe të regjistruara në një nivel shumë të detajuar.

Një datawarehouse është një sistem që merr dhe i konsolidon të dhënat në mënyrë periodike, nga sisteme të tjera në modele dimensionale ose të normalizuara. Zakonisht, mbajnë të dhëna për vite të ndryshme dhe janë të modeluar të lehtësojnë kërkimin e informacionit për business intelligence (BI). Analizat mund të kërkojnë dhe një akses në kohë reale të të dhënave. Meqë një analist të dhënash përparon vazhdimisht në përvetësimin dhe kuptueshmërinë e të dhënave, një raport i thjeshtë nuk është kurrë i mjaftueshëm, kështu që për të rritur shkallën e vendimmarrjeve të tij, janë gjithmonë të nevojshme ad hoc query-të dhe analizat e sofistikuar. Në kuadër të përparimit të vazhdueshëm, shumica e datawarehouse-ve nuk konsiderohen kurrë të përfunduara, andaj ato janë gjithnjë në zhvillim e sipër për t'iu përgjigjur sa më mirë ndryshimeve që ndodhin në fushat për të cilat ato janë ndërtuar, e sidomos në rrafshin e menaxhimit të emergjencave që nevojiten çdo herë ndryshime (Çela, 2014).

ETL, OLTP

Brenda një sipërmarrje ka sasi të mëdha aplikimesh të ndryshme dhe burime të dhënash, të cilat duhet të integrohen. Përpunimi i transaksionit on-line (OLTP) dhe DWH nuk mund të bashkekzistojnë në mënyrë të efektshme në të njëjtin mjedis databazë, meqenëse databazat OLTP mbajnë të dhëna aktuale në shumë detaje, kurse DWH-të merren me të dhënat e përmbledhura dhe shpesh historike të bashkërenduara globalisht. OLTP thekson efikasitetin

e transaksioneve të përditësimit të shkurtër, ku secila mbulon një pjesë të vogël të databazës, ndërsa DWH kërkon studim të

gjatë të një pjese të madhe të databazës.

Procesi ETL ekstrakton të dhënat nga sistemet burim, i transformon të dhënat sipas rregullave të biznesit dhe i ngarkon rezultatet në DWH të synuar. Aktivitetet ETL përfshijnë:

(a) identifikimin e informacioneve përkatëse në anën burim; (b) ekstraktimin e këtij informacioni; (c) personalizimin dhe integrimin e informacionit që vijnë nga disa burime në një format të vetëm; (d) pastrimin e setit të të dhënave që vjen si rezultat mbi bazën e rregullave të databazës dhe biznesit; dhe (e) përhapjen e të dhënave në DWH / data mart-e.

Cilësia e të dhënave dhe efektiviteti i një DWH-je varet direkt nga efikasiteti i procesit ETL.

Prandaj, një proces ETL cilësor sjell fuqi vendimmarrjeje cilësie. Studimi ka zbuluar se 70% e përpjekjeve për implementimin e softuerëve dhe të mirëmbajtjes së DWH-së harxhohet në sistemin ETL. Qasjet tradicionale në ETL bazohen në përhapjen e përditësimeve të vërteta të të dhënave dhe jo në ngarkimet në rritje.

Sinkronizimi i të dhënave në sistemet ETL në kohë reale

Përdorimi i DWH-së është rritur ndjeshëm viteve të fundit dhe luan një rol themelor në shumë procese vendim mbështetëse, sidomos të AME. Më shumë AME, që tani duhet të integrojnë përditësimin e të dhënave në kohë, përditësimi midis databazave transaksionale dhe analitike është bërë një çështje kritike.

Për të arritur efikasitetin maksimal, disa burime të të dhënave të AME-së kërkojnë futjen e të dhënave transaksionale në një DWH në kohë reale ose në periudha të rregullta më të shkurtra se ngarkimet statike tradicionale. Shumica e ndërtimeve të DWH-së e trajtojnë të gjithë DWH-në si një entitet të vetëm, homogjen sa i përket sinkronizimit.

Të gjitha të dhënat e ruajtura rifreskohen në të njëjtën kohë gjatë të njëjtave ngarkesa periodike. Sidoqoftë, kjo qasje nuk është më e vlefshme pasi porcionet e DWH-së, që përbëjnë modele të ndryshme agjencish të emergjencave, kanë kërkesat e tyre të përditësimit, të cilat ndryshojnë me kalimin e kohës. Për të adresuar këtë problem, paraqesim një strukturë që përdor sete parametri, për të përcaktuar opsionin më të përshtatshëm të përditësimit për një model të caktuar të AME-së.

Segmentet e DWH-së përditësohen në intervale të ndryshme në varësi të burimit të informacionit në mjedisin e përpunimit të transaksionit online, si edhe në kërkesat e aplikimeve analitike. Shpeshtësia e procesit të ngarkimit të DWH-së përcakton pikat e përditësimit midis sistemeve të transaksionit dhe DWH-në me aplikime analitike. Normalisht, DWH mbështetet në përditësime statike, me procese ngarkimi në grup që ndodhin çdo ditë, javë, muaj ose në intervale të rregullta. Sidoqoftë, nevojat e sotme të AME-së kërkojnë një mjedis analitik që siguron:

- integrim të vazhdueshëm të të dhënave me periudha më të shkurtra për kapjen dhe ngarkimin nga burimet operacionale,
- një vegël aktiv vendimi që bën rekomandime dhe
- disponueshmëri e lartë.

Sinkronizimi i DWH-së në kohë reale me sistemet transaksionale kërkon reduktimin e intervalit midis pikave të përditësimit. Për të arritur këtë opsion dinamik, sistemi i databazës analitike duhet të reflektojë menjëherë në përditësimin e të dhënave transaksionale. Familjet e ndryshme të algoritmeve implementojnë mirëmbajtjen e DWH-së në kohë reale. Megjithatë këto algoritma dhe derivatet e tyre të shumta rritin saktësinë dhe performancën, ato nuk

marrin parasysh se cili opsion i përditësimit të të dhënave është më i përshtatshëm për një model AME të caktuar.

Për shembull, kontrolli i të dhënave nga Instituti Sizmologjik i Kosovës kërkon marrje dhe konsolidim të të dhënave në kohë reale, ndërsa kontrolli i regjistrimit civil kërkon ngarkim të dhënash çdo muaj. Nevoja të ndryshme bashkekzistojnë në të njëjtin ent të agjencisë së emergjencave, duke ndryshuar nëpër departamente të ndryshme. Për më tepër, jo të gjitha modelet e AME-së kërkojnë sinkronizim të dhënash në kohë reale. E njëjta DWH duhet të jetë në gjendje të mbështetë disa opsione përditësimi bazuar në një agjenci të menaxhimit të fatkeqësive natyrore në rrethana teknike të caktuara (Cristina&Ferreira, 2006:53).

Aplikimet mobile

Mungesa e informacionit të saktë, preciz, mbi lëvizjen dhe komunikimin e popullsisë gjatë fatkeqësive natyrore, mund të kufizojë efikasitetin e reagimit me qëllim humanitar. Sidoqoftë, prezenca në rritje e pajisjeve mobile ka shpalosur mundësi të reja për akses në informacione të tilla.

Të dhënat në kohë reale të pajisjeve mobile mund të na ofrojnë vëzhgime me vlerë mbi sjelljen e popullsisë së prekur gjatë një fatkeqësie. Duke ekzaminuar të dhënat e aktivitetit të pajisjeve mobile përpara dhe pas një fatkeqësie natyrore, mund të krijohet një pikënisje e të kuptuarit të sjelljes në momente emergjence dhe mund të krijohet kapaciteti për të matur shkallën e rimëkëmbjes nga fatkeqësia. (WFP, 2014).

Ky kapitull shpjegon zgjidhjen e propozuar të bazuar në pajisje mobile për situatën pas fatkeqësisë. Pra, ky kapitull bazohet në një sistem operativ Android, që është një softuer me burim të hapur nga Google.

Android SDK i jep mundësinë zhvilluesit të aplikimit që të krijojë aplikacione plotësisht të hapura për pajisje mobile, që do të thotë se çdo përdorues aplikacioni mund t'i përdorë të gjitha funksionet e pajisjes mobile, për shembull, bërja e thirrjeve, dërgimi i sms-ve, përdorimi i kamerës së celularit dhe ruajtja e të dhënave në memorien e celularit.

Android bazohet në Kernel Linux, por ka mundësi të zhvillohet edhe për MAC OS X, Linux dhe Microsoft Windows. Një epërsi e madhe e Android-it është dhënia e mundësisë për të pasur akses në librari të ndryshme për zhvillim të pasur të aplikacioneve. Aspekti i mosfunksionimit të një aplikacioni që zhvillohet në Android e ndihmon përdoruesin të kombinojë të dhëna nga burime të ndryshme. Për shembull, një përdorues mund të kombinojë të dhëna nga ueb-i dhe nga telefoni celular i dikujt për të pasur një përvojë përdoruesi më relevante.(Muhammad, 2012).

1.7 Struktura e punimit

Ky punim është ndarë në 8 kapituj, ku përshkruhet struktura si më poshtë:

Kapitulli I: Paraqet hyrjen në problematikën e menaxhimit të emergjencave dhe fatkeqësive natyrore, duke paraqitur me theks të veçantë komunikimin ndërmjet sistemeve të ndryshme dhe me këtë rast edhe integrimin e këtyre sistemeve përmes datawarehouse DWH.

Kapitulli II: Jep një përshkrim të sistemeve DWH dhe problemet që ai i zgjidhë në mbështetje të vendimmarrjes (DSS). Duke marrë për bazë hulumtimet shkencore dhe literaturën, është një paraqitje e sistemeve operacionale dhe informative. Po ashtu, edhe një përmbledhje metodologjike dhe arkitekturore që aplikohet sot në ndërtimin dhe projektimin e DWH-së.

Kapitulli III: Vijon me aplikimin e modelit shumëdimensional, duke ndërtuar një analizë të detajuar të regjistrave të agjencisë së emergjencave, në të paraqitet hap pas hapi modeli i një data mart-i. Në kuadër të ndërtimit të data mart-ëve kemi edhe indeksimin. Para së gjithash, me këtë do të krijohet një ide më e qartë se si këto sisteme do të ndikojnë në analiza më të thella në vendimmarrjen strategjike të menaxhimit të emergjencave.

Kapitulli IV: Paraqet një pasqyrë të gjendjes së emergjencave dhe fatkeqësive natyrore në vendin tonë dhe më gjerë në Evropë, në tri dekadat e fundit. Në këtë kapitull trajtohen edhe format e integritetit të të dhënave, ku si rezultat i përparimeve në teknologjinë e informacionit ka ndikuar aplikimi i sistemeve DWH, ku edhe bëhet përshkrimi teorik për vlerat që sjell ndërtimi i sistemeve të integruara.

Kapitulli V: Trajton ndërtimin e proceseve ETL, në fakt, kjo konsiderohet njëra nga detyrat më esenciale në ndërtimin e një sistemi DWH. Talend open studio Big Data është në vetvete zgjidhje software-ike, e cila është përgjegjëse për tërheqjen e të dhënave nga burime të ndryshme për pastrimin e tyre, përshtatjen dhe ngarkimin në databazën e centralizuar DWH. Në këtë kapitull trajtohet edhe ndërtimi i DWH-së që sinkronizohet në kohë reale, e që ka synim zvogëlimin e kohës së replikimit të të dhënave, pra jepen shpjegime teorike të bazuara fillimisht në arkitekturën e DWH-së në kohë reale, por trajtohet ETL si përgjegjës për tërheqjen, transformimin dhe ngarkimin e DWH-së.

Kapitulli VI: Trajton zhvillimin dhe aplikimin e pajisjeve mobile, duke pasur synim gjuhën programuese Java në platformën Android. Me këtë synohet krijimi i aplikacioneve të pajisjeve mobile, duke pasuar të dhënat nga databaza ORACLE, SQLite në aplikimet mobile, siç është rasti i fatkeqësive natyrore.

KAPITULLI II: METODOLOGJIA DHE ARKITEKTURA E DATAWAREHOUSE (DWH)

2.1. Hyrje

Viteve të fundit rëndësia e analizës së të dhënave është rritur ndjeshëm, ndërsa nga organizatat, në të gjithë sektorët, kërkohet të përmirësojnë proceset e tyre të vendimmarrjes, me qëllim që të ruajnë avantazhin e tyre konkurrues. Sistemet tradicionale të databazave nuk i plotësojnë kërkesat e analizës së të dhënave. Ato mbështetin operacionet ditore të një organizate dhe shqetësimi i tyre kryesor është të sigurojnë qasje të shpejtë të të dhënave në prani të shumë përdoruesve, e cila e bën të nevojshme përpunimin e transaksioneve, kontrollin e konkurrencës dhe teknikat e rikuperimit që garantojnë qëndrueshmërinë e të dhënave. Këto sisteme njihen si databaza operacionale ose sisteme të përpunimit të transaksioneve online (OLTP). Databazat tipike operacionale përmbajnë të dhëna të detajuara, nuk përfshijnë të dhëna historike dhe, meqenëse zakonisht janë shumë të normalizuara, performojnë dobët kur ekzekutojnë kërkesa komplekse që duhet të bashkojnë shumë tabela relative së bashku ose të përmbledhin volume të mëdha të dhënash. Për më tepër, kur përdoruesit duan të analizojnë sjelljen e një organizate si një e tërë, duhet të integrohen të dhënat nga sisteme të ndryshme operacionale. Kjo mund të jetë një detyrë e vështirë për t'u arritur shkaku i diferencave në përkufizimin e të dhënave dhe përmbajtjen.

Janë propozuar DWH me qëllim që t'i përgjigjen më së miri kërkesave në rritje të përdoruesve të vendimmarrjes. Një DWH është një koleksion i të dhënave të orientuara nga subjekti, të integruara, të pandryshueshme dhe të ndryshueshme në kohë për të mbështetur vendimet e menaxhimit (Malinoëski&Zim'anyi, 2008).

2.2. Koncepti datawarehouse (DWH)

Termi datawarehouse (DWH) është shpikur nga W. H. Inmon: "Datawarehouse është një grumbullim të dhënash i orientuar nga subjekti, i integruar, i varur nga koha dhe i pandryshueshëm në mbështetje të procesit të marrjes së mendimeve nga menaxhimi".

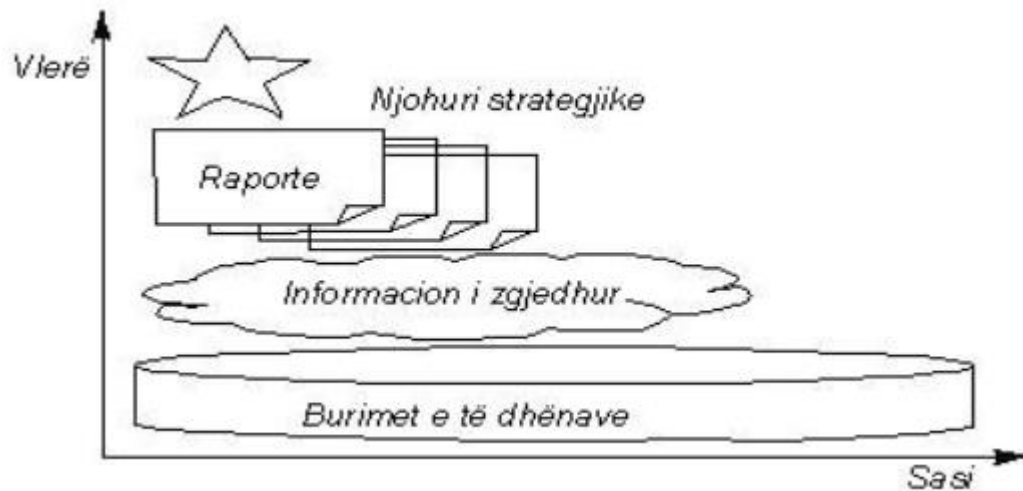
Në kontrast me aplikimet On-Line Transaction Processing (OLTP) (përpunimi i transaksioneve-On-Line), mbështetja e vendimit bën disa kërkesa të ndryshme në teknologjinë e databazës. Datawarehouse-t (DWH) sigurojnë ruajtje, funksionalitet dhe shpejtësi për kërkesat që janë përtej kapacitetit të databazave OLTP. Ato përmbajnë të dhëna historike, të përmbledhura dhe të konsoliduara mundësisht në periudha të gjata kohe. Përmasat e tyre mund të jenë nga qindra gigabajt deri në qindra terabajt. Ekziston një nevojë shumë e madhe për t'u siguruar vendimmarrësve në të gjitha nivelet e menaxhimit informacione në nivelin e detajeve të dëshiruara, për të ndihmuar vendimmarrjen e tyre. Përveç kryerjes së aktiviteteve të zakonshme të paracaktuara të riparimit, një numër përdoruesish paralel po paraqesin kërkesa 'ad hoc', komplekse. Këto kërkesa kërkojnë akses në miliona regjistrime dhe shkaktojnë veprime të shumta skanimit, bashkimi dhe grumbullimi në DWH (Stolba, 2007).

2.3. Historiku i sistemeve në mbështetje të vendimmarrjes (DSS)

Deri në vitet 1980 kompanitë ruanin vetëm informacionin e sistemeve Online Transaction Processing (OLTP) në databazat e tyre. Për çdo kërkesë raportimi apo analize mbi të dhënat duhej të nxirrej informacioni nga databazat operationale, nëpërmjet përmbledhjes së tij, nga

raportet që ndërtoheshin mbi të dhënat burim, si figura 2.1, më poshtë. Sipas kësaj mënyre, sa më shumë shtohet informacioni aq më të pakta bëheshin njohuritë.

Figura: 2.1 Vlera e informacionit si funksion i sasisë së të dhënave (Golfarelli&Rizzi, 2009).



Rritja e vazhdueshme e këtyre databazave kërkonte analiza që nevojiteshin për një vendimmarrje, prandaj lindi nevoja e krijimit të sistemeve të ndërtuara për mbështetjen e vendimmarrjeve (DSS). Sistemet e mbështetjes së vendimeve kanë ekzistuar shumë më përpara konceptimit të DWH-së, shpesh të iniciuara nga departamentet e marketingut, shitjeve ose financave, ku secili grup i merrte të dhënat që kërkonte nga sistemet e vjetruara dhe i vendoste në baza të vogla të dhënash në serverët e tyre lokalë.

Këto zgjidhje zbatoheshin në mënyrë të pavarur dhe pa ndihmën e shërbimeve të informacionit. Si rrjedhim, rezultatet kishin prirjen të ishin jokonsistente, veçanërisht kur krahasoheshin me rezultatet e marra nga organizata simotra me sisteme të ngjashme, por individuale në vendime.

Koncepti i sistemeve DSS përfshin dy fusha kërkimesh:

- Studime teorike të proceseve të vendimmarrjes në një organizatë;

- Studime teknike të ndërlidhjes së sistemeve IT për të ndihmuar vendimet.

Nga pikëpamja e arkitekturës një sistem DSS përfshin një sistem menaxhimi të informacionit, të bazuar në modele që lidhet me njohuritë dhe, gjithashtu, shfaqet me një ndërfaqe grafike interaktive.

Sistemet DWH kanë menaxhuar sistemet DSS prej viteve 1990. Ato duhet të nxjerrin informacionin e duhur nga gjithë ai informacion sasi i ruajtur në disa sisteme heterogjene. Në këtë mënyrë mund të gjenerohen shumë analiza komplekse pa ndikuar në performancën e sistemeve OLTP (Juliana, 2014).

2.4. Nevoja për t'i veçuar sistemet operacionale dhe informative

Sistemi operacional është një sistem që përdoret për të drejtuar një biznes në kohë reale, në bazë të të dhënave aktuale. Shembuj të sistemeve operacionale janë përpunimi i porositë të shitjeve, sistemet e rezervimit dhe sistemet e regjistrimit të pacientëve. Sistemet operacionale duhet të përpunojnë volume të mëdha të leximit/shkrimit relativisht të thjeshtë të transaksioneve dhe të sigurojnë përgjigje të shpejtë. Sistemet operacionale quhen, gjithashtu, sisteme regjistrimi. Sistemet informative janë ndërtuar për të ndihmuar marrjen e vendimeve në bazë të të dhënave historike dhe të parashikimit. Ato janë ndërtuar, po ashtu, për kërkesa komplekse ose aplikacionet e gjurmimit të të dhënave. Shembuj të sistemeve informative janë sistemet për analizimin e trendit të shitjeve, segmentimit të klientëve dhe planifikimin e burimeve njerëzore.

Diferencat kyçe midis sistemeve operacionale dhe informative janë treguar në tabelën 2.1a. Këto dy lloje përpunimesh kanë karakteristika shumë të ndryshme në pothuajse çdo kategori krahasimi. Në veçanti, ato kanë komunitete përdoruesish krejt të ndryshme.

Tabela 2.1 Krahasimi i sistemeve operacionale dhe i sistemeve informative (Khan, 2012).

Karakteristikat	Sistemet operacionale	Sistemet informative
Qëllimi kryesor	Drejtuar biznesit në një bazë të tanishme	Mbështetur vendimmarrjen menaxheriale
Lloji i të dhënave	Përfaqësimi i fushës së biznesit	Shënime me bazë të të dhënave historike
Përdoruesit kryesor	Nëpunësit, shitësit, administratorët	Menaxherët, ekzekutivët, analistët e biznesit
Fushat e përdorimit	Të ngushta, të planifikuara dhe të reja	Të gjerë, ad hoc, pyetje dhe analiza komplekse
Qëllimi i projektimit	Performanca: disponueshmëria	Lehtësia e qasjes fleksibile në përdorim
Vëllimi	Update-ime të shumta e të vazhdueshme në rreshtat ose në disa tabela	Update-ime periodike që kërkojnë shumë ose të gjithë rreshtat

Sistemet operacionale përdoren nga shitësit, administratorët, agjentët e shitjes dhe të tjerë që duhet të përpunojnë transaksione biznesi. Sistemet informative përdoren nga menaxherët, ekzekutivët, analistët e biznesit dhe, gjithnjë e më tepër, nga klientët, të cilët kërkojnë informacione statusi ose të cilët janë marrësit e vendimeve.

Nevoja për t'i veçuar sistemet operacionale dhe informative bazohen në tre faktorë kryesorë:

1. Një DWH centralizon të dhënat që janë shpërndarë përmes sistemeve operacionale të ndryshme dhe i vë në dispozicion për aplikacionet që ndihmojnë vendimmarrjen.

2. Një DWH, i ndërtuar mirë, u shton vlerë të dhënave, duke përmirësuar cilësinë dhe konsistencën e tyre.

3. Një DWH i veçantë eliminon shumë mosmarrëveshjet për burimet që rezultojnë kur aplikacionet informative përzihen me përpunimet operacionale (Hoffer&Ramesh, 2011).

2.5. Metodologjia e zhvillimit të datawarehouse (DWH)

Bill Inmon dhe Ralph Kimball janë dy liderët kryesorë të metodologjisë brenda fushës DWH, të cilët kanë nisur rritjen e shquar të zhvillimit DWH në dhjetëvjeçarin e fundit. Gjatë 10-15 vjetëve të fundit, metodologjitë e tyre kanë krijuar një qendër solide datawarehouse (DWH) të ngulitura me sukses.

Sipas **Inmonit** *datawarehouse (DWH)* është ndërtuar në një model përsëritës, nga DWH të ndërmarrjes në databazat e departamenteve, duke përdorur një mjet databazë relativ të zakonshëm për t'i shërbyer shumicës së nevojave për marrjen e vendimeve. Inmoni thekson tri premisat bazë për ndërtimin e datawarehouse-ve.

- E para, normalizimi i të dhënave dhe përkufizimi i metadatave për modelimin e të dhënave dhe proceseve, siç është e zakonshme në zhvillimin e sistemit të menaxhimit të databazave.
- E dyta, ndihmohen vendimet e përdoruesve finalë, duke përdorur kërkesa arbitrare në databazë.

- E treta, pastrimi dhe unifikimi i të dhënave në fazën e përgatitjes së të dhënave (skenimi i të dhënave) për transformimin e të dhënave operacionale në të dhëna për qëllime “informative”.

Zhvillimi i DWH-së së synuar përfshin një numër detyrash, duke filluar me një analizë të madhësisë/copëzimit, e cila përfshin tre nivele copëzimi: atomike, departamentesh dhe DWH individuale.

Sipas **Kimballit** përdoret një metodë e nxitur nga kërkesa, e kjo fillon me planifikimin e projektit për vlerësimin e gatishmërisë, studimit dhe justifikimit.

Qasja e Kimballit është e tipit nga poshtë-lart, e cila e vë analizën e kërkesave të biznesit në krye të fazave të zhvillimit të saj. Arkitektura e Kimballit formon DWH të ndërmarrjes me anë të arkitekturës bus, që grumbullon data marts në bazë të dimensioneve dhe fakteve të vërtetuara. Secili data mart përfaqëson një proces biznesi në organizatë, me anë të modelimit dimensional (p.sh., skema yll). Modelet e ndërtimit dimensional përfshijnë një proces me katër hapa:

- Zgjedhja e procesit të biznesit, duke filluar me procesin me ndikimin më të madh;
- Deklarimi i njësive, i cili shërben për të vendosur se çfarë niveli detajesh do të përmbajë DWH;
- Zgjedhja e dimensioneve;
- Identifikimi i fakteve.

Pastaj, zhvillohet një ndërtim fizik, duke u kushtuar vëmendje strategjive të përmbledhjes dhe indeksimit. Për qëllime funksionaliteti, krijohet një strukturë analitike e qëndrueshme,

që përmbush kërkesat analitike të përdoruesit. (Tabela e cila i thjeshton diferencat kryesore midis qasjeve të Inmonit dhe Kimballit).

Tabela 2.2 Qasja sipas Inmonit dhe Kimballit

	Inmon	Kimball
Qasja sipas emrit	Waterfall, CIE, Hub and Spoke	Data mart, arkitekturës bus dhe dimensioneve
Kërkesat e biznesit	Top down	Bottom Up
Modeli DWH	Modeli ER	Modeli shumëdimensional
Qasja e zhvillimit	Qasja operative për zhvillimin e DB të departamenteve nga EDW	Grumbullon data mart në bazë të dimensioneve dhe fakteve
Pjesëmarrja kryesore	Profesionistë të IT-sës	Përdoruesit fundorë

Ndërkohë, rritja e vazhdueshme në adoptimet DWH-së ka rritur ndërgjegjësimin e studiuesve dhe zhvilluesve për të propozuar dhe zhvilluar metodologji të reja që komprometohen dhe shkëmbehen midis Inmonit dhe Kimballit ose përshtatin teknika a arkitektura të reja.

2.6. Arkitektura e sistemeve të DWH-së

Koncepte të shumta arkitekturore janë prezantuar dhe diskutuar me kalimin e kohës, të cilat nxjerrin një numër idesh dhe pikëpamjesh. Janë propozuar arkitektura DWH konceptuale në mënyra të ndryshme. Ato klasifikohen, zakonisht, në tre grupe, sipas numrit të niveleve, nëse është arkitekturë me një nivel, dy nivele a tre nivele.

- **Në një arkitekturë me një nivel**, të quajtur zakonisht si magazinë të dhënash virtuale, çdo pjesë të dhënash ruhet njëherë, ndërsa një nivel “i mesëm” vepron si ndërfaqe midis përdoruesve dhe databazave operacionale pa asnjë DWH aktuale, por stimulohet duke përdorur shikimet. Ky lloj arkitekture karakterizohet nga zhvillimi i shpejtë me kosto të reduktuar dhe krijon nevojën që aktivitetet e planifikimit (p.sh., identifikimi i të dhënave burim, transformimi i të dhënave etj.) të bëhen për çdo kërkesë. Gjithashtu, ka një mungesë të historizimit të të dhënave dhe një kohë aksesi mjaft të paparashikueshme për përdoruesin final.
- Nga ana tjetër, **arkitekturat me dy nivele** veçojnë të dhënat burim (niveli i parë) nga të dhënat e nxjerra ose niveli i dytë i arkitekturës. Të dhënat e nxjerra mund të jenë goftë një kopje e thjeshtë e të dhënave burim ose të marra prej tyre nga ndonjë proces nxjerrjeje/përmbledhjeje. Ky lloj është praktik, veçanërisht kur burimet operacionale janë homogjene. Në kontrast me këtë, ka dyfishim të konsiderueshëm të dhënash, pasi që çdo aplikim i ndihmës për vendimmarrje ka kopjen e vet të të dhënave të nxjerra zakonisht.
- **Arkitektura me tre nivele** përbëhet nga niveli i të dhënave operacionale, niveli i të dhënave të bashkërenduara dhe niveli që mbështet të dhënat e vendimmarrjes (Arwa, 2011).

Arkitekturat standarde, që përdoren në ndërtimin e sistemeve DWH, së bashku me shtresat bazë që secila arkitekturë ka në përmbajtjen e saj, paraqiten në tabelën e mëposhtme.

Tabela 2.3 Elementët e arkitekturave për ndërtimin e sistemeve DWH (Juliana, 2014).

	Shtresa e burimeve	Shtresa Stage e rakordimit	Shtresa e DWH/datamart
Arkitektura me një shtresë	X		
Arkitektura me dy shtresa	X		X
Arkitektura me tri shtresa	X	X	X

Nga tabela e mësipërme shohim:

1. **Shtresa e burimit** përfaqëson burimet e informacioni, të cilat mund të jenë databaza relacionale të kompanisë, fajllave të tipave të ndryshëm ose të dhëna të marra nga një aplikacion i jashtëm.
2. **Shtresa “Stage”** - të dhënat e marra nga burimet e ndryshme duhet të pastrohen, transformohen dhe përshtaten në mënyrë të tillë që prej tyre të ndërtohet një skemë e vetme dhe e përbashkët të dhënash. Kjo shtresë krijohet në rastin kur proceset ETL i ruajnë informacionet në databaza fizike.
3. **Shtresa DWH/data mart** - informacioni që mblidhet nga të gjitha burimet ngarkohet në një sistem të vetëm të centralizuar apo data mart (Juliana, 2014).

Nxjerrja e të dhënave për mbështetjen e vendimmarrjes kryhet në dy faza:

- Faza e parë, integrimi i të dhënave operacionale;

- Faza e dytë, nxjerrja e të dhënave të mbështetjes së vendimmarrjes nga burimet e integruara.

Sen sugjeron llojet arkitekturore DWH të zakonshme:

- Lloji i parë është data mart i pavarur, i cili sugjeron data mart të ndryshëm për funksione të ndryshme biznesi. Megjithëse kjo qasje ndan të njëjtën ETL për burimet e zakonshme të të dhënave, nuk jep asnjë version të së vërtetës, e cila e bën të vështirë të analizohen të dhënat në data mart-e.
- Lloji i dytë është DWH e centralizuar me data mart të pavarur (Hub&Spoke). Kjo arkitekturë përdor shikimin e sipërmarrjes të të dhënave; zhvillohet në mënyrë përsëritëse, ku një fushë subjekti ndjek tjetrën. Data mart-ët e pavarur zhvillohen për qëllime departamenti, funksionale ose speciale; këta data mart marrin të dhëna nga magazina.
- Lloji i tretë është DWH e centralizuar pa data mart të pavarur. Kjo arkitekturë është e ngjashme me atë të mësipërmen, por pa data mart të pavarur.
- Lloji i katërt është arkitektura DWH virtuale, e cila nuk përfshin ndërtimin DWH tipik; në vend të saj përdor një “program urë” për të përdorur të dhënat operacionale. Kjo arkitekturë rekomandohet për firmat që nuk duan të rindërtojnë.
- Lloji i fundit është arkitektura bus e data mart-ëve me data mart dimensional të lidhur.

Për çdo proces të caktuar biznesi ndërtohet një data mart me përdorimin e dimensioneve të konfirmuara për të siguruar data mart të integruar në mënyrë logjike dhe një pamje ndërmarrje e të dhënave (Arwa, 2011).

2.7. Konkluzion

Një sistem i DWH-së ofron qasje të efektshme në informacionet e përmbledhura të cilësisë së lartë, që ndihmojnë menaxhimin për vendimmarrje më të mirë. Një DWH lejon që menaxhimi të njohë tendencat kyçe dhe të parashikojë rezultatet e pritura. Gjithashtu, lejon menaxhimin të kuptojë shkaqet dhe pasojat e asaj që ka ndodhur tashmë, e cila në këmbim i ndihmon organizatat të jenë më konkurruese dhe të përmbushin më shpejt kërkesat e tregut. Sipas Benmecchiche dhe Chouinard, dobia më e madhe e mundshme e një DWH-je ndodh kur i jep menaxhimit një botëkuptim logjik të ngjarjeve të tyre, e cila lehtëson përgjigjet e vendimmarrësve të rindërtojnë një proces menaxhimi të emergjencave, të një organizate dhe të mbështetin objektivat strategjike të QOAME-së për dobinë e tyre.

KAPITULLI III :ANALIZA E DATA MART-ëve PËR KRIJIMIN E QENDRËS SË EMERGJENCAVE NË KOSOVË

3.1 Hyrje

Agjencitë për menaxhimin e emergjencave dhe të fatkeqësive natyrore e kanë të domosdoshme të grumbullojnë dhe analizojnë të dhëna masive të fatkeqësive natyrore. Rekordet e tërmeteve, vërshimeve të detajuara apo edhe të stuhive e temperaturave ekstreme, janë fusha më e nxehtë për eksplorimin e njohurive në menaxhimin e emergjencave dhe të fatkeqësive natyrore.

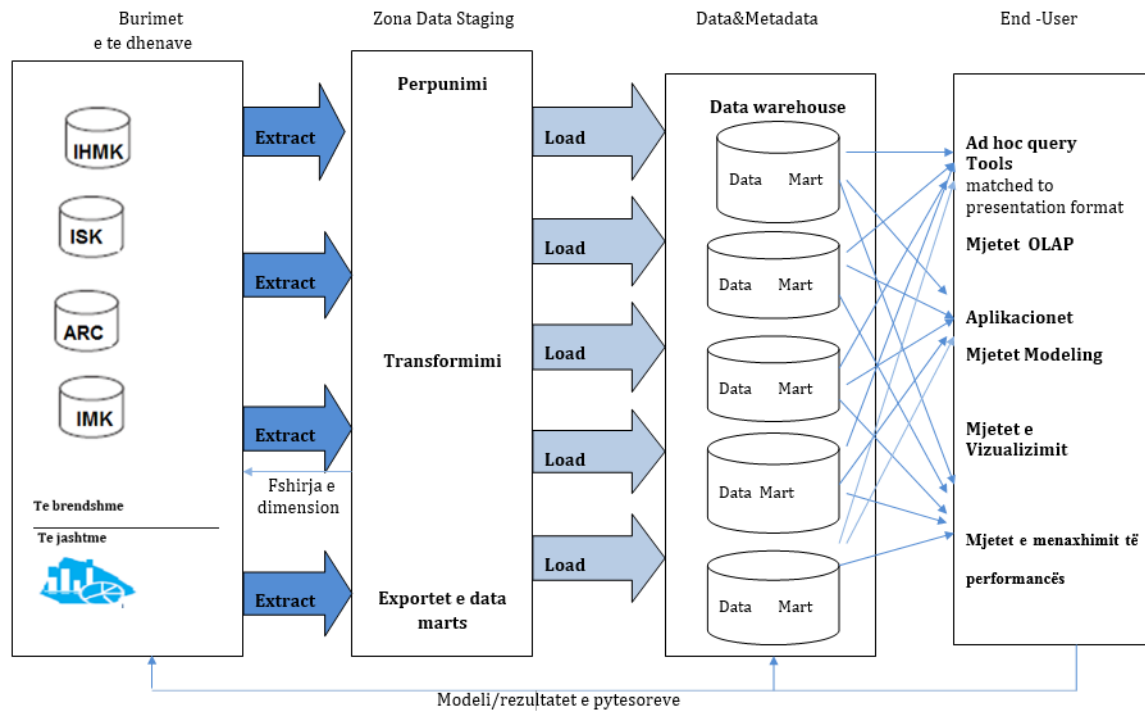
Analizat që duhet të gjenerohen prej tyre janë shumë të larmishme. Për këtë arsye, modeli i një data mart-i për këto të dhëna është shumë interesant për t'u zhvilluar. Për ndërtimin e tij është i rëndësishëm të shfrytëzohet një model që është shumë i lehtë për t'u kuptuar edhe nga përdorues të thjeshtë, sepse ata do të luajnë një rol aktiv në gjenerimin e rezultateve. Ky kapitull përmend të gjithë hapat, në detaje, që ndjekim në ndërtimin e një data mart-i mbi këto të dhëna.

Në të përmendën pikat dhe problemet kryesore që duhen të zgjidhen gjatë ndërtimit të modeleve. Për të përfituar më shumë fleksibilitet, në lidhje me analizat mbi informacionin e ruajtur, përdorën mjetet OLAP të teknologjisë BI.

3.1.1 Arkitektura e data mart-ëve

Arkitektura data-mart, e pavarur për një DWH, është treguar në figurën 3.1.

Figura 3.1 Arkitektura data-mart e pavarur për një DWH (Hoffer&Ramesh, 2011).



Ndërtimi i kësaj arkitekture kërkon katër hapa bazë. Duke lëvizur nga e majta në të djathtë, në figurën 3.1:

1. Të dhënat nxirren nga skedat e sistemit dhe databazat, burime të ndryshme të brendshme e të jashtme. Në një organizatë të madhe ato mund të jenë me dhjetëra ose qindra skeda dhe databaza të tilla.
2. Të dhënat nga sistemet e ndryshme të burimeve transformohen dhe integrohen para se të ngarkohen në data mart-e. Transaksionet mund të dërgohen në sistemet burim për t'i

korrigjuar gabimet e gjetura në skenimin e të dhënave. DWH konsiderohet si koleksion i data mart-ëve.

3. DWH është një set databazash të ndryshme fizikisht, të organizuara për ndihmë vendimmarrjeje. Përmban të dhëna të detajuara dhe të përmbledhura së bashku.

4. Përdoruesit përdorin DWH përmes një sërë gjuhësh kërkimi dhe mjetes analitike. Rezultatet (p.sh., parashikimet, planifikimet) mund të kalohen nga një DWH dhe databaza operacionale.

Më pastaj, do të flasim për proceset më të rëndësishme të nxjerrjes, transformimit dhe ngarkimit (ETL) të të dhënave nga sistemet burim në DWH me më shumë hollësi në këtë kapitull. Do të japim më poshtë, gjithashtu, mjete të ndryshme prezantimi të përdoruesve finalë. Nxjerrja dhe ngarkimi ndodhin rregullisht: disa herë çdo ditë, javë ose muaj. Prandaj, DWH shpesh nuk ka ose nuk ka nevojë të ketë të dhëna aktuale.

Të mos harrojmë se DWH nuk mbështet (direkt) përpunimin operacional të transaksioneve, megjithëse mund të përmbajë të dhëna transaksionesh (më shpesh përmbledhje të transaksioneve dhe pamje çasti të variableve të statusit, të tilla si bilancet e llogarive dhe nivelet e inventarit). Për shumicën e aplikacioneve të DWH-së, përdoruesit nuk kërkojnë reagim në një transaksion të veçantë, por më tepër tendencat dhe strukturat në gjendjen e organizatës nëpër një nëngrup të madh të DWH-së. Si minimum, pesë tremujorë fiskalë të dhënash mbahen në një DWH, në mënyrë që të shihen, të paktën, tendencat dhe strukturat vjetore. Të dhënat më të vjetra mund të pastrohen ose të arkivohen.

Do të shohim më vonë se një arkitekturë e avancuar e DWH-së, DWH në kohë reale, bazohet në një supozim të ndryshëm rreth nevojës për të dhëna aktuale. Ndryshe nga shumë parime të diskutuara deri në këtë kapitull, qasja e data mart-ëve të pavarur nuk krijon një DWH. Por,

në vend të kësaj, kjo qasje krijon shumë data mart të veçantë, secili i bazuar në DWH, jo teknologji databazë të përpunimit të transaksioneve.

Një data mart është një DWH, i cili është i kufizuar në shtrirje, i personalizuar për aplikacionet e vendimmarrjes së një grupi të caktuar përdoruesish final.

Përmbajtjet e saj merren qoftë nga proceset ETL të pavarura, siç tregohet në figurën 3.1 për një data mart të pavarur ose merren nga DWH, të cilin do ta diskutojmë në dy pjesët më tej. Është e mundur që çdo data mart të ndërtohet duke përdorur mjete të ndryshme, për shembull, një data mart financiar mund të ndërtohet duke përdorur një mjet multidimensional si Hyperion's Essbase dhe një data mart i shitjeve mund të ndërtohet mbi një platformë DWH, me qëllim më të përgjithshëm, ta zëmë, Teradata, duke përdorur MicroStrategy dhe mjete të tjera për raportim, kërkesa dhe shikim të dhënash.

Do të tregojmë një krahasim të arkitekturave të ndryshme të DWH-së më vonë, por mund të shihet një karakteristikë të dukshme të strategjisë data mart të pavarur: kompleksitet për përdoruesit finalë, kur duhet t'i përdorin të dhënat në data (të evidencuara nga vijat e përshkuara që lidhin gjithë data mart-ët me mjetet e prezantimit të përdoruesit final).

Ky kompleksitet vjen jo vetëm nga detyrimi për t'i hapur të dhënat nga databaza të veçanta data mart-ësh, por gjithashtu nga mundësia e një gjenerate të re sistemesh të dhënash kontradiktore data mart-ësh.

Nëse ka një set metadatash në gjithë data mart-ët dhe nëse të dhënat janë bërë të qëndrueshme në data mart, përmes aktiviteteve në fushën e skenimit të të dhënave, p. sh., nga ato çka quhen "dimensione konform" në kutinë e fushës së skenimit të të dhënave në figurën 3.1), atëherë kompleksiteti për përdoruesit reduktohet. Jo kaq i dukshëm në figurën 3.1 është kompleksiteti

për proceset ETL, sepse transformimet dhe ngarkesat e veçanta duhet të ndërtohen për secilin data mart të pavarur.

Sidoqoftë, projektimi i mjedisit të DWH në grupet e ndryshme të objektivave afatshkurtër do të thotë që humbisni elasticitetin për atë afatgjatë dhe aftësinë për të reaguar ndaj ndryshimeve . Dhe, të qenit në gjendje për të reaguar ndaj ndryshimit është kritik për të ndihmuar vendimmarrjen. Nga ana organizative dhe politike mund të jetë më e lehtë të kesh DWH të dhënash të veçanta, të vogla, në vend se t'i mbledhësh të gjitha palët organizative të bien dakord për pikëpamjen e një organizate në një DWH qendrore të dhënash. Gjithashtu, disa teknologji të DWH-së kanë kufizime teknike për madhësinë e DWH-së, që mund të mbështetin, e të cilën do ta quajmë më vonë çështje shkallëzimi.

Prandaj, teknologjia mund të diktojë në një arkitekturë DWH nëse e mbyllni më parë veten në një grup teknologjish të DWH-së, para se t'i kuptoni kërkesat tuaja për DWH. I diskutojmë avantazhet edhe disavantazhet e arkitekturës së data mart-it të pavarur me kulmin e vet, duke konkurruar arkitekturën në pjesën vijuese.

3.1.2 Arkitektura e data mart-it të pavarur dhe e ruajtjes së të dhënave operacionale

Arkitektura e data mart-it të pavarur në figurën 3.1, ka kufizime të ndryshme të rëndësishme:

1. Është zhvilluar një proces i veçantë ETL, për secilin data mart mund të japë të dhëna të tepërta dhe mundime përpunimi të kushtueshme.
2. Data mart-ët mund të mos jenë të pajtueshëm me njëri- tjetrën shkaku se shpesh janë zhvilluar me teknologji të ndryshme dhe, prandaj, mund të mos japin një shikim të qartë të të dhënave në lidhje me subjektet e rëndësishme.

3. Nuk ka asnjë aftësi për të gërmuar në më shumë detaje ose fakte përkatëse në data mart të tjerë ose një depozitë të dhënash të ndarë, prandaj analiza është e kufizuar ose më e shumta shumë e vështirë (p.sh., bërja e bashkimeve nëpër platforma të veçanta për data mart të ndryshëm). Në thelb, lidhja e të dhënave nëpër data mart është një detyrë e kryer nga përdoruesit jashtë DWH-së.

4. Kostot e shkallëzimit janë të tepërta, meqë çdo aplikacion i ri që krijon një data mart të veçantë, përsërit të gjithë hapat e nxjerrjes dhe ngarkimit. Zakonisht, sistemet operacionale kanë dritare kohore të kufizuara për nxjerrjen e të dhënave në grup, prandaj në njëfarë pike ngarkimi në sistemet e operacioneve mund të nënkuptojnë që teknologjia e re është e nevojshme, me kosto të tjera.

5. Nëse ka një përpjekje për t'i bërë data mart-ët më të pajtueshëm, kostoja për ta bërë këtë është shumë e lartë.

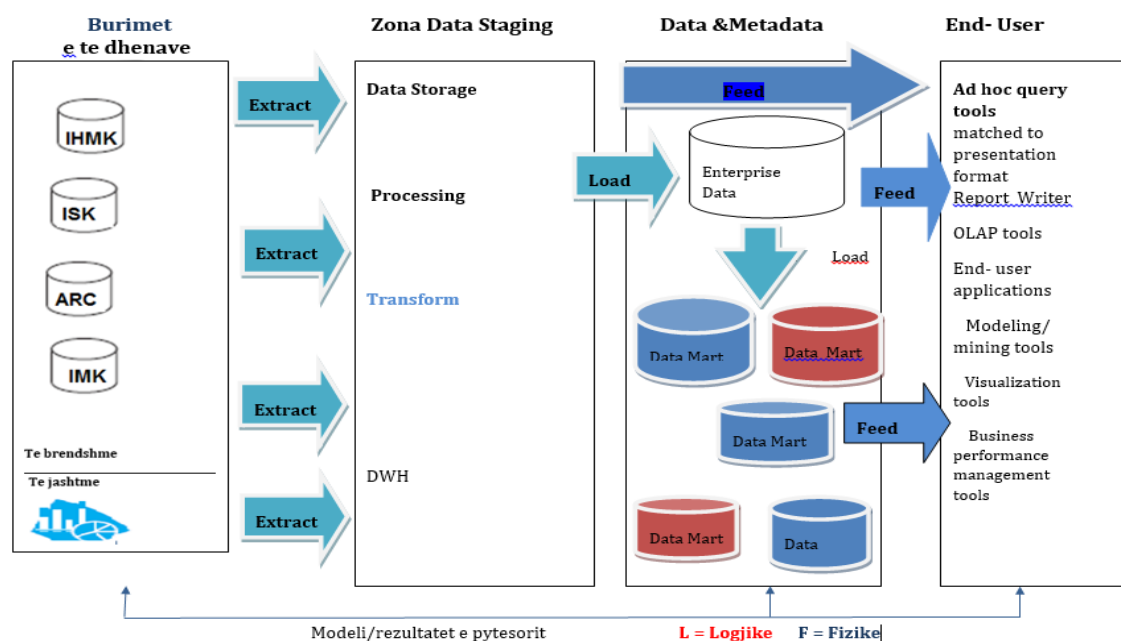
Vlera e data mart-ëve të pavarur është debatuar mjaft. Kimball (1997) mbështet fort zhvillimin e data mart-ëve të pavarur, si një strategji të zbatueshme për një zhvillim në faza të sistemeve për mbështetjen e vendimmarrjes. Armstrong (1997), Inmon (1997, 2000) dhe Marco (2003), tregojnë pesë ide të gabuara të përmendura më parë dhe shumë të tjera.

Ka dy debate në lidhje me vlerën reale të data mart-ëve të pavarur:

1. Një debat merret me natyrën e qasjes në faza në implementimin e mjedisit të DWH-së. Thelbi i këtij debati është nëse çdo data mart duhet ose nuk duhet të zhvillohet në një model poshtë-lart nga një nëngrup të dhënash mbarëndërmarrjesh për mbështetjen e vendimmarrjes.
2. Debati tjetër merret me arkitekturën e përshtatshme të databazës për përpunimin analitik. Ky debat përqendrohet në shkallën në të cilën duhet të normalizohet një databazë data mart.

Thelbi i këtyre dy debateve adresohet më tej në këtë kapitull. Një nga qasjet më të njohura për adresimin e kufizimeve të data mart-ëve të pavarur, të përmendur më sipër, është përdorimi i një qasjeje me tre nivele, e përfaqësuar nga arkitektura e data mart-ëve të pavarur dhe e ruajtjes së të dhënave operacionale (shih figurën 3.2).

Figura 3.2 Arkitektura trishtresore e data mart-it të pavarur (Hoffer & Ramesh, 2011).



Këtu niveli i ri është burimi (IHMK,IMK,ARC,ISK) i të dhënave operacionale dhe niveli i ruajtjes të të dhënave dhe metadatave është rikonfiguruar.

Kufizimi i parë dhe i dytë adresohet duke ngarkuar data mart-ët e pavarur nga DWH të ndërmarrjes (EDW), i cili është një DWH qendror, i integruar.

Është pika e kontrollit dhe "versioni i së vërtetës" i vetëm i vënë në dispozicion për përdoruesit finalë për aplikacionet e mbështetjes së vendimmarrjes. Data mart-ët e pavarur akoma kanë një qëllim për të siguruar një mjedis të thjeshtuar dhe me performancë të lartë, i

cili akordohet sipas nevojave të vendimmarrjes së grupeve të përdoruesve. Një data mart mund të jetë një databazë e veçantë fizike (data mart-ët e ndryshëm mund të jenë në platforma të ndryshme) ose mund të jetë një data mart logjik.

Një grup përdoruesish mund të hapin data mart-in e vet dhe kur nevojiten të dhënat e tjera, përdoruesit mund të hapin EDW. Teprica nëpër data mart-ët e pavarur është e planifikuar dhe të dhënat e tepërta janë të pajtueshme, sepse çdo data mart ngarkohet në një mënyrë të sinkronizuar nga një burim i përbashkët të dhënash (ose është një reflektim i DWH-së). Integrimi i të dhënave është përgjegjësia e stafit IT, që menaxhon DWH të ndërmarrjes; nuk është përgjegjësi e përdoruesve finalë të integrojnë të dhënat nëpër data mart të pavarur për çdo kërkesë ose aplikim.

Kufizimi i tretë adresohet duke siguruar një burim të integruar për të gjitha të dhënat operationale në një bazë të dhënash operationale. Një bazë të dhënash operationale (Operational Data Store) (ODS) është një databazë e integruar, e orientuar nga subjekti, e përditësueshme vazhdimisht, me vlerë aktuale (me histori të fundit), mbarë-organizate, e detajuar dhe e ndërtuar për t'u shërbyer përdoruesve operationalë, ndërsa bëjnë përpunimin e mbështetjes të vendimmarrjes (Imhoff, 1998; Inmon, 1998). Një ODS, zakonisht, është një databazë relative dhe e normalizuar si databazat në sistemet e regjistrimit, por është e akorduar për aplikimet e vendimmarrjes.

Për shembull, indekset dhe elementët e tjerë të ndërtimit të databazës relative janë akorduar për kërkesat që gjejnë grupe të mëdha të dhënash, në vend se për përpunimin e transaksioneve ose kërkimin individual dhe regjistrimet e lidhura direkt. Për shkak se është e ndryshueshme, aktuale dhe vetëm e dhënë e historisë të fundit, e njëjta kërkesë kundrejt një ODS-je me shumë mundësi do të jep rezultate të ndryshme në kohë të ndryshme. Një ODS, zakonisht,

nuk përmban histori "të thellë", ndërsa një EDWH mban, zakonisht, një histori me shumë shtresa të pamjeve të gjendjes së organizatës. Një ODS mund të përkrahet nga databaza e një aplikacioni ERP, por shkakut se shumica e organizatave nuk kanë vetëm një databazë ERP dhe nuk i bëjnë të gjitha veprimet kundrejt një ERP-je, një databazë ODS, zakonisht, është e ndryshme nga një databazë ERP.

ODS, gjithashtu, shërben si fushë skenimi për ngarkimin e të dhënave në EDWH. ODS mund t'i marrë të dhënat menjëherë ose me njëfarë vonese nga sistemet e regjistrimit, cilado që është më praktike dhe e pranueshme për kërkesat e vendimmarrjes që mbështet. Arkitektura e data mart-ëve të pavarur dhe e ruajtjes së të dhënave operacionale quhet, gjithashtu, fabrikë korporate informacioni (Corporate Information Factory) (CIF) (shih Imhoff, 1999). Konsiderohet të jetë një pamje gjithëpërfshirëse e të dhënave organizative në mbështetje të të gjitha kërkesave të përdoruesve për të dhënat.

Drejtuesit e ndryshëm në këtë fushë përkrahin qasje të ndryshme për datawarehouse. Ata që përkrahin qasjen e data mart-ëve të pavarur argumentojnë se kjo qasje ka dy dobi të mëdha:

1. Lejon që koncepti i Datawarehouse (DWH) të tregohet duke punuar në një seri projektesh të vogla;
2. Koha derisa fillojnë dobitë nga Datawarehouse (DWH) reduktohet, sepse organizata nuk vonohet derisa të gjitha të dhënat të centralizohen.

Shkencëtarët e CIF (Armstrong, 2000; Inmon, 1999) ngrenë çështje serioze me qasjen e pavarur, këto çështje përfshijnë pesë kufizime të data mart-ëve të pavarur, të përmendura më sipër. Inmon sugjeron se një avantazh i data mart-ëve fizikisht të veçantë, *të pavarur*, është se ata mund të akordohen sipas nevojave të secilit komunitet përdoruesish.

Në veçanti, ai sugjeron nevojën për një DWH eksplorimi, i cili është një version i veçantë i EDWH së optimizuar për gjurmimin e të dhënave dhe inteligjencën e biznesit, duke përdorur modelime statistikore, matematikore dhe mjete vizualizimi. Armstrong (2000) dhe të tjerë shkojnë më tej, duke argumentuar se dobitë e pretenduara nga shkencëtarët e data mart-ëve të pavarur janë, vërtet, në dobi të marrjes së një qasjeje në faza për zhvillimin e DWH-së.

Një qasje në faza mund të arrihet, gjithashtu, brenda strukturës CIF dhe lehtësohet nga arkitektura finale e DWH-së, që do ta shqyrtojmë në pjesën vijuese.

3.2 Shtrimi i problemit për QOAME

Që shtrimi i problemit të jetë më domethënës i ndajmë të dhënat, duke u bazuar në periudhë kohore, si: të dhëna historike, të dhëna momentale dhe të dhëna të së ardhmes. Në fakt, të dhënat e së ardhmes janë të dhëna të kalkuluar dhe për t'i kalkuluar ato duhen bërë analiza . Dhe, kjo, në fakt, paraqet arsyen kryesore të analizës së të dhënave. Të dhënat historike që posedon QOAME, janë një burim shumë i rëndësishëm dhe bazë për analizat numerike, nga edhe rrjedhin indikatorët e performancës, shifrat kyçe të QOAME-së. Në vijim paraqesim gjendjen e përgjithshme të Agjencisë për Menaxhim të Emergjencave (CIO.al, 2011):

a) Analizat mbi të dhënat e Institutit të Mjekësisë së Kosovës (IMK)

Sistemi informativ shëndetësor është, kryesisht, manual në letër dhe shumica e të dhënave për pacientë dhe trajtim regjistrohen në formularë letre, të cilët grumbullohen në institucionet gjegjëse dhe pastaj futen në formatin elektronik nga ana operatorëve në spitale dhe qendra kryesore të mjekësisë familjare. Në sektorin e kujdesit parësor

shëndetësor (KPSH), të dhënat nga Qendrat e mjekësisë familjare (QMF) dhe ambulantet, grumbullohen në formë të letrave nga Qendrat kryesore të mjekësisë familjare (QKMF), e pastaj futen në bazën elektronike nga ana e SISH-njësive dhe transferohen në departamentin e SISH-it, në kuadër të IKSHPK. Në shumicën e rasteve, ky transfer bëhet me anë të mjetit elektronik “flash memory”. Situata e njëjtë qëndron edhe në Qendrën Klinike Universitare të Kosovës (QKUK), ku të dhënat nga repartet dhe klinikat e ndryshme transferohen në njësitë e SISH-it, në formë të kopjeve të letrës.

Të dhënat për personelin dhe pacientët (të ashtuquajturat “statistikat e mesnatës”), kryesisht futen në programin Excel. Megjithatë, qysh nga viti 2002 të dhënat për aktivitetet, sëmundjet dhe vdekshmërinë në kujdesin parësor, dytësor dhe tretësor, janë duke u regjistruar gjithnjë e më tepër në bazën kryesore të të dhënave “Access-Master databaza”. Deri më sot Access-Master databaza implementohet në të gjitha QKMF-të, në 7 spitale (5 spitale regjionale dhe 2 spitale të qytetit) dhe në QKUK. Mirëpo, të dhënat raportohen me vonesë prej gjashtë muaj e më tepër. Analizat e të dhënave të njëjta japin rezultate të ndryshme dhe jo të gjitha institucionet shëndetësore raportojnë (MSH, 2014).

b) Analizat nga Instituti Sizmologjik i Kosovës (ISK)

Pjesa dërmuese e të dhënave nga Instituti Sizmologjik i Kosovës (ISK) ndodhen të arkivuara në trajtën e raporteve, tabelave, skicave e hartave. Shfrytëzimi i tyre bëhet me metoda tradicionale. Mjetet e komunikimit janë manuale, mungon një rrjedhë e vazhdueshme dhe e përditësuar për mbledhjen, vlerësimin dhe administrimin. Përdoruesit e kanë të pamundur të

sigurojnë në kohë reale informacion të freskët dhe të besueshëm për problemet për të cilat interesohen në fushën sizmike.

c) Analiza nga Instituti Hidrometeorologjik i Kosovës (IHMK)

Bazuar në analizat e gjeneruara të raportimit dhe në rrethanat që gjendet Instituti Hidrometeorologjik i Kosovës, ai ka grumbulluar të dhëna mjedisore nga institucionet monitoruese, kompanitë, operatorët, ndërmarrjet e ndryshme, publikimet, raportet dhe nga burime të tjera. Për t'i përmbushur kërkesat e raportit, të dhënat e grumbulluara janë përpunuar në informata kualitative mjedisore, që më pas janë prezantuar në raporte.

Këto të dhëna janë në formë teksti, tabelash, hartash dhe paraqitjesh grafike (Agjencia, 2011).

Cilësia e të dhënave dhe besueshmëria janë shumë të rëndësishme. Të dhënat e pasakta ose jo të plota të reshjeve janë më të dëmshme se mungesa e të dhënave.

Pa të dhëna asgjë nuk mund të parashikohet, të planifikohet ose të menaxhohet (Zyra e Kryeministrit, 2014).

Nga Instituti Hidrometeorologjik i Kosovës deri me tani nuk kemi një bazë të të dhënave që mund t'i referohemi në kohë reale dhe të kemi freskim të të dhënave nga Instituti Hidrometeorologjik.

3.3 Modelimi i të dhënave të përzgjedhura

Analizimi i kërkesave të mësipërme udhëheq procesin e modelimit të të dhënave për ndërtimin e një zgjidhjeje. Modelimi shumëdimensional u përzgjedh për t'u aplikuar në rastin tonë për tri arsye:

- Këto modele ofrojnë një thjeshtësi në përdorim edhe nga personat pa njohuri në IT;

- Plotëson kërkesa komplekse me performancë të lartë;

- Përshtatet shumë mirë me sasi të mëdha informacioni, siç është rasti ynë (Juliana, 2014).

Data mart-ët që do i ndërtojmë në vijim janë për të optimizuar performancën, për përdorimet e përcaktuara mirë dhe të parashikueshme, ndonjëherë qoftë edhe një ose me pak kërkesa.

Për shembull, rasti i menaxhimit të emergjencave mund të ketë një data mart për Institutin e Mjekësisë të Kosovës (IMK), një data mart për Agjencinë e Regjistrimit Civil (ARC), një data mart për Institutin Hidrometeorologjik të Kosovës (IHMK), e kështu me radhë, për të ndihmuar përpunimin analitik të raporteve (Hoffer&Ramesh, 2011).

Sipas të njëjtës mënyrë ndërtohen edhe për tipat e tjerë të rekordeve. Gjithashtu, duke u nisur nga përshkrimi i problemit dhe duke krahasuar elementët e dy metodikave kryesore për ndërtimin e zgjidhjes, u përzgjedh metodika Kimball, që bazohet mbi data mart-ët.

Elementët kriter të përzgjedhjes janë paraqitur në tabelën e mëposhtme.

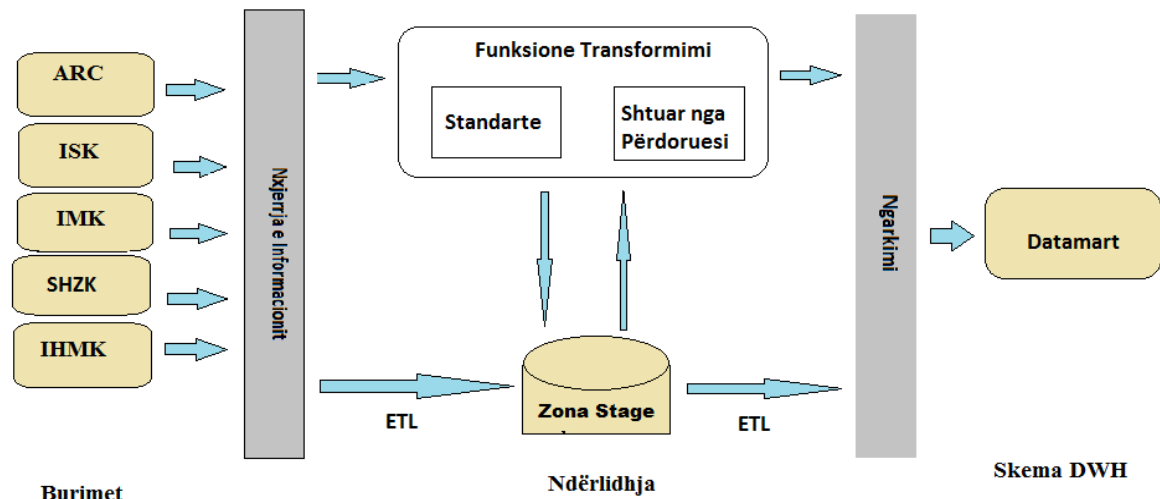
Tabela 3.1 Elementët kriter për përzgjedhjen e modelimit

Elementet	Kimball	Inmon
Nevoja	E menjëhershme	E zgjatur në kohë
Fokusi	Fusha specifike e biznesit	Gjithë sipërmarrja
Buxheti	Buxheti i vogël	Buxheti i madh
Përdoruesit	Grupe specifike të biznesit	Korporata

Duke u bazuar pikërisht në rekomandimet e Kimball-it u përzgjedh arkitektura me dy shtresa, për të përmbushur veçanërisht tiparin e ndarjes, ku të dhënat fillimisht do të ngarkohen në

një zonë Stage, ku edhe do të transformohen, për t'u ngarkuar më pas në skemën e data mart-it (Juliana, 2014).

Figura 3.3 Skema e arkitekturës së përzgjedhur (Ali, 2011:91).



Më pastaj, shqyrtojmë edhe njëërën nga teknikat kyçe, me të cilën janë ndërtuar dhe implementuar strukturat e të dhënave brenda DWH-së, e kjo ka të bëjë me modelimin dimensional.

3.4 Modelimi dimensional

Modelimi dimensional është një teknikë e ndërtimit logjik dhe bën strukturimin e të dhënave për përdoruesit e biznesit, meqë ka performancë kërkimi të shpejtë. Modelimi dimensional pranohet gjerësisht si qasje e preferuar për prezantimin DW/BI. Ekspertët kanë dalluar se të dhënat e prezantuara në mjetet e BI-së duhet të bazohen në thjeshtësi për t'i qëndruar çdo mundësie suksesi. Herë pas herë, organizatat IT, konsulentët, përdoruesit dhe shitësit, kanë zbritur në një strukturë të thjeshtë dimensionale, për t'iu përshtatur nevojës njerëzore bazë për thjeshtësi. Modelimi dimensional e ndanë botën në matje dhe kontekst. Matjet kapen nga

procedurat e biznesit të organizatës dhe sistemet e tyre burime operacionale mbështetëse.

Matjet, zakonisht, janë vlera numerike dhe i quajmë ato si fakte (Kimball&Becker, 2008).

Tabela e fakteve është e përbërë nga çelësat kryesorë me shumë pjesë, ku faktet numerike dhe shitesë janë faktet më të dobishme në tabelën e fakteve.

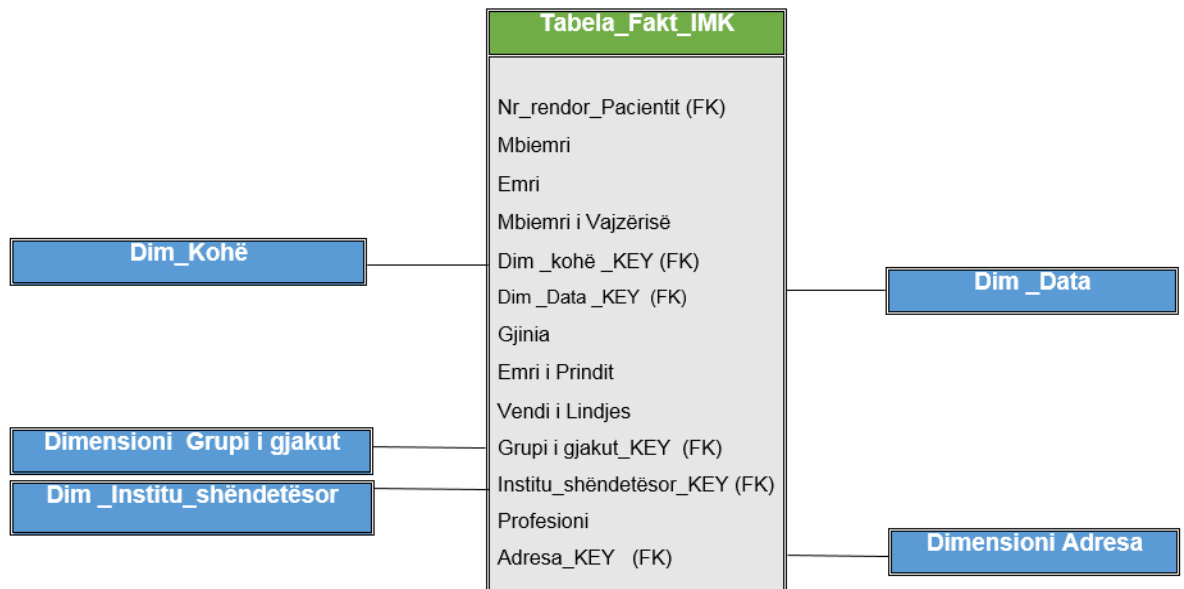
Secila tabelë faktesh tregon një numër tabelash më të vogla të quajtura dimensione, çelësi kryesor i të cilave i përket saktësisht një prej çelësave me shumë pjesë në tabelën e fakteve.

Një përshkrim i tillë është quajtur si modeli i skemës yll, i cili shërben fillimisht për një data mart. Teksa DWH rritet, mund të shtohen më shumë data mart, të cilët japin më shumë mundësi për të krijuar dimensione, të cilat janë të përbashkëta për më shumë se një data mart.

Këto dimensione të përbashkëta (DP) veprojnë si "ura të dhënash" midis tabelave të fakteve, duke lejuar kërkesa të cilat veprojnë që të dhënat nga faktet e shumëfishta të shkruhen. Shembuj tipikë të dimensioneve të përbashkëta përfshijnë datën dhe vendndodhjen.

Si rrjedhojë, një hap shumë i rëndësishëm në ndërtimin e modelit multidimensional është krijimi i një dimensionit të konfirmuar, me qëllim që DWH të funksionojë si një depozitë e vetme. Secila prej proceseve të menaxhimit të emergjencave, si në rastin e Institutit të Mjekësisë së Kosovës (IMK), mund të tregohet nga një model dimensional, i cili përbëhet nga një tabelë faktesh, e që përmban matjet numerike të rrethuara nga një sërë tabelash dimensionesh që përmbajnë kontekstin tekstual, siç ilustruhet në figurën në vijim.

Figura 3.4 Proceset e modelimit dimensional



Kjo strukturë shumë herë e ngjashme me yllin, term ky që datohet në ditët më të hershme të databazave relacionale. Modelet dimensionale të ruajtura në platformën e databazës relacionale, zakonisht, quhen “Skema yll”, ndërsa modelet dimensionale të ruajtura në strukturat e përpunimit analitik online multidimensional (OLAP) quhen kuba.

3.4.1 Modelimi i formës së tretë normale (3NF)

Modelimi 3NF është një teknikë ndërtimi që kërkon të eliminojë teprinë e të dhënave. Të dhënat ndahen në shumë entitete të dallueshme, secila prej tyre bëhet një tabelë në databazën relacionale. Ky normalizim është shumë i dobishëm për përpunimin e transaksioneve, pasi e bën ngarkimin dhe përditësimin e transaksionit thjesht dhe shpejt.

Modelimi dimensional dhe modelimi 3NF rezultojnë të dy në tabela fizike, të arsyetuara në një sistem databazë relacionale. Strukturat e tabelave që vijnë si rezultat ndryshojnë, por të gjitha kanë të njëjtat karakteristika databazë relacionale, pavarësisht nga teknika e modelimit.

Vlen të theksohet se konvertimi nga strukturat normale në modelet dimensionale është relativisht i drejtpërdrejtë.

- Hapi i parë është të përcaktohen relacionet “shumë-në-shumë” në modelin e normalizuar që përmban vlera numerike dhe shtuese metrike jo kyçe si tabela faktesh. Tabelat e fakteve, zakonisht, normalizohen në 3NF në një model dimensional, sepse konteksti përkatës hiqet në tabelat e dimensionit.
- Hapi i dytë është të denormalizohen tabelat e mbetura në tabela dimensionit me çelësa me një pjesë, të cilat lidhen direkt me tabelën e fakteve. Tabelat e dimensionit më shpesh u ngjajnë tabelave të formës normale të dytë me shumë përshkrues me kardinalitet të ulët (Kimball&Becker, 2008).

3.4.2 Tabelat e fakteve

Tabelat e fakteve ruajnë matjet e performancës, të gjeneruara nga aktivitetet ose evenimentet e institutit mjekësor të Kosovës . Termi fakt i referohet çdo matjeje performance, që, zakonisht, nuk dihet vlera e një fakti më përpara, sepse ajo është e ndryshueshme.

Vlerësimi i faktit ndodh në kohën e evenimentit të matjes, si, për shembull, kur merret një porosi, dërgohet një ngarkesë ose regjistrohet një problem shërbimi.

Pothuajse çdo fakt është numerik. Faktet më të dobishme janë numerike dhe shtuese. Shtimi është i rëndësishëm, sepse aplikimet BI rrallë marrin një rresht të vetëm tabele faktesh. Kërkesat, zakonisht, zgjedhin qindra ose mijëra rreshta faktesh njëherësh dhe gjëja e vetme e dobishme për t'u bërë me shumë rreshta është që ato të mblidhen.

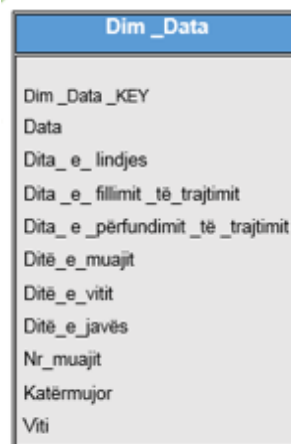
Tabelat e fakteve janë gjigante, me miliona ose miliarda rreshta, por janë efikase. Shkaku se tabelat e fakteve shpesh ruajnë më shumë se 90 për qind të të dhënave në një model

dimensional, duhet të jemi të kujdesshëm t'i minimizojmë të dhënat e tepërta në një tabelë faktesh. Gjithashtu, ne përpiqemi t'i ruajmë matjet që dalin nga institutit mjekësor në një tabelë të vetme faktesh, e cila është e ndarë në gjithë organizatën. Për shkak se të dhënat e matjeve janë të dhënat më voluminoze, e shmangim qëllimisht dyfishimin e metrikës të detajuar në tabelat e shumëfishta të fakteve në korporatë, ndërsa përshkruajmë më tej me arkitekturën bus të DWH të korporatës.

3.4.3 Tabelat e dimensioneve

Tabelat e dimensioneve janë shoqërues integral të një table faktesh. Tabelat e dimensioneve përmbajnë përshkrues tekstual të një rasti, siç është QOAME tabelë dimensionale, që lidhet me të dhënat e datave, siç ilustruhet në fig. 3.5.

Figura 3.5 Mostër e tabelës dimensionale

A diagram of a dimension table structure. It consists of a blue header box labeled 'Dim_Data' and a light gray body box containing a list of attributes. The attributes are: Dim_Data_KEY, Data, Dita_e lindjes, Dita_e fillimit_të_trajtimit, Dita_e përfundimit_të_trajtimit, Ditë_e_muajit, Ditë_e_vitit, Ditë_e_javës, Nr_muajit, Katërmujor, and Viti.

Dim_Data
Dim_Data_KEY
Data
Dita_e lindjes
Dita_e fillimit_të_trajtimit
Dita_e përfundimit_të_trajtimit
Ditë_e_muajit
Ditë_e_vitit
Ditë_e_javës
Nr_muajit
Katërmujor
Viti

Në një model dimensional të ndërtuar mirë, tabelat e dimensioneve kanë shumë kolona ose attribute. Atributet e dimensioneve shërbejnë si burimi kryesor i kufizimeve të kërkimit, grupeve dhe emërtimeve të raportit.

Në një kërkim ose kërkesë raporti atributet dallohen nga fjalët. Atributet e tabelës së dimensioneve luajnë një rol jetësor në DWH. Meqenëse janë burim virtualisht i të gjitha kufizimeve interesante dhe emërtimeve të raporteve, ato janë çelës për ta bërë DWH-në të përdorshme dhe të kuptueshme. Në shumë mënyra, DWH vlen vetëm aq sa atributet e dimensioneve. Atributet më të mira janë tekstuale dhe të dallueshme. Atributet duhet të përbëhen nga fjalë të vërteta, në vend se nga shkurtime kriptike.

Atributet tipike për një dimension Data do të përfshinin një përshkrim të shkurtër (10 deri 15 karaktere), një përshkrim të gjatë (30 deri 50 karaktere), dita e lindjes, dita e fillimit të trajtimit, dita e përfundimit të trajtimit, viti dhe shumë karakteristika të tjera të dimensionit data.

Megjithëse madhësia është me shumë mundësi numerike, përsëri është një atribut dimensionali, sepse sillet më shumë si një përshkrim tekstual sesa si një matje numerike. Madhësia është një përshkrues i qartë dhe konstant i një produkti specifik.

Ndonjëherë, kur ndërtojmë një databazë, është e paqartë nëse një fushë të dhënash numerike, e nxjerrë nga një burim të dhënash prodhimi, është fakt apo atribut dimensionesh. Ne shpesh mund të marrim vendim, duke pyetur nëse fusha është një matje që merr shumë vlera dhe merr pjesë në llogaritje (duke e bërë një fakt) apo është një përshkrim me vlerë të qartë, i cili është pak a shumë konstant dhe merr pjesë në kufizime (duke e bërë një atribut dimensional).

3.4.4 Çelësat e tabelës së dimensioneve

Tabelat e fakteve kanë një çelës me shumë pjesë, e rreshtat e dimensioneve dallohen në mënyrë unike nga një fushë çelësi e vetme. Meqë çelësat kryesorë të tabelave të

dimensioneve janë, thjesht, numra të plotë me radhë, duke filluar nga 1, që do të thotë se kur krijojmë një çelës për një regjistrim dimensionit të ri, ne, thjesht, shtojmë 1 në vlerën e fundit që kemi përdorur. Këta çelës zëvendësues janë pa kuptim, ata thjesht shpërbejnë si fusha bashkimi midis tabelave të fakteve dhe dimensioneve.

Meqenëse jemi përqendruar në dobritë, është e drejtë të pranojmë se ka, gjithashtu, një kosto në përdorimin e çelësve zëvendësues. Caktimi dhe menaxhimi i çelësve zëvendësues të tabelës së dimensioneve dhe zëvendësimi i mirë i tyre për çelësat natyralë operacional në rreshtat e fakteve, vendos një barrë në sistemin ETL.

Fatmirësisht, përpunimi është virtualisht identik në tabelat e fakteve dhe dimensioneve, prandaj mund të ndërtohet edhe njëherë dhe, pastaj, të ripërdoret. Gjithashtu, shumë ofrues të ETL japin aftësi të integruara për të lehtësuar përpunimin e çelësve zëvendësues.

3.4.5 Dimensionet e konfirmuara

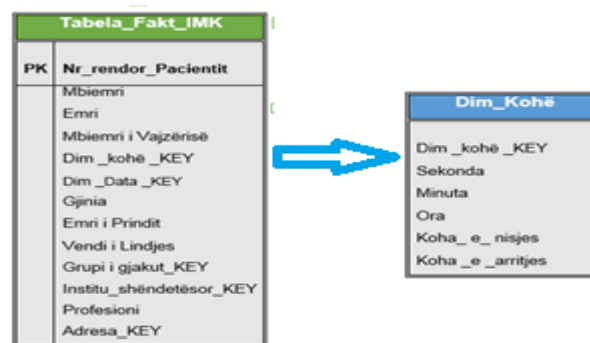
Dimensionet e konfirmuara të standardizuara janë synim për çdo DWH të arkitektuar mirë. Ndonjëherë i referuar si dimensione master ose dimensione referencë të përbashkëta, dimensionet e konfirmuara ndahen nëpër mjedisin e DWH të korporatës, duke iu bashkuar tabelave të fakteve të shumëfishta, që përfaqësojnë procese të ndryshme biznesi. Dimensionet e konfirmuara vijnë në dy forma: Në formën bazë, dy dimensionet e konfirmuara janë absolutisht identike, me të njëjtët çelës, emra atributesh, përkufizime atributesh dhe vlera të dhënash, pavarësisht nga tabela e fakteve me të cilat ato bashkohen. Për shembull, dimensionit Kohë nga tabela e fakteve i Institutit të Mjekësisë së Kosovës (IMK) është identike me dimensionin Data, të bashkuar në faktet e inventarëve.

Kjo tabelë dimensionesh jep të njëjtën përmbajtje, interpretim dhe prezantim, pavarësisht nga procesi i biznesit që përfshihet.

Ndryshe, dimensionet konfirmohen kur një dimension është një nëngrup perfekt i një tabele dimensionesh më të detajuar, më të copëzuar.

Në këtë rast, atributet që janë të përbashkëta për dimensionin e detajuar dhe dimensionin nëngrup të tkurrur, kanë të njëjtët emra attributesh, përkufizime dhe vlera të dhënash, siç ilustron në fig. 3.6

Figura 3.6 Tabelat Fakt-dimension



Tabelat e konfirmuara të detajuara dhe të tkurrura të dimensioneve korrespondojnë me tabelat e fakteve në copëzime të ndryshme.

Dimensionet e konfirmuara janë jashtëzakonisht të rëndësishme për natyrën AME të sistemit DWH/BI, pasi ato japin dobiti e mëposhtme:

- Qëndrueshmëri - Dimensionet e konfirmuara sigurojnë që çdo tabelë faktesh të filtrohet në mënyrë të qëndrueshme dhe grupet e përgjigjeve të kërkesës që vijnë si rezultat të emërtohen vazhdimisht.
- Integrim - Përmbajtja rreptësisht me dimensionet e konfirmuara lejon mjedisin DWH/BI, që të operojë si një i tërë i integruar. Dimensionet e konfirmuara lejojnë kërkesat të gërmojnë në tabelat e fakteve që përfaqësojnë procese të ndryshme, duke lëshuar kërkesa të veçanta të

çdo tabele të veçantë faktesh dhe, pastaj, duke iu bashkuar grupeve të rezultateve në atributet e dimensioneve të përbashkëta.

- Reduktim i kohës të zhvillimit në treg – Ndërkohë, ka një investim organizate, që duhet përcaktuar dhe menaxhuar dimensionet e konfirmuara.

Pasi janë ndërtuar, ato mund të pakësojnë në mënyrë dramatike kohën e zhvillimit për një projekt, meqë dimensionet e përbashkëta janë në dispozicion pa rikrijuar vazhdimisht. (Kimball&Becker, 2008). Më parë ka qenë një traditë për t'i ndjekur të tri etapat e modelimit gjatë ndërtimit të një databaze.

Modeli i parë, që duhet të ndërtohet, është ai konceptual, i cili fokusohet drejt biznesit për të përmbledhur kërkesat e tij në nivel të lartë. Zakonisht, prezantohet nga një skemë dhe duhet të jetë i pavarur nga tipi i databazës së përdorur. Modeli vijues është ai logjik, i cili duhet të përshkruajë skemën e saktë të përdorur nga databaza. Së fundi, është modeli fizik, i cili është përbërë nga një bashkësi skriptesh, i cili gjeneron dhe krijon tabelat, indekset, lidhjet etj. Në kohët e sotme shumë teknikë i bashkojnë në një fazë të vetme të tre modelet. Ata, thjesht, ndërtojnë një model (ER) dhe e konvertojnë direkt në një bashkësi objektesh të databazës.

Megjithatë, një DWH është i ndryshëm nga tipat e tjerë të databazave dhe për rrjedhojë rekomandohet të përdoret triologjia konceptuale/logjike/fizike (Lula & Kika, 2012:230).

3.4.6 Sjellja së bashku e fakteve dhe dimensioneve

Tani, që i kuptojmë tabelat e fakteve dhe dimensioneve, sjellim së bashku dy blloqe ndërtimi në një model dimensional, siç ilustruhet në fig. 3.7

Figura 3.7 Tabela e fakte dhe dimensioneve në modelin dimensional.

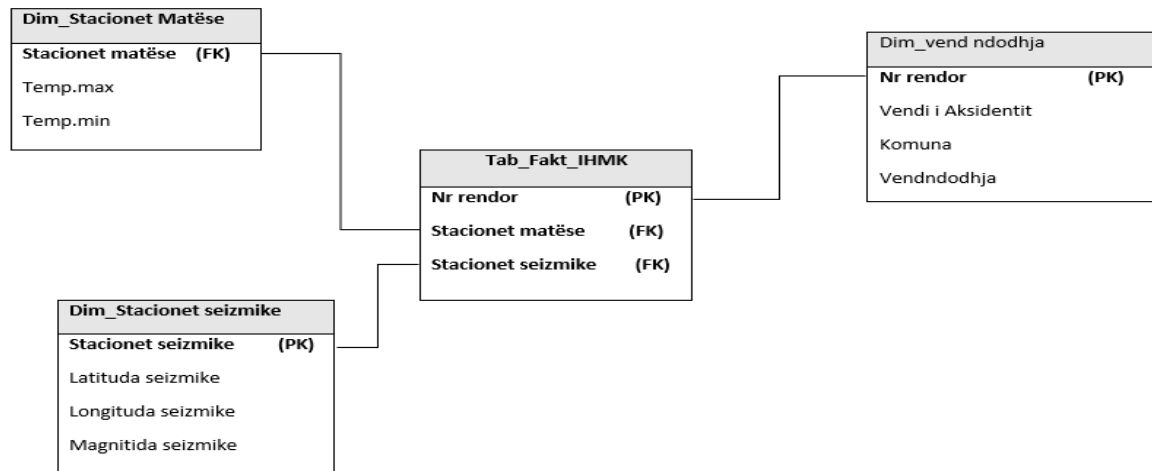


Tabela e fakteve, që përbëhet nga matje numerike, i bashkohet një grupi tabelash dimensionesh të mbushura me attribute përshkruese.

Kjo strukturë karakteristike, ngjashëm me yllin, shpesh quhet skema e bashkimit yll. Ky term datohet që në ditët e hershme të databazave relacionale. Gjëja e parë që vëmë re rreth skemës dimensionale është thjeshtësia dhe simetria. Pa dyshim, përdoruesit e biznesit përfitojnë nga thjeshtësia, meqë të dhënat janë më të lehta për t'u kuptuar dhe naviguar.

Struktura e ndërtimit në fig. 3.7 është fakti se është shumë e dallueshme për përdoruesit e menaxhimit të emergjencave. Ne kemi vërejtur me qindra raste, ku përdoruesit bien dakord menjëherë se modeli dimensional është mundësi për menaxhim efektiv të fatkeqësive natyrore. Për më tepër, numri i reduktuar i tabelave dhe përdorimi i përshkruesve kuptimplotë të menaxhimit e bën më pak të mundur ndodhjen e gabimeve. Thjeshtësia e një modeli dimensional ka, gjithashtu, dobi performance. Optimistët e databazave i përpunojnë këto skema të thjeshta në mënyrë më të efektshme me më pak bashkime.

Një vegël database mund të bëjë supozime shumë të forta, së pari rreth detyrimit të tabelave të dimensioneve të indeksuara, e pastaj edhe të tabelës së fakteve, njëherësh me produktin Kartezian të çelësave të tabelës së dimensioneve, duke plotësuar detyrimet e përdoruesit. Çuditërisht, duke përdorur këtë qasje, është e mundur të vlerësohen bashkimet n -way në një tabelë faktesh me një kalim të vetëm përmes indeksit të tabelës së fakteve.

Së fundi, modelet dimensionale janë shumë të zgjerueshme për të akomoduar ndryshimet. Struktura e parashikueshme e një modeli dimensional i reziston ndryshimeve të papritura në sjelljen e përdoruesit.

Çdo dimension është ekuivalent, të gjitha dimensionet janë pika hyrjeje të njëjta simetrikisht në tabelën e fakteve. Modeli logjik nuk ka paragjykim të integruar në lidhje me strukturat e pritura të kërkimit.

Me modelet dimensionale mund të shtojmë dimensione krejt të reja për sa kohë përcaktohet një vlerë e vetme e atij dimensionit për çdo rresht ekzistues faktesh.

Po kështu, mund të shtojmë fakte të reja, të papritura në tabelën e fakteve, duke supozuar që niveli i detajeve përputhet me tabelën e fakteve ekzistuese.

Ne mund të shtojmë tabela dimensionesh para-ekzistuese me attribute të reja, të papritura. Gjithashtu, mund t'i ndajmë rreshtat ekzistues të dimensioneve në një nivel më të ulët copëzimi nga një pikë e caktuar në kohë. Në çdo rast, tabelat ekzistuese mund të ndryshohen në vend, thjesht duke shtuar rreshta të rinj të dhënash në tabelë ose duke ekzekutuar një komandë SQL ALTER TABLE. Të dhënat nuk do të kishin nevojë të ngarkoheshin përsëri. Të gjitha aplikimet ekzistuese të aksesit të të dhënave vazhdojnë të punojnë pa nxjerrë rezultate të ndryshme (Kimball&Ross, 2002).

3.5 Modelimi konceptual

Modeli konceptual është një përfaqësim i koncepteve dhe marrëdhënieve midis tyre. Është kryesisht për të kapur kërkesat e përdoruesve të vendimmarrjes pa u shqetësuar për detajet e implementimit. Ka modele konceptuale që ekzistojnë për databazat relacionale dhe speciale, por ato nuk shkallëzohen mirë në DWH speciale shkaku i pranisë së hierarkive, përmbledhjeve, matjeve dhe dimensioneve (Garg&Mithal).

Në këtë modelim identifikohen entitetet dhe relacionet midis tyre dhe përcaktohet niveli më i ulët i detajimit të ruajtjes së informacionit, që plotëson kërkesat e analizimit të informacionit. Përcaktimi i nivelit më të ulët të informacionit lidhet ngushtë me një përlogaritje të volumeve fizike të informacionit, që nevojiten për ruajtjen e të dhënave burim (Juliana, 2014).

Modelet konceptuale lehtësojnë komunikimin midis përdoruesve dhe ndërtuesve, meqenëse nuk kërkojnë njohuri për veçoritë specifike të platformës së implementimit. Për më tepër, skemat e zhvilluara, duke përdorur modelet konceptuale, mund të hartohen në modele logjike të ndryshme, si modelet relacionale, objekt-relacionale ose të orientuara nga objekti, duke thjeshtuar kështu përgjigjet për ndryshimet në teknologjinë e përdorur.

Kështu, modelet konceptuale lehtësojnë mirëmbajtjen dhe evolucionin e databazës, meqë përqendrohen në kërkesat e përdoruesve, e si pasojë japin mbështetje më të mirë për ndryshimet në skemat logjike dhe fizike (Alejandro&Esteban, 2014).

Bashkë me këtë vendim u bën parashikimet përkatëse të volumeve fizike të nevojshme për ruajtjen e informacionit. Ruajtja në nivel shumë të detajuar të rekordeve të gjeneruar nga çdo

platformë kërkon investim të lartë në hardware dhe, gjithashtu, një strategji të mirë përcaktuar të arkivimit të tyre. Informacioni i shtuar çdo ditë në databazën destinacion është disa GB.

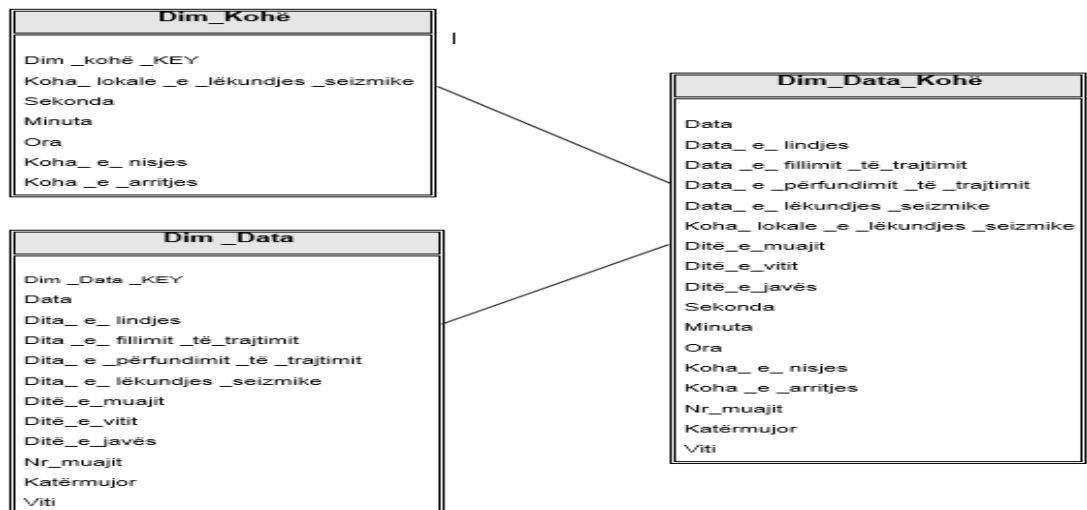
Nga analiza e përbashkët e kërkesave të grumbulluara u përcaktuan elementët, mbi bazë të të cilave kërkohej matje. Këto, sipas modelit shumëdimensional, do të quhen dimensionet e skemës. Gjatë modelimit konceptual një vend të veçantë zë analiza e elementit kohë. Dimensioni kohë është pjesë e çdo analize të mundshme, ai mund të jetë i nevojshëm deri në nivel sekonde për analiza të detajuara (Juliana, 2014).

Megjithëse dimensionin e datës është dimensionin më të rëndësishëm të kohës, duhet, gjithashtu, një dimension muaji kur shtrirja kohore e tabelës së fakteve është një muaj.

Në disa mjedise duhet të ndërtojmë dimensionet javore, tremujore ose vjetore, si edhe nëse ka tabela faktesh në secilën prej këtyre shtrirjeve (Kimball&Caserta, 2004).

Duke u bazuar në modelet shumëdimensionale, elementi kohë u konsiderua si një dimension që ndryshon shpesh dhe duhet të ndahet. Sipas arsyetimit të mësipërm, dimensionin e datës_kohë duhet të ndahet në dy tabela të skemës sonë (Juliana, 2014).

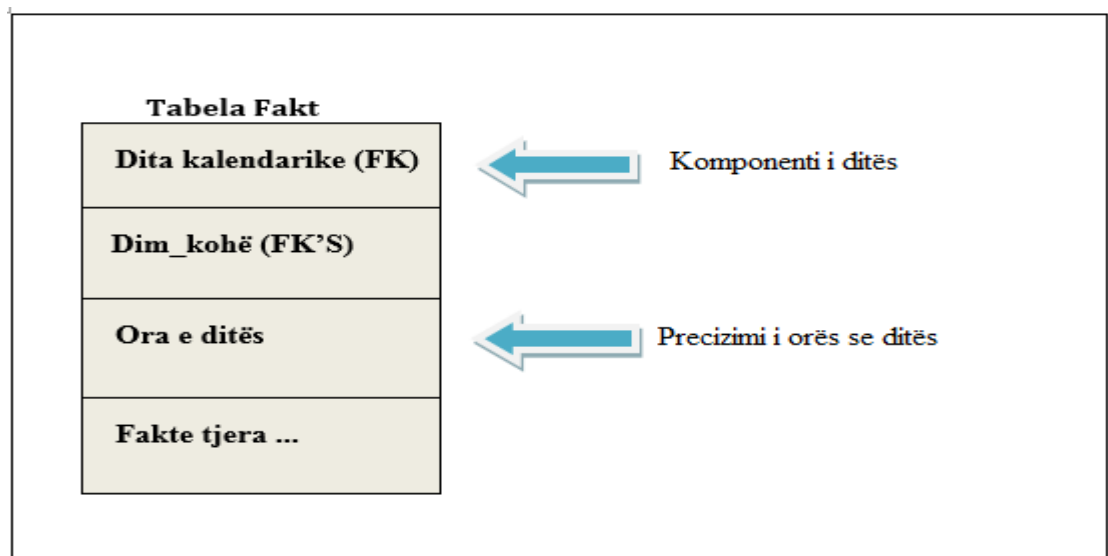
Figura 3.8 Ndarja e dimensioneve në skemën (Juliana, 2014).



Dimensioni i muajit duhet të jetë një tabelë fizike e veçantë dhe duhet të krijohet duke eliminuar fizikisht rreshtat dhe kolonat e zgjedhura nga dimensionin e ditës kalendarike.

Në disa tabela faktesh, koha matet në nivelin e ditës kalendarike, deri në minutën, madje edhe sekondën. Nuk mund të ndërtohet një dimension kohe me çdo minutë ose çdo sekondë që ekziston. Ka më shumë se 31 milionë sekonda në vit. Ne duam të ruajmë dimensionin e fuqishëm të datës kalendarike dhe të mbështetim, njëkohësisht, kërkimin e saktë deri në minutën ose sekondën. Gjithashtu, mund të llogaritim intervale kohore shumë të sakta, duke krahasuar kohën e saktë të dy regjistrimeve të tabelës së fakteve. Për këto arsye, rekomandojmë ndërtimin e treguar në figurën 3.9

Figura 3.9 Projektimi i tabelës fakt për trajtimin e matjeve në kohë të sakta (Kimball&Caserta, 2004).



Komponenti i ditës kalendarike të kohës së saktë mbetet si çelës i huaj referencë për dimensionin tonë të njohur të ditës kalendarike. Mendojeni si një lloj të veçantë fakti, jo si një dimension.

Në këtë rast është e dobishme të bëjmë një dimension komponentë minutash ose sekondash të kohës të saktë, sepse llogaritja e intervaleve kohore në rekordet e tabelës së fakteve bëhet shumë e ngatërruar kur përpiqemi të merremi me dimensione të veçanta dite dhe ore-të-ditës (Kimball&Caserta, 2004).

Pjesa e kodit të procedurës Oracle që popullon tabelën Dim_Kohe dhe Dim_Data është si më poshtë:

```
CREATE OR REPLACE PROCEDURE POPULLIMI_KOHE_DIM
AS
    v_koha_plote    DATE;
    v_sekonda       NUMBER;
    v_minuta        NUMBER;
    v_ore           NUMBER;
    v_key           NUMBER;
    v_koha_e_nisjes VARCHAR (40);
    v_koha_e_arritjes VARCHAR (40);
    v_koha_plotel   DATE;

BEGIN
    DELETE FROM DIM_KOHE;
    v_koha_plote := TO_DATE ('06-04-15 00:00:00', 'DD-MM-YY HH24:MI:SS');
    v_koha_plotel:= TO_DATE ('28-10-15 00:00:00', 'DD-MM-YY HH24:MI:SS');
    v_key := 1;
    WHILE v_koha_plote < v_koha_plotel
    LOOP
        BEGIN

            v_sekonda := TO_CHAR (v_koha_plote, 'SS');

            v_minuta := TO_CHAR (v_koha_plote, 'MI');

            v_ore := TO_CHAR (v_koha_plote, 'HH24');

            v_koha_e_nisjes:= TO_CHAR (v_koha_plote, 'HH24:MI:SS');
            v_koha_e_arritjes:= TO_CHAR (v_koha_plote, 'HH24:MI:SS');

            INSERT INTO DIM_KOHE
                (Dim_kohe_Key,
                 sekonda,
                 minuta,
                 ore,
                 kohe_e_nisjes
                 kohe_e_arritjes)
            VALUES (v_key,
                    v_sekonda,
                    v_minuta,
```

```

v_ore,
v_kohe_e_nisjes
v_kohe_e_arritjes);
v_koha_plote := v_koha_plote + 1 / (24 * 60 * 60);
v_key := v_key + 1;
END;
END LOOP;
END;

```

Pjesa e kodit të procedurës Oracle që popullon tabelën Dim_Kohe është si më poshtë:

```

CREATE OR REPLACE PROCEDURE POPULLIMI _DATA_DIM
(p_data_filllimit DATE,
 p_data_mbarimit DATE)
AS

```

```

        v_data_plote DATE;
        v_dite_e_muajit NUMBER;
        v_dite_e_vitit NUMBER;
        v_dite_e_javes VARCHAR2 (30);
        v_dita_e_lindjes VARCHAR2 (20);
        v_dita_e_filllimit_te_trajtimit VARCHAR2(40);
        v_dita_e_perfundimit_te_trajtimit VARCHAR2 (40);
        v_dita_e_lekundjes_seizmike VARCHAR2(30);
        v_nr_muaji NUMBER;
        v_viti NUMBER;
        v_katermujor NUMBER;
        v_key NUMBER;

BEGIN
v_data_plote := p_data_filllimit;
v_key := 1;
WHILE v_data_plote < p_data_mbarimit
LOOP
BEGIN
v_dite_e_muajit := TO_CHAR (v_data_plote, 'DD');
v_dite_e_vitit := TO_CHAR (v_data_plote, 'DDD');
v_dite_e_javes := UPPER (TO_CHAR (v_data_plote, 'DAY'));
v_dita_e_lindjes:= UPPER (TO_CHAR (v_data_plote, 'DAY'));
v_dita_e_filllimit_te_trajtimit:= UPPER (TO_CHAR (v_data_plote, 'DAY'));
v_dita_e_perfundimit_te_trajtimit:= UPPER (TO_CHAR(v_data_plote, 'DAY'));
v_dita_e_lekundjes_seizmike:= UPPER (TO_CHAR(v_data_plote, 'DAY'));
v_muaji := UPPER (TO_CHAR (v_data_plote, 'MONTH'));
v_nr_muaji:= TO_CHAR (v_data_plote, 'MM');
v_viti:= TO_CHAR (v_data_plote, 'YYYY');
v_katermujor:= TO_CHAR (v_data_plote, 'Q');
INSERT INTO Dim_Data (Dim_data_Key,

```

Data

Dita_ e_lindjes

Dita_ e_filllimit_te_trajtimit

Dita_ e _perfundimit_te_trajtimit

Dita_ e_lekundjes _seizmike

```

Dite_e_muajit
Dite_e_vitit
Dite_e_javes
Nr_muajit
Katermujor
Viti)

VALUES (v_key,
v_data_plote,
v_Dita_e_lindjes

v_Dita_e_filllimit_te_trajtimit
v_Dita_e_perfundimit_te_trajtimit
v_Dita_e_lekundjes _seizmike

v_dite_e_muajit,
v_dite_e_vitit,
v_dite_e_javes,
v_muaji,
v_nr_muaji,
v_katermujor,
v_viti);
v_data_plote := v_data_plote + 1;
v_key := v_key + 1;
END;
END LOOP;
END;

```

Ky dokument i përgatitur do të përbëjë thelbin e këtij data mart-i dhe do të shërbejë si një fjalor (metadata). Pas përfundimit të modelit konceptual do vazhdojmë me ndërtimin e modelit logjik (Juliana, 2014).

3.6 Modelimi logjik

Modelimi logjik transformon modelin e të dhënave konceptuale në një model dimensional, i njohur zakonisht si skema yll. Përbëhet nga një tabelë e madhe faktesh (e njohur si tabela e fakteve), me një numër tabelash të tjera që e rrethojnë, të cilat përmbajnë të dhëna

përshkruese, të quajtura dimensione. Modeli i të dhënave logjike përfshin veçoritë e mëposhtme:

- Përfshin të gjitha entitetet dhe marrëdhëniet midis tyre.
- Të gjitha atributet për çdo entitet janë të përcaktuara.
- Çelësi kryesor për çdo entitet është i përcaktuar.
- Çelësat e huaj ‘foreign keys’ (çelësat që përcaktojnë marrëdhënien midis entiteteve të ndryshme) janë të përcaktuara.

Në këtë nivel, modeluesi i të dhënave përpiket të përshkruajë të dhënat me sa më shumë detaje që të jetë e mundur, pa u shqetësuar për mënyrën se si do të implementohen fizikisht në databazë.

Secili prej emrave të dimensioneve, i treguar në diagramin me pika, zakonisht, bëhet një tabelë dimensionesh e veçantë në skemën yll.

Sidoqoftë, çelësat natyralë duhet të zëvendësohen me të ashtuquajturit çelësa zëvendësues ‘surrogate key’. Çelësi zëvendësues është një çelës për rreshtin e databazës, që është unik për çdo rresht në një tabelë dimensionesh dhe nuk nxirret nga të dhënat e rreshtit.

Secili prej këtyre çelësave duhet të jetë një numër dhjetor i thjeshtë, duke filluar me 1 dhe duke shkuar në numrin më të madh që nevojitet (Kimball&Caserta, 2004).

Lidhjet midis tabelave dimension dhe tabelës fakt duhet të bazohen ekskluzivisht në këta çelësa. Zakonisht, një ‘primary key’ i tabelës fakt është një bashkim i çelësave. Ka disa arsye pse duhet të përdoren ‘surrogate keys’ dhe një prej tyre është nevoja për të regjistruar të dhëna të pasigurta. Kjo situatë, zakonisht, njihet me termin ‘nuk e di’ dhe mundëson të shtosh

një rresht në tabelën e dimensioneve, që do të ketë si vlerë të barabartë me -1. Këto mund të përfshijnë rekordet problematike, me informacione jo të plota.

Gjithashtu, ekziston një tjetër problem që lidhet me përfshirjen e kohës dhe duhet të konsiderohet me kujdes gjatë modelimit logjik. Në rastet kur duhet të kontrollohen ndryshimet, është e papranueshme të ndërtohet çdo dimension si i varur nga koha. Dimensionet e pavarura nga koha janë të njohura me emrin SCD (dimensione shumë pak të ndryshueshme).

Analiza e problemit dhe zgjidhja e tij

Duke pasur parasysh shumicën e teknikave për ndërtimin e DWH /data mart, analizimi i kërkesave cilësohet njëra nga prioritetet parësore të procesit të ndërtimit të DWH/data mart. Duke ndjekur modelin Kimball, identifikimi i elementeve bazë të modeleve shumëdimensionale mund të përkufizohet përmes sistemit matricë e lidhjes dimension/fakt/matje (tabela 3.2).

Pra, ndërtimi i data mart-it do të bazohet mbi matricën që përcakton qartë elementet e skemës.

Tabela 3.2 Matrica dimension/fakt

Dimensione/fakte	Dim_Kohe	Dim_Grupi i gjakut	Dim_Adresa	Dim_Data	Dim_Instituti shëndetësor
Tabela Fakt_IMK	X	X	X	X	X

Për rastin tonë dimensionet janë:

Koha - përbën informacionin kalendarik (ditën, muajin, vitin, orën, minutën, sekondën) dhe vlerësohet njëra nga dimensionet më të detajuara.

Grupi i gjakut - paraqet njërën nga dimensionet më të detajuara lidhur me pacientët.

Adresa - zakonisht është njëra nga më të veçantat në ndërtimin e një data mart-i.

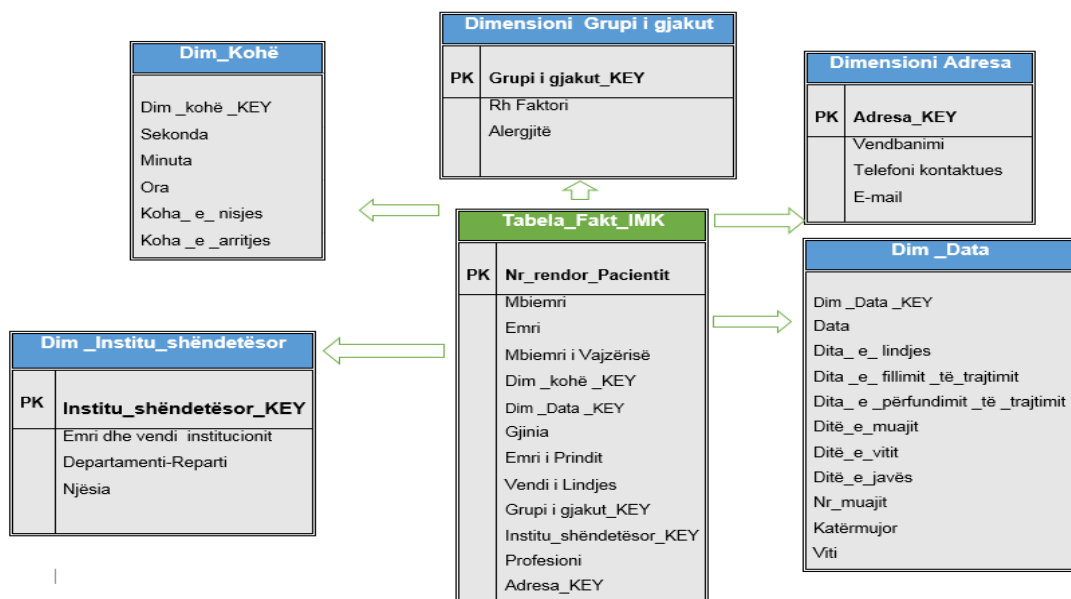
Data - informacioni më i detajuar si dimension, që ruan të dhëna në bazë të datave.

Instituti shëndetësor - paraqet institucionin që menaxhon me një numër të madh të pikave që ofrojnë shërbime në fushën shëndetësore.

Ndërsa, tabela Fakt është:

Tabela Fakt_IMK ruan informacione nga institucioni mjekësor i Kosovës, që në vete përbën një gamë të gjerë të të dhënave, ku përfshihen të dhënat nga shumë burime të të dhënave të fushës së përkujdesjes mjekësore. Skema e data mart-it paraqitet si më poshtë:

Figura 3.10 Skema e data mart-it të IMK-së



Në rastin e data mart-it të IMK-së një ‘dimension junk’ do të përfshinte të dhënat bazike të mbiemrit, emrit, gjinia, emri i prindërit, vendi i lindjes, profesioni, adresa etj. Në skemë do të përfaqësojnë një dimension të veçantë.

Përveç strukturimit të dimensioneve duhet të ndërtohen edhe tabela fakt. Zakonisht, ky tip tabele përmban të gjitha ‘foreign keys’, si dhe të gjitha matjet numerike të quajtura si fakte. Në rastin e tabelës fakt të modelit do të jetë vetëm një matje numerike, që ka të bëjë me kohëzgjatjen e komunikimeve në IMK. (Lula&Kika, 2012:230).

Për çdo atribut në një tabelë dimension duhet të përcaktohet një strategji për të ruajtur ndryshimet. Ka gjithsej tre tipa teknikash:

3.6.1 Tipi 1 - SCD

Megjithëse vendosja e rekordeve të reja në një SCD Tipi 1 kërkon gjenerimin e çelësave të rinj të dimensionit, ndryshimet e përpunimit në një SCD Tipi 1 nuk ndikon asnjëherë në çelësat e tabelës së dimensioneve ose çelësat e tabelës së fakteve dhe, në përgjithësi, ka ndikimin më të vogël në të dhënat e tre llojeve SCD.

SCD Tipi 1 mund të ketë një ndikim në ruajtjen e tabelave të fakteve agregate, nëse ndonjë agregat ndërtohet direkt në atributin që u ndryshua.

Bazuar në çelësin natyral të nxjerrë nga sistemi burim, çdo rekord i ri i caktohet një çelësi zëvendësues të ri dhe i lidhur me të dhënat e dimensioneve ekzistuese. Rekordet ekzistuese update-ohen në vend. Performanca e kësaj teknike mund të jetë më e dobët në krahasim me ngarkimin në një ngarkues në shumicë, por nëse tabelat janë me madhësi të arsyeshme, impakti duhet të jetë i pranueshëm.

Disa mjete ETL ofrojnë transformime të specializuara, që mund të zbulojnë nëse një duhet të futet apo të update-ohet një rekord. Sidoqoftë, ky utilitet duhet të ngulë tabelën, duke përdorur çelësin kryesor të regjistrimit të kandidatit për të parë nëse ekziston.

Kjo qasje është proces intensiv dhe duhet të shmanget. Për të minimizuar goditjen e performancës, kur përdoret SQL për të ngarkuar një dimension Tipi 1, procesi ETL duhet të veçojë qartë të dhënat ekzistuese që kërkojnë deklarata UPDATE nga të dhënat që kërkojnë INSERT (Kimball&Caserta, 2004). Sipas kësaj teknike, atributi i vjetër do të mbivendoset nga vlera e re. Kjo është shumë e lehtë për t'u aplikuar, por nuk mirëmban asnjë historik të vlerave të hershme për këto dimensionet. Prandaj, është më e përshtatshme që kjo teknikë të përdoret për të ruajtur vlera të korrigjuara pa u shqetësuar për historikun (Lula&Kika, 2012:230).

3.6.2 Tipi 2 - SCD

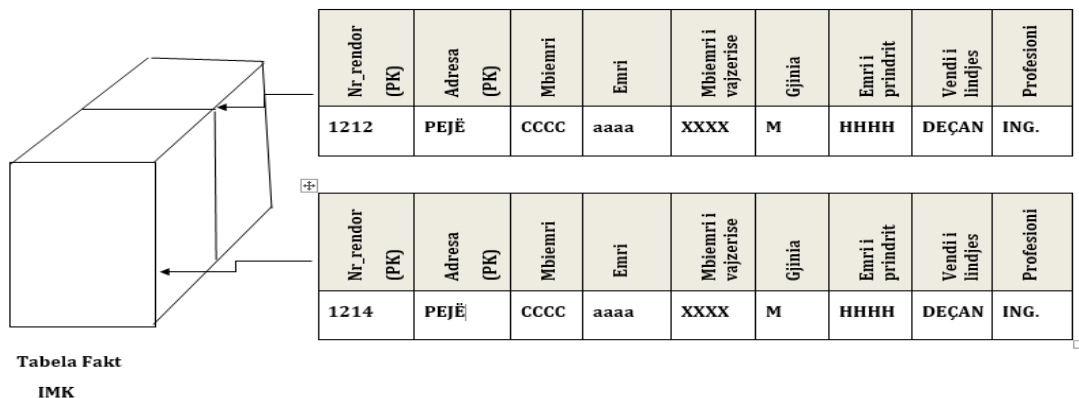
Kur DWH njoftohet se një regjistrim dimensionit ekzistues duhet të ndryshohet, në vend se të mbishkruhet, DWH lëshon një regjistrim dimensionit të ri në momentin e ndryshimit.

Këtij regjistrimi, dimensionit të ri, i caktohet një çelës kryesor zëvendësues i freskët dhe ai çelës përdoret nga ai moment e tutje në të gjitha tabelat e fakteve që kanë atë dimension si çelës të huaj.

Për sa kohë çelësi i ri zëvendësues caktohet menjëherë në momentin e ndryshimit, asnjë çelës ekzistues në çdo tabelë faktesh nuk ka nevojë të përditësohet ose ndryshohet dhe asnjë tabelë faktesh agregate nuk ka nevojë të rillogaritet.

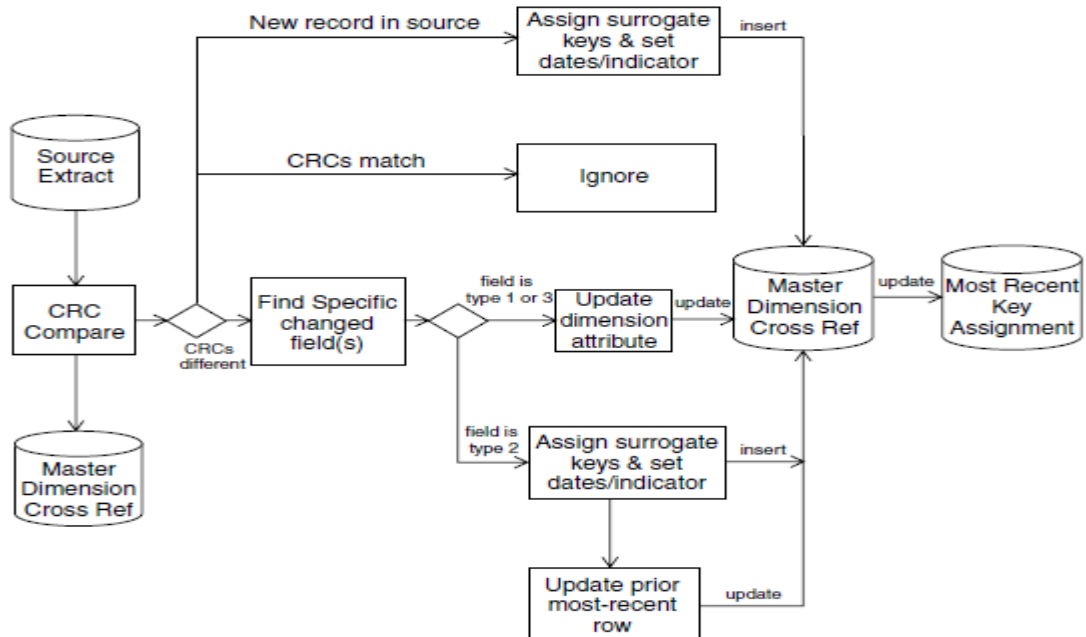
Ne themi se SCD Tipi 2 e ndanë historinë në mënyrë perfekte, sepse çdo version i detajuar i entitetit dimensional është i lidhur saktë me shtrirjen e regjistrimeve të tabelës së fakteve për të cilin ai version është shumë i saktë. Në figurën 3.11 ilustron...

Figura 3.11 Tipi 2 – SCD, historiku perfekt i partizioneve



Për një dimension të vogël me disa mijëra rekordesh dhe një dhjetë fusha, si, për shembull, një fajll të thjeshtë pacienteve, zbulimi i ndryshimeve i treguar në figurën 3.12 mund të bëhet me forcë.

Figura 3.12 Menaxhimi i tabelës dimension ‘surrogate key’ (Kimball&Caserta, 2004).



Duke krahasuar çdo fushë në çdo rekord në shkarkimin e sotëm me çdo fushë në çdo rekord nga i djeshmi, duhet të gjenden shtimet, ndryshimet dhe fshirjet.

Por, për një dimension të madh, një listë e tillë me miliona pacientë të siguruar me 100 fusha përshkruese në çdo rekord, qasja e forcës së krahasimit të çdo fushe në çdo rekord është shumë joefikase (Kimball&Caserta, 2004).

Kjo teknikë përdoret më së shumti në dimensionet SCD. Kjo metodë mundëson krijimin e një rreshti të ri, sa herë që duhet të ndryshohet ndonjë atribut i dimensionit. Gjithë atributet e tjerë do të jenë me të njëjtat vlera. Normalisht, një tjetër ‘surrogate key’ do të gjenerohet për këtë rresht të ri (Lula&Kika, 2012:230)

3.6.3 Tipi 3 - SCD

Përdoret kur ndodh një ndryshim në një dimension regjistrimi, por vlera e vjetër e atributit mbetet e vlefshme si zgjedhje e dytë.

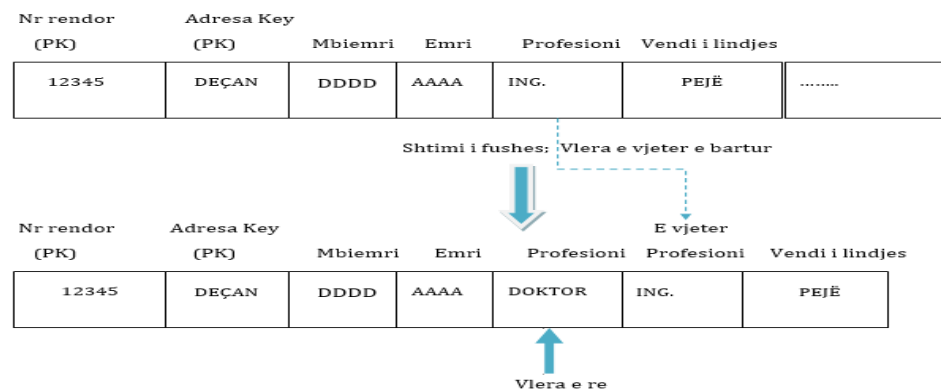
Dy situatat më të zakonshme të IMK, ku ndodh kjo, janë ndryshimet, e ku caktimi i territorit të vjetër duhet të vazhdojë të jetë në dispozicion si zgjedhje e dytë dhe ndryshimet në caktimet nr. rendor-profesioni, ku caktimi profesioni i vjetër duhet të vazhdojë të jetë i disponueshëm si zgjedhje e dytë. Arkitekti i DWH-së duhet t’i dallojë fushat që kërkojnë administrimin e Tipit 3.

Në SCD Tipi 3, në vend se të lëshohet një rresht i ri kur bëhet një ndryshim, krijohet një kolonë e re (nëse nuk ekziston tashmë) dhe vlera e vjetër vendoset në këtë fushë të re para se të mbishkruhet vlera primare.

Për shembullin nr. rendor-profesioni, ne supozojmë se fusha kryesore quhet profesioni. Për të implementuar SCD Tipi 3, ndryshojmë tabelën e dimensioneve për të shtuar fushën

profesioni i vjetër. Në kohën e ndryshimit ne e marrim vlerën origjinale të profesionit dhe e shkruajmë atë në fushën profesioni i vjetër; pastaj e mbishkruajmë fushën profesioni sikur të ishte ndryshim Tipi 1. (Shiko figurën 3.13).

Figura 3.13 Implementimi i Tipit 3-SCD në përshkrimin nr. rendor-profesioni



Nuk ka nevojë të ndryshohen çelësat në asnjë tabelë dimensionesh ose në asnjë tabelë faktesh. Njëlloj si SCD Tipi 1, nëse tabelat agregate janë ndërtuar direkt në fushën poshtë ndryshimit Tipi 3, këto tabela agregate duhet të rillogariten.

Kur shtohet një regjistrim i ri në një dimension që përmban fusha Tipi 3, duhet të nxirret një rregull i IMK-së për të vendosur si të popullohet fusha e vlerës të vjetër.

Vlera aktuale mund të shkruhet në këtë fushë ose mund të jetë NULL, në varësi të rregullit të IMK-së. Ne shpesh e përshkruajmë SCD Tipi 3 si mbështetës të një realiteti alternativ. Në shembullin tonë nr. rendor-profesioni, përdoruesi final mund të zgjedhë midis dy versioneve të hartimit të nr. rendor në profesion.

Qasja SCD Tipi 3 mund të zgjatet në shumë realitete alternative, duke krijuar një numër arbitrar fushash alternative bazuar në atributin origjinal (Kimball & Caserta, 2004).

Sipas kësaj teknike duhet të regjistrohen vlerat para dhe pas ndryshimit. Kjo realizohet me shtim të një kolone të re për vlerën e shtuar. Kjo teknikë përdoret shumë rrallë. Zakonisht, në praktikë kombinohen teknika e parë dhe e dytë për të njëjtin rresht informacioni, sepse mund të ketë fusha për të cilat mund të ruhen ose jo ndryshimet (Lula&Kika, 2012:230).

Një situatë tjetër shumë e zakonshme, kur bëhet fjalë për modelimin logjik, është kur ka nevojë për t'i përfshirë disa flamuj dhe tregues, që janë të lidhur dukshëm me dimensionet, por nuk përshtaten në asnjë prej tabelave të dimensioneve të konfirmuara.

Në teori ka katër zgjidhje për këtë problem:

- Flamujt dhe treguesit janë pjesë e tabelës së fakteve.
- Nxirr të gjithë flamujt dhe treguesit nga ndërtimi.
- Vendos çdo flamur dhe tregues në dimensionin e tij të veçantë (përkatës).
- Bashkohen të gjithë flamujt dhe treguesit në një ose më shumë 'dimensione jun'.

Në praktikë vetëm zgjidhja e fundit është e përshtatshme.

Një dimension shumicë është një grupim i mirë i flamujve dhe treguesve (kjo mund të bëhet, gjithashtu, për të dhëna të tjera jo përkatëse, të cilat do të kishin kuptim të kombinoheshin).

Përfshin të gjitha kombinimet e vlefshme të attributeve, të cilat nënkuptojnë që seti i të dhënave finale do të ishte një produkt kartezian i të gjithë flamujve dhe treguesve individualë.

Nëse të gjitha vlerat e mundshme rezultojnë në një set të dhënash të vogël, është mirë që rreshtat të krijohen më përpara (Kimball&Caserta, 2004).

3.7 Modelimi fizik

Në këtë nivel modeluesi i të dhënave përcakton mënyrën se si do të realizohet modeli i të dhënave logjike në skemën e databazës.

Objektivat e procesit të ndërtimit fizik nuk përqendrohen në struktura. Modeluesi është më i shqetësuar për mënyrën si do të punojë modeli në vend se me mënyrën si do të duket. Sipas Poniah, ka disa hapa të ndryshëm në procesin e ndërtimit fizik për një DWH:

Zhvillimi i standardeve

Standardet variojnë nga mënyra si të emërtohen fushat në databazë deri në bërjen e intervistave me përdoruesit finalë për përkufizimin e kërkesave.

Standardi për emërtimin e objekteve merr domethënie të veçantë, sepse përdorimi i emrave të objekteve nuk është i kufizuar vetëm për specialistët IT.

Përdoruesit i referohen, gjithashtu, objekteve me emra, kur krijojnë dhe bëjnë kërkimet e tyre. Kjo është arsyeja pse vetë emri duhet të jetë në gjendje të përcjellë kuptimin dhe përshkrimin e objektit. Gjithashtu, është e nevojshme të adoptohen standarde efektive për emërtimin e strukturës së të dhënave në fazën e skenimit, si edhe të gjitha llojet e skedave (jo vetëm skedat e të dhënave dhe indekseve, por edhe skedat që përmbajnë kode dhe skripte, skeda databazë dhe dokumente aplikacionesh).

Krijimi i planeve të të dhënave të grumbulluara

Në këtë hap duhet të shqyrtohen mundësitë për ndërtimin e tabelave të të dhënave të grumbulluara. Është e mundur që shumë prej të dhënave të grumbulluara do të tregohen në sistemin OLAP. Sidoqoftë, nëse rastet OLAP nuk janë për përdorim universal nga të gjithë përdoruesit, atëherë të dhënat e grumbulluara duhet të tregohen në DWH.

Janë dy faktorë kryesorë që duhet të merren në konsideratë kur merret një vendim për planin e të dhënave të grumbulluara:

- Së pari, modelet e aksesit të përdoruesve të agjencioneve duhet të merren parasysh - të dhënat që ata po përmbledhin shpeshherë duhet të paraqiten në DWH.
- Së dyti, duhet të vlerësohet shpërndarja statistikore e të dhënave (p.sh., sa raste unike ekzistojnë në secilin nivel hierarkie dhe cila është forma e shpërndarjes në rast të lëvizjes nga një nivel në tjetrin).

Përcaktimi i skemës së ndarjes së të dhënave

Duhet të merren parasysh mundësitë e ndarjeve për tabelat e fakteve, si edhe tabelat e përmasave. DWH, zakonisht, mban tabela databazë shumë të mëdha.

Tabelat e fakteve janë në miliona rreshta dhe shumë tabela dimensionesh (p.sh., tabelat e pacienteve). Ato përmbajnë një numër të madh rreshtash. Duke pasur tabela me madhësi shumë të mëdha, hasen disa probleme të caktuara. Ngarkimi i tabelave të mëdha dhe ndërtimi i indekseve kërkon shumë kohë. Kërkimet, gjithashtu, zgjatin më shumë, rezervimi dhe rikuperimi i tabelave të mëdha, po ashtu. Kjo është arsyeja pse ndarja në particione është e rëndësishme.

Ndarje në particione do të thotë ndarje e qëllimshme e një table dhe e të dhënave të indeksit në pjesë të menaxhuara. Sistemi i menaxhimit të databazës (DBMS) mbështet dhe siguron mekanizmin për këtë. Çdo particion i një table trajtohet si objekt më vete. Particionet mund të shpërndahen nëpër disa disqe për të arritur performancë optimale.

Një tabelë shumë e madhe mund të ndahet vertikalisht ose horizontalisht. Në ndarjen vertikale njëra ndan particionet, duke i grupuar kolonat e zgjedhura së bashku. Në këtë skenar çdo tabelë e ndarë në particione përmban të njëjtin numër rreshtash si tabela origjinale. Zakonisht, tabelat me dimensione të mëdha janë kandidatë për ndarjen vertikale.

Në rast të ndarjes horizontale, tabela ndahet duke i grupuar rreshtat e zgjedhur së bashku. Ndarja horizontale e tabelës së madhe të fakteve të bazuara në datat kalendarike, zakonisht, punon mirë në një mjedis të DWH-së.

Skema e ndarjes në particione duhet të përfshijë si më poshtë: tabelat e fakteve dhe tabelat e dimensioneve të zgjedhura për ndarjen në particione, lloji i ndarjes për çdo tabelë (horizontal ose vertikal), numri i particioneve për çdo tabelë, kriteret për të ndarë çdo tabelë dhe përshkrimi se si të bëhen kërkime që janë në dijeni të particioneve (një kërkim ka nevojë të përdorë vetëm partitionet e nevojshme). Çdo DBMS ka mënyrën e vet të caktuar të implementimit të particioneve fizike. Një nga shumë konsideratat e rëndësishme kur zgjidhet DBMS është mbështetja e saj për ndarjen në particione dhe indeksimi në rastin Tabela_Fakt_IMK.

```
Create table Tabela_Fakt_IMK
(
    Nr_rendor_pacienti    NUMBER (8) not null,
    Mbiemri              VARCHAR2 (50) not null,
    Emri                 VARCHAR2 (50) not null,
    Mbiemri i vajzerise   VARCHAR2 (50) not null,
    Dim_Data_KEY          NUMBER(4) notnull,
    Dim_kohe_KEY          NUMBER (5) not null,
    Gjinia               CHAR (3) not null,
    Emri i prindit        VARCHAR2 (50) not null,
    Vendi i lindjes       VARCHAR2 (50) not null,
    Grupi i gjakut_Key    NUMBER (5) not null,
    Instituti_shendetesor_Key    NUMBER (5) not null,
    Profesion            VARCHAR2 (25) not null,
    Adresa_Key            NUMBER (5) not null,
)
PARTITION BY RANGE (Dim_Data_KEY)
SUBPARTITION BY Hash (ROUTING_CATEGORY)

(
PARTITION Qershor_2015 values less than (TO DATE ('01-JAN-2015','dd-MON-
YYYY'))
TABLESPACE IMK_tbl
PCTFREE 10
INITRANS 1
```

```

MAXTRANS 255
(
    SUBPARTITION p1 TABLESPACE IMK_tbl,
    SUBPARTITION p2 TABLESPACE IMK_tbl,
    SUBPARTITION p3 TABLESPACE IMK_tbl,
    SUBPARTITION p4 TABLESPACE IMK_tbl,
    SUBPARTITION p5 TABLESPACE IMK_tbl,
    SUBPARTITION p6 TABLESPACE IMK_tbl,
    SUBPARTITION p7 TABLESPACE IMK_tbl,
    SUBPARTITION p8 TABLESPACE IMK_tbl,
)
...
...
ALTER TABLE Tabela_Fakt_IMK
ADD constraint trafiku_PRIM primary key (Nr_rendor_Pacientit)
using index
tablespace trafiku_index
pctfree 10
initrans 2
maxtrans 255
storage
(
    initial 64K
    next 1M
    minextents 1
    maxextents unlimited);

```

Themelimi i opsioneve të bashkimit të copave

Bashkimi i copave përfshin vendosjen dhe menaxhimin e njësive përkatëse të të dhënave në të njëjtin bllok fizik memorieje. Në këtë mënyrë, ato njësi mbledhen së bashku në një veprim të vetëm. Për të krijuar mundësitë e mira për bashkimin e copave, tabelat duhet të ekzaminohen me kujdes dhe duhet të gjenden çiftet që janë të lidhur me njëri-tjetrin.

Rreshtat nga tabelat përkatëse, zakonisht, mbledhen së bashku për përpunim dhe është mençuri që ato të ruhen së bashku në të njëjtën fajllë.

Caktimi i strukturave të memories

Plani i memories duhet të përcaktojë ku do të vendosen të dhënat në mjetin e memories fizike, fajllat fizike, si dhe plani për caktimin e çdo tabele në fajllin specifike, por edhe mënyra se si çdo fajllë fizike duhet të ndahet në blloqe të dhënash.

Përfundimi i modelit fizik

Ky është hapi final që rishikon dhe konfirmon plotësimin e aktiviteteve dhe detyrave të mëparshme. Triologjia konceptuale/logjike/fizike, e përshkruar në këtë pjesë, sipas opinionit të autorit, është praktika më e mirë për modelimin e DWH-së. DWH siguron platformën për analizën OLAP.

Menaxhimi i particioneve

Particionet e lejojnë një tabelë (dhe treguesit e saj) të ndahen fizikisht në *mini-tabela* për qëllime administrative dhe për të përmirësuar performancën e kërimit. Dobia finale e copëzimit të një tabele është se një kërkim, që kërkon një muaj të dhënash nga një tabelë që ka dhjetë vjet të dhënash, mund të shkojë direkt në partitionin e tabelës që përmban të dhënat për muajin pa skanuar të dhënat e tjera. Particionet e tabelave mund të përmirësojnë në mënyrë dramatike performancën e kërimit në tabela të mëdha faktesh. Particionet e një tabele janë të mbuluara, të fshehura nga përdoruesit. Vetëm stafi DBA dhe ETL duhet të jetë në dijeni të particioneve. Meqenëse përdoruesit, zakonisht, kufizohen në kolonat në dimensionin e datës, ju duhet ta ndani tabelën e fakteve në kyçin që i bashkohet dimensionit të datës, në mënyrë që optimizuesi të njohë kufizimin. Tabelat që janë ndarë nga një interval date, zakonisht, ndahen sipas vitit, tremujorit ose muajit. Faktet shumë voluminoze mund të

ndahen sipas javës ose ditës. Zakonisht, ndërtuesi i DWH-së punon me stafin DBA për të përcaktuar strategjinë më të mirë të copëzimit në bazë tabelë-pas-tabele. Stafit ETL duhet t'i tregohet për çdo particion table që duhet të mirëmbahet (Kimball&Caserta, 2004).

3.8 Përgatitja e një strategjie indeksimi

Plani i indeksimit duhet të tregojë indeksat për çdo tabelë. Për më tepër, për çdo tabelë, duhet të përcaktohet sekuenca në të cilin do të krijohen indeksat. Gjithashtu, indeksat që janë pranuar të ndërtohen në çdo rast të parë të databazës duhet të përshkruhen. Ofruesit e DBMS-së ofrojnë një sërë teknikash indeksimi. Zgjedhja nuk është më e kufizuar vetëm në fajllat e indeksit në sekuencial. Të gjithë ofruesit mbështetin indeksat B-tree për gjetjen efikase të të dhënave. Një mundësi tjetër është indeksi bit-map. Tabelat e dimensioneve duhet të kenë një indeks unik në çelësin kryesor me një kolonë. Kimball, gjithashtu, rekomandon një tregues B-tree në kolonat e attributeve me kardinalitet të lartë të përdorura për kufizimet dhe vendosjen e indekseve bitmap në të gjitha atributet me kardinalitet të mesëm e të lartë. Në rastin e tabelave të fakteve, çelësi kryesor është pothuajse gjithmonë një nënset i çelësave të huaj.

Në këtë rast është e mençur të vendoset një indeks i vetëm, i përqendruar në dimensionet kryesore të tabelës së fakteve. Pasja e një çelësi date ndihmon, po ashtu, në përmirësimin e performancës. Meqenëse mund të përdoret më shumë se një indeks në të njëjtën kohë, duke zgjidhur një kërkim, mund të ndërtohen, gjithashtu, indekse të veçanta në çelësat e huaj të pavarur në tabelën e fakteve. Indeksi është, ndoshta, mekanizmi më i efektshëm për përmirësimin e performancës. Sidoqoftë, pasja e një numri shumë të madh indeksesh ngadalëson ngarkimin e të dhënave në magazinë, veçanërisht gjatë ngarkimeve fillestare.

Ky problem mund të zgjidhet duke rrëzuar indeksat para se të bëhen punët e ngarkimit. Në këtë rast, ato nuk do të krijojnë hyrje indeksi gjatë procesit të ngarkimit, por ndërtimi i fajllave të indeksit mund të bëhet pasi të përfundojë procesi i ngarkimit.

Ka edhe diçka tjetër që duhet mbajtur në mend e ai është fakti se tabelat e mëdha (me miliona rreshta) nuk mund të mbështetin shumë indekse. Kur një tabelë është shumë e madhe, pasja e më shumë se një indeksi vetë mund të shkaktojë vështirësi. Nëse është i nevojshëm më shumë se një indeks, tabela duhet të ndahet para se të përcaktohen më shumë indekse. Kolonat më të përshtatshme për indeksimin janë ato që përdoren shpesh për të kufizuar kërkimet. Sidoqoftë, praktika më e mirë është të fillohet me indeksat vetëm në çelësat kryesorë dhe të huaj të çdo table. Pastaj, pas monitorimit të performancës me kujdes mund të shtohen disa indeks të tjerë (Kimball&Caserta, 2004).

3.8.1 Indeksi i datawarehouse

Në një sistem kërkim-qendror, si mjedisi i DWH-së, dominon nevoja për t'i përpunuar kërkesat më shpejt. Për përdoruesin në një seancë analize, që kalon nëpër një seri të shpejtë kërkesash komplekse, ju duhet të përshtatni ritmin e rezultateve të kërkimit me shpejtësinë e mendimit. Mes metodave të ndryshme për të përmirësuar performancën, indeksimi radhitet shumë lart. Çfarë lloje indeksesh duhet të ndërtoni në DWH?

Ofruesit DBMS ofrojnë një sërë mundësish. Zgjedhja nuk është e kufizuar vetëm në fajllat e indeksit në sekuencial. Të gjithë ofruesit mbështetin indeksat B-Tree për gjetjen efikase të të dhënave. Një mundësi tjetër është indeksi bitmap. Siç do shohim më tej në këtë pjesë, kjo teknikë indeksimi është shumë e përshtatshme për mjedisin e DWH-së. Disa ofrues po e

shtojnë fuqinë e indeksimit sipas kërkesave specifike. Këto përfshijnë indekse mbi tabela të copëzuara dhe tabela të organizuara me indekse.

3.8.2 Përmbledhja e indeksimit

Le të marrim në konsideratë teknikën e indeksimit nga këndvështrimi i DWH-së. Tabelat e të dhënave janë vetëm të lexueshme. Kjo veçori do të thotë se pothuajse asnjëherë nuk i përditësoni regjistrimet apo t'i fshini regjistrimet. Dhe, regjistrimet nuk futen në tabela pas ngarkimeve. Kur bëhen shtime, përditësime ose fshirje, krijohet ngarkesë shtesë për manipulimin e fajllave të indeksit. Por, në një DWH nuk ndodh kështu.

Kështu, mund të krijohen një numër indeksesh për çdo tabelë. Shumica e indeksimit bëhet në tabelat e dimensioneve. Në përgjithësi, do të shikohen shumë më tepër indekse në një DWH sesa në një sistem OLTP. Kur një tabelë rritet në volum, indekset, po ashtu, rriten në madhësi, duke kërkuar më shumë memorie. Rregulli i përgjithshëm është se numri maksimal i indeksëve ndryshon anasjelltas me madhësinë e tabelës. Numrat e mëdhenj të indeksëve ndikojnë në procesin e ngarkimit, pasi indekset krijohen për regjistrimet e reja në atë moment. Andaj, duhet të balancohen faktorët e ndryshëm dhe të vendosen mbi numrin e indeksëve për tabelë.

Në pjesën e mbetur të kësaj pjese do t'i studiojmë më thellë teknikat specifike të indeksimit. Para se ta bëjmë këtë, vini re parimet e përgjithshme të mëposhtme. Kur keni një numër të madh indeksesh, ngarkimi i të dhënave në DWH ngadalësohet ndjeshëm. Kjo ndodh, meqë kur shtohet çdo regjistrim në tabelën e të dhënave, duhet të krijohet çdo indeks përkatës. Problemi është më i madh gjatë ngarkesave fillestare. Mund ta adresoni këtë problem, duke

i rrëzuar indekset para se t'i bëni punët e ngarkimit. Duke bërë këtë, punët e ngarkimit nuk do të krijojnë hyrje indeksi gjatë procesit të ngarkimit.

Pasi të mbarojë procesi i ngarkimit, mund të bëni punë të veçanta për të ndërtuar fajllat e indeksit. Ndërtimi i fajllave të indeksit kërkon kohë të madhe, por jo aq sa krijimi i hyrjeve të indeksit gjatë procesit të ngarkimit.

3.8.3 Indeksimi për tabelat e mëdha

Tabelat e mëdha, me miliona rreshta, nuk mund të mbështetin shumë indekse. Kur një tabelë është shumë e madhe, nevoja e më shumë se një indeksi vetë mund të shkaktojë vështirësi. Nëse duhet të kemi më shumë indekse për tabelë, merret parasysh ndarja e tabelës para se të përcaktohen më shumë indekse.

3.8.4 Indeksimi vetëm me lexim

Siç dihet, në procesin e gjetjes së të dhënave, regjistrimi i indeksit më parë lexohet dhe pastaj bëhet leximi i të dhënave përkatëse. DBMS zgjedh indeksin më të mirë mes shumë indekseve. Le të themi se DBMS përdor një indeks të bazuar në katër kolona në një tabelë dhe shumë përdorues në mjedisin tuaj kërkojnë të dhëna nga ato katër kolona dhe një kolonë më shumë në tabelë. Andaj, si bëhet gjetja e të dhënave? DBMS përdor këtë indeks për të gjetur regjistrimin përkatës të të dhënave.

Duhet, të paktën, dy veprime input-output (I/O). Në këtë rast, DBMS duhet të gjejë regjistrimin e të dhënave vetëm për një kolonë shtesë. Në raste të tilla merret parasysh shtimi i asaj kolone ekstra në indeks. DBMS do të lexojë indeksin dhe do të zbulojë se të gjitha

informacionet që duhen ndodhen në vetë regjistrimin e indeksit, prandaj nuk do ta lexojë regjistrimin e të dhënave në mënyrë të panevojshme.

3.9 Efektet e indekseve

DWH përdoret sepse është shumë më i shpejtë dhe më i besueshëm se sistemi i transaksionit.

Avantazhi i shpejtësisë, që ofron DWH, është për shkak të një numri veçorish kyçe:

- Modeli i të dhënave dimensionale, i cili lejon indeksimin e ndërtuar qëllimisht të dimensioneve dhe fakteve;
- Strategjia e indeksit agresiv;
- Regjistrimet e ruajtura fizikisht;
- Paralelizmi i kërkimit.

Teknikat e ndërtimit dhe dobisë të të dhënave dimensionale janë treguar me hollësi në këtë literaturë dhe janë në dispozicion në një gamë tjetër të gjerë. Në këtë pjesë do të flasim shkurtimisht rreth indekseve. Indekset janë shtylla e kohës kërkesë-përgjigje në DWH.

Çdo kërkim që paraqitet në databazë përdor, të paktën, një indeks. Fatkeqësisht, në atë sa indekset ndihmojnë përdoruesit të kërkojnë DWH, stafi ETL është i ngarkuar me barrën për të menaxhuar indekset ekzistuese gjatë procesit ETL.

Menaxhimi i indekseve llogaritet për një pjesë të konsiderueshme të shumicës së proceseve ETL. Para se të zhytemi në teknikat për menaxhimin e indekseve gjatë ngarkesës së DWH-së, është e rëndësishme që t'i shqyrtojmë llojet e ndryshme të indekseve që janë në dispozicion në shumicën e databazave.

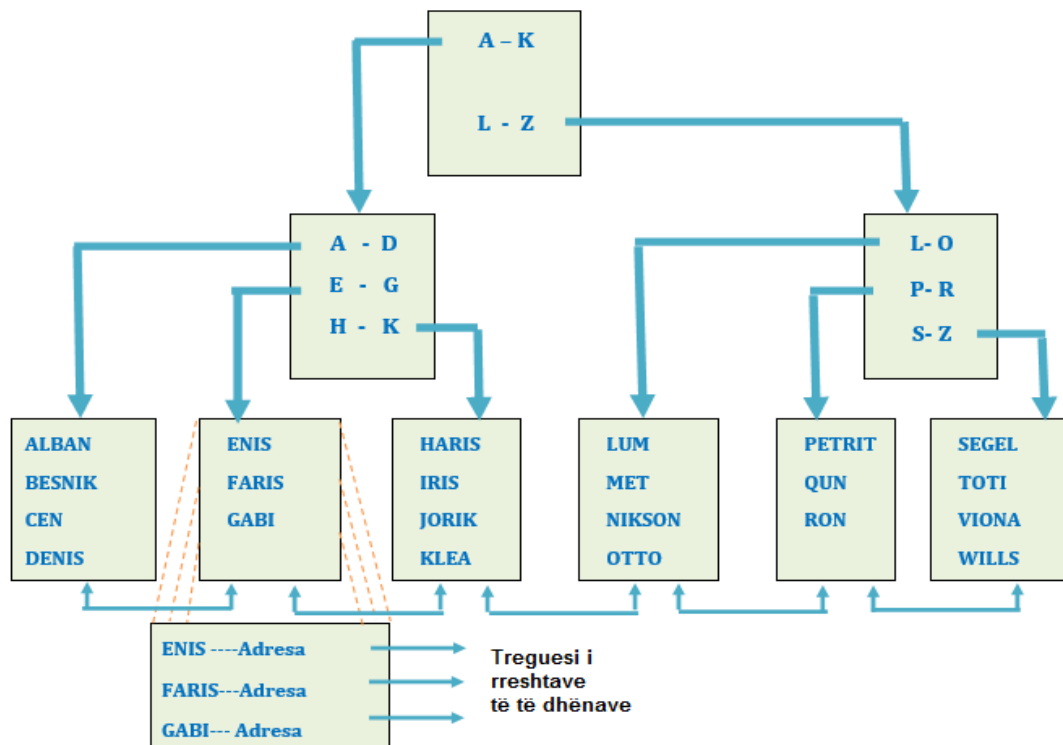
Kryesisht, do të gjeni dy lloje indeksesh dhe është e rëndësishme të kuptohen dallimet midis dy llojeve (Kimball&Caserta, 2004).

3.9.1.Indeksi B-Tree

Shumica e sistemeve të menaxhimit të databazave kanë teknikën e indeksit B-Tree, si metodë standarde të indeksimit. Kur kodoni deklarata, duke përdorur gjuhën e përkufizimit të të dhënave të softuerit të databazës për të krijuar një indeks, sistemi krijon indeksin B-Tree. DBMS, gjithashtu, krijon automatikisht indekse mbi vlerat kryesore kyçe.

Teknika e indeksit B-Tree është superiore ndaj teknikave të tjera, për shkak të shpejtësisë së gjetjes së të dhënave, lehtësisë së mirëmbajtjes dhe thjeshtësisë. Figura 3.14 tregon një shembull të një indeksi B-Tree.

Figura 3.14 Struktura e një indeksi B-Tree (Ponniah, 2010).



Indeksi përbëhet nga një strukturë B-Tree (një pemë binare e balancuar), e bazuar mbi vlerat e kolonës së indeksuar. Në shembullin përkatës kolona e indeksuar është “emri”. Kjo B-Tree krijohet duke përdorur të gjithë emrat ekzistues, të cilat janë vlerat e kolonës së indeksuar.

Blloqet në nivelin më të ulët tregohen në rreshtat në tabelën e të dhënave. Nëse një kolonë në një tabelë ka shumë vlera unike, atëherë selektiviteti i kolonës thuhet të jetë i lartë. Në një tabelë dimensionit, territori, kolona për “qytetin” përmban shumë vlera unike. Prandaj, kjo kolonë është shumë selektive.

Indekset B-Tree janë më të përshtatshmet për kolonat shumë selektive. Shkaku se vlerat e nyjave do të jenë unike, ato çojnë në rreshta të qartë të dhënash dhe jo në një zinxhir rreshtash. Por, nëse një kolonë e vetme nuk është shumë selektive, si mund ta përdorni indeksimin B-Tree në këto raste? Për shembull, kolona e parë e emrit në një tabelë punonjësish nuk është shumë selektive. Ekzistojnë shumë emra të njëjtë. Mund ta përmirësoni selektivitetin duke lidhur emrin me mbiemrin.

Kombinimi është shumë më selektiv nëse krijohet një indeks B-Tree, i lidhur mbi të dy kolonat së bashku. Indekset rriten në përpjesëtim të drejtë me rritjen e tabelës së indeksuar të të dhënave. Kurdoherë që indекset përmbajnë një lidhje të disa kolonave, ato kanë prirjen të rriten në madhësi.

Derisa DWH merret me volume të mëdha të dhënash, madhësia e fajllave të indeksit mund të jetë shkak për t'u shqetësuar. Çfarë mund të themi për selektivitetin e DWH-së? A janë shumica e kolonave shumë selektive? Jo, tamam. Nëse i kontrolloni kolonat në tabelën e dimensioneve, do të vini re se një numër kolonash që përmbajnë indekse të data-Tree me selektivitet të ulët nuk punojnë mirë me të dhënat e atyre që kanë selektivitet të ulët. Cila është alternativa? Kjo na çon në një lloj tjetër teknike indeksimi (Ponniiah, 2010).

Indekset B-tree janë optimalë për kolonat me kardinalitet shumë të lartë. Me kardinalitet të lartë nënkuptojmë numrin e vlerave të ndryshme. Struktura B-tree përdor një teknikë jashtëzakonisht të efektshme në rastin e analizimit të të dhënave, për të gjetur një vlerë të caktuar.

Andaj, treguesit B-tree janë të shkëlqyeshëm për të optimizuar kërkimet e njohura, por janë paksa joelastikë për të mbështetur mjediset ad-hoc, si, për shembull, DWH-në. Me këtë treguesit B-tree janë përcaktuar si joelastikë, sepse nuk mund të kombinohen me kolonat e indeksuara për të krijuar në formë dinamike të indekseve të përbëra për të zgjidhur kërkime të reja dhe të papritura.

Të gjitha indekset duhet të bëhen më përpara. DBA duhet të përpiqet t'i gjejë se cilat kolona mund të kufizohen. Për më tepër, rendi në të cilat pozicionohen kolonat përcakton nëse ato po përdoren apo jo. Rezultati është se stafi DBA duhet të bëjë shumë indekse B-tree të përbëra, shumë prej të cilave përmbajnë të njëjtat kolona në rende të ndryshme.

3.9.2 Indeksi bitmap

Indekset bitmap janë idealisht të përshtatshëm për të dhënat me selektivitet të ulët. Një bitmap është një seri bitesh të radhitur, një për çdo vlerë të qartë të kolonës së indeksuar. Supozohet se kolona për ngjyrën ka tri ngjyra të qarta: e bardha, bezha dhe e zeza.

Çdo hyrje në bitmap përmban tre bite. Le të themi se biti i parë i referohet të bardhës, i dyti bezhës dhe i treti të zezës. Nëse një kolonë Mbiemri ka ngjyrë të bardhë, hyrja bitmap për kolonën Mbiemri përbëhet nga tre bite, ku biti i parë është vendosur në 1, biti i dytë është vendosur në 0 dhe biti i tretë është vendosur në 0.

Nëse kolona Mbiemri ka ngjyrë bezhë, hyrja bitmap për atë kolonë Mbiemri përbëhet nga tre bite, ku biti i parë është vendosur në 0, biti i dytë është vendosur në 1 dhe biti i tretë është vendosur në 0. Në vazhdim japim shembullin e indeksit bitmap, i treguar në figurën 3.15

Figura 3.15 Shembull i Indeksi bitmap (Ponniah, 2010).

Të dhënat e Agjencionit të regjistrimit civil

ADRESA ID RRESHTAT	MBIEMRI	EMRI	DATE LINDJA	KOMUNA	ADRESA
00001BFE.0012.0111	KRASNIQI	JORIK	04.06.2009	PRISHTINE	"Lidhja e Prizrenit"
00001BFE.0013.0114	HALILI	LUM	25.03.2011	PRIZREN	"Bajram Bahtiri"
00001BFF.0012.0115	GASHI	BUKURIJE	25.08.1999	MITROVICE	"Ismail Dumoshi"
00001BFF.0012.0138	GASHI	ISA	23.04.1988	PEJE	"Lidhja e Prizrenit"
00001BFF.0012.0145	HALILI	RRON	14.05.2007	PEJE	"Ismail Dumoshi"
00001BFF.0012.0157	KRASNIQI	VIONA	28.10.2005	PRISHTINE	"Ismail Kadare"
00001BFF.0014.0165	GASHI	RRON	23.04.1989	PEJE	"Lidhja e Prizrenit"

Indeksi Bitmap për kolonën Mbiemri
Caktohet për bitat: Krasniqi, Halili, Gashi

ADRESA ID RRESHTAVE	BITMAP
00001BFE.0012.0111	001
00001BFE.0013.0114	010
00001BFF.0012.0115	100
00001BFF.0012.0138	100
00001BFF.0012.0145	010
00001BFF.0012.0157	001
00001BFF.0014.0165	100

Indeksi Bitmap për kolonën Komuna
Caktohet për bitat: Prishtinë, Prizren, Pejë

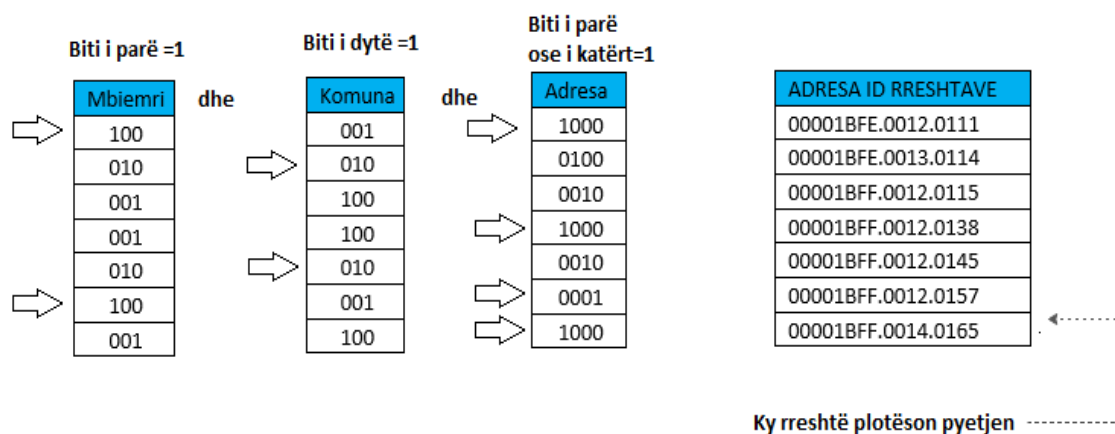
ADRESA ID RRESHTAVE	BITMAP
00001BFE.0012.0111	100
00001BFE.0013.0114	010
00001BFF.0012.0115	001
00001BFF.0012.0138	001
00001BFF.0012.0145	010
00001BFF.0012.0157	100
00001BFF.0014.0165	001

Indeksi Bitmap për kolonën Adresa
Caktohet për bitat: "Lidhja e Prizrenit", "Bajram Bahtiri", "Ismail Dumoshi", "Ismail Kadare"

ADRESA ID RRESHTAVE	BITMAP
00001BFE.0012.0111	1000
00001BFE.0013.0114	0100
00001BFF.0012.0115	0010
00001BFF.0012.0138	1000
00001BFF.0012.0145	0010
00001BFF.0012.0157	0001
00001BFF.0014.0165	1000

Figura tregon një ekstrakt të tabelës së Agjencisë së regjistrimit civil dhe të indekseve bitmap për tri kolona të ndryshme. Vini re se si çdo hyrje në një indeks përmban bitet e radhitur për të treguar vlerat e qarta në kolonë. Krijohet një hyrje për çdo rresht në tabelën bazë. Çdo hyrje mban adresën e rreshtit të tabelës bazë. Duke marrë në konsideratë një kërkesë kundrejt tabelës së Agjencisë së regjistrimit civil në shembullin më sipër, zgjidhni rreshtat nga tabela e Agjencisë së regjistrimit civil ku kolona Mbiemri është “Rondele” dhe ngjyra është “bezhë” dhe ndarja është “Lindje” ose “Jug”. Kjo është paraqitur në figurën 3.16

Figura 3.16 Indeksi bitmap: rikthim i të dhënave



Pika e fortë e kësaj teknike qëndron në efektshmërinë e saj, duke përdorur kallëzuesit mbi kolonat me selektivitet të ulët në kërkimet. Indekset bitmap marrin shumë më pak hapësirë se indекset B-Tree për kolonat me selektivitet të ulët.

Në një DWH shumë aksese të dhënash bazohen në kolonat me selektivitet të ulët. Gjithashtu, analiza, duke përdorur skenarët "po nëse", kërkon kërkime që përfshinë disa kallëzues. Do të zbulohet se indекset bitmap janë më të përshtatshëm për një mjedis DWH sesa për një sistem OLTP. Nga ana tjetër, nëse prezantohen vlera të reja për kolonat e indeksuara, indекset bitmap duhet të rindërtohen.

Një disavantazh tjetër lidhet me nevojën për të aksesuar tabelat e të dhënave gjithë kohës pasi janë aksesuar indекset bitmap. Indекset B-tree nuk kërkojnë akses table nëse informacionet e kërkuara janë tashmë në indeks (Ponniah, 2010).

3.9.3 Treguesit bitmap

Treguesit bitmap funksionojnë krejt ndryshe nga treguesit B-tree. Treguesit bitmap janë të përshtatur më mirë për kolonat me kardinalitet të ulët. Shumë tregues bitmap me një kolonë mund të bashkohen në formë dinamike së bashku, për të krijuar indekse të përbëra për të mbështetur kërkimet ad-hoc. Për shkak të elasticitetit të tyre, është praktikë e zakonshme të krijohen indekse bitmap me një kolonë në çdo çelës zëvendësues në tabelat e fakteve në datawarehouse (DWH). Tani të kthehemi në mënyrën se si indekset ndikojnë në transaksionet DML. Çdo e dhënë e një indeksi B-tree përmban një dhe vetëm një numër rreshti që tregon regjistrimin përkatës në tabelën bazë.

Përndryshe, indekset bitmap përmbajnë një gamë numrash, rreshtash për çdo vlerë në indeks. Nëse një vlerë përditësohet, çdo regjistrim që i përket gamës së numrave të rreshtave kyçet. Së fundi, çdo herë që përditësohet një regjistrim, vendoset një barrë shumë e madhe në databazë për të menaxhuar të gjitha kyçjet e rreshtave që ndodhin. Fatkeqësisht, tabelat e fakteve, zakonisht, përbëhen vetëm nga indekse bitmap dhe bërja e përditësimeve masive krijon vonesa masive për shkak të ngarkesës së madhe dhe performancës shumë të dobët.

3.9.4 Indeksi i tabelës së fakteve

Çelësi kryesor i tabelës së fakt_IMK, siç paraqitet në figurën 3.4, përbëhet nga çelësat kryesorë të të gjitha dimensioneve të lidhura. Nëse është tabela me pesë dimensionë: Dim_Data, Dim_Kohë, Dim_grupi_gjakut, Dim_instit_shendetsor dhe Dim_adresa, atëherë çelësi i plotë kryesor i tabelës së fakteve është seria e çelësve kryesorë të tabelave të Data, Kohë, Grupi i gjakut, Instituti shëndetësor dhe Adresa. Kolonat e tjera janë metriket, të tilla si njësia, sekonda, ora, e kështu me radhë.

Këto janë llojet e kolonave që duhet të merren në konsideratë për indeksimin e tabelave të fakteve. Kur bëhet planifikimi i krijimit të indekseve për tabelat e fakteve?

- Nëse DBMS nuk krijon një indeks mbi çelësin kryesor, krijohet qëllimisht një indeks B-Tree në çelësin e plotë kryesor.
- Caktohet me kujdes rendi i elementëve kyçe individuale në çelësin e plotë të lidhur për indeksim. Në rendin e lartë të çelësit të lidhur vendosen çelësat e tabelave të dimensioneve të referuar shpesh gjatë kërkimit.
- Rishikohen komponentët individualë të çelësit të lidhur. Më pas krijohen indekset mbi kombinimet, bazuar në kërkesat e përpunimit të kërkimeve.
- Nëse DBMS mbështet kombinimet inteligjente të indekseve për akses, atëherë mund të krijohen indekse mbi çdo komponentë të veçantë të çelësit të lidhur.
- Mos anashkalimi i mundësive të indeksimit të kolonave që përmbajnë attribute metrike. Për shembull, nëse shumë kërkime kërkojnë njësinë brenda gamave të dhëna, atëherë kolona Dim_instituti_shendetsor është një kandidat për indeksimin.
- Indeksi bitmap nuk aplikohet për tabelat e fakteve. Pothuajse nuk ka kolona me selektivitet të ulët.

3.9.5 Indeksi i tabelave të dimensioneve

Kolonat në tabelat e dimensioneve përdoren në kallëzuesit e kërkimeve. Kolonat Dim_Kohë, Dim_Data dhe Dim_instituti_shendetsor, janë tri tabela të ndryshme dimensionesh që janë kandidate për indeksim. Më pas kontrolloni kolonat e çdo table dimensionesh me kujdes dhe planifikoni indekset për këto tabela. Me këtë proces nuk arrijmë përmirësim të

performancës, duke indeksuar kolonat në tabelat e fakteve, por kolonat në tabelat e dimensioneve ofrojnë mundësi shumë të mëdha për të përmirësuar performancën me anë të indeksimit.

Ja disa këshilla mbi indeksimin e tabelave të dimensioneve:

- Krijimi i një indeksi unik B-Tree mbi çelësin kryesor me një kolonë.
- Ekzaminohen kolonat që përdoren, zakonisht, për të kufizuar kërkimet. Këto janë kandidatë për indekse bitmap.
- Kërkohen kolona që përdoren shpesh së bashku në tabela dimensionesh të mëdha. Përcaktohen se si këto kolona mund të rregullohen dhe përdoren për të krijuar indekse me shumë kolona. Kolonat që përdoren më shpesh ose kolonat në nivele hierarkike më të larta në tabelën e dimensioneve, vendosen në rendin e lartë të indekseve me shumë kolona. (Ponniah, 2010).

3.9.6 Rëndësia e indeksimit

Një strategji indeksimi e planifikuar mirë lejon akses të shpejtë dhe eficient në të dhënat bazë. Indekset mund të krijohen në tabela ose views dhe idealisht lejojnë DWH-në të vendosë dhe menaxhojë të dhënat në mënyrë më efektive. Kur eficienta përmirësohet, koha në dispozicion që çdo operacion kryen reduktohet së bashku me koston e lidhur me kryerjen e operacionit. Dizajnimi i indeksit, zakonisht, kryhet gjatë zhvillimit të aplikimit të bazës së të dhënave. Arsyeja është te aftësia për të krijuar indekse efektive dhe e bazuar në të kuptuarit se si “application queries” është koduar dhe se si e dhëna është ruajtur (stored) në bazën e të dhënave. Megjithatë, indekset kërkojnë menaxhim pasi “database application” është

vendosur dhe se si modelet e përdorura shfaqen ose ndryshojnë. Menaxhimi dhe optimizimi i indekseve është një proces në vazhdim dhe lejon përmirësime të performancës së mundshme pa kërkuar ndryshime në skemën themelore (Dasho, 2014).

3.10 Konkluzione

Të dhënat e tërmeteve, vërshimeve të detajuara apo edhe të stuhive e temperaturave ekstreme, janë fusha më e nxehtë për eksplorimin e njohurive në menaxhimin e emergjencave dhe të fatkeqësive natyrore. Për këtë arsye, modelimi i një data mart-i për këto të dhëna është shumë interesant për t'u zhvilluar.

Ky kapitull përmend të gjithë hapat, në detaje, që ndjekim në ndërtimin e një data mart-i mbi këto të dhëna, po ashtu edhe Kimball mbështet fort zhvillimin e data mart-ëve të pavarur, si një strategji të zbatueshme për një zhvillim në faza të sistemeve për mbështetjen e vendimmarrjes. Duke u bazuar pikërisht në rekomandimet e Kimball-it, u përzgjedh arkitektura me dy shtresa, për të përmbushur veçanërisht tiparin e ndarjes, ku të dhënat fillimisht do të ngarkohen në një zonë Stage, ku edhe do të transformohen, për t'u ngarkuar më pas në skemën e data mart-it.

Më pastaj, shqyrtojmë edhe njëërën nga teknikat kyçe, me të cilën janë ndërtuar dhe implementuar strukturat e të dhënave brenda DWH-së, e kjo ka të bëjë me modelimin dimensional. Secila prej proceseve të menaxhimit të emergjencave, si në rastin e Institutit të Mjekësisë së Kosovës (IMK), mund të tregohet nga një model dimensional, i cili përbëhet nga një tabelë faktesh, e që përmban matjet numerike të rrethuara nga një sërë tabelash dimensionesh që përmbajnë kontekstin tekstual.

Duke pasur parasysh shumicën e teknikave për ndërtimin e DWH /data mart, analizimi i kërkesave cilësohet njëra nga prioritet parësore të procesit të ndërtimit të DWH/data mart. Duke ndjekur modelin Kimball, identifikimi i elementeve bazë të modeleve shumëdimensionale mund të përkufizohet përmes sistemit matricë e lidhjes dimension/fakt/matje.

Të gjithë ofruesit mbështetin indekset B-Tree për gjetjen efikase të të dhënave. Një mundësi tjetër është indeksi bitmap. Shumica e sistemeve të menaxhimit të databazave e kanë teknikën e indeksit B-Tree si metodë standarde të indeksimit.

KAPITULLI IV: APLIKIMI I SISTEMEVE TË DWH-së NË RASTIN E EMERGJENCAVE DHE FATKEQËSIVE NATYRORE

4.1 Hyrje

Në ditët e sotme, më shumë se kurrë, shohim nevojën për të ndërtuar një sistem institucional të koordinuar me të gjitha burimet e disponueshme dhe kapacitetet e planifikuara për një reagim të menjëhershëm në raste të katastrofave natyrore dhe incidenteve të tjera. Është koha të punojmë në mënyrë më gjithëpërfshirëse për t'i koordinuar përpjekjet tona në menaxhimin e përgjigjeve kundrejt situatave të emergjencës, përfshirë katastrofat natyrore, incidentet me rrezik të lartë, terrorizmi dhe incidente të tjera të shkaktuara nga qeniet njerëzore.

Me fjalë të tjera, ky sistem është një qasje sistematike drejt udhëheqjes dhe përgatitjes më të mirë të të gjitha agjencive dhe departamenteve përkatëse në të gjitha nivelet e qeverisjes dhe të institucioneve të tjera joqeveritare, që t'u shërbejnë qytetarëve si është më së miri në këto emergjenca, pavarësisht kompleksitetit, shtrirjes, vendit apo shkakut. Incidentet dhe katastrofat natyrore nuk njohin kufij. Implementimi i këtij sistemi ofron një zgjidhje të koordinuar komunikimi dhe ushtrimi të avancuar në fushë në nivel shtetëror. Për më tepër, ky sistem do të çojë në bashkëpunim më të mirë rajonal dhe ndërkombëtar (SIME, 2010).

4.2 Gjendja e emergjencave dhe fatkeqësive natyrore në vendin tonë

Tre rreziqet kryesorë natyrorë që Kosova u nënshtrohet janë: tërmetet, vërshimet dhe zjarret pyjore. Veç këtyre rreziqeve, rrezik të konsiderueshëm, gjithashtu, paraqesin rrëshqitjet e tokës, thatësira, reshjet e mëdha të borës dhe plasja e pendëve të rezervuarëve të ujit.

Për më tepër, ekzistojnë edhe fatkeqësi mjedisore të shkaktuara nga ish-fabrika e plumbit në Mitrovicë, në të cilën erërat e fuqishme dhe stuhitë kontribuojnë në ndotjen e ujit për pije dhe kontaminimin e ushqimit me pasoja, siç janë, sëmundjet dhe epidemitë. Këto rreziqe përbëjnë një kërcënim të përhershëm për qytetarët e Kosovës.

Nisur nga kjo, për qëllim analize, është shqyrtuar gjendja aktuale lidhur me kapacitetet e komunave për intervenim në rastet emergjente, qoftë në kuadrin ligjor, qoftë në kapacitetet teknike e njerëzore që disponojnë institucionet relevante, si dhe përgatitjen e vendit tonë për menaxhimin e emergjencave. Kur kësaj i shtohet edhe mungesa e analizave të kësaj problematike, qoftë nga institucionet, qoftë nga mekanizmat tjerë joqeveritarë, e konsiderojmë të nevojshëm vështtrimin e kapaciteteve institucionale dhe të nevojave të institucioneve relevante për reagim në situata emergjente (Forumi për siguri, 2013).

4.2.1 Analizimi i të dhënave në rastin e tërmeteve

Për këtë dëshmojnë qartë tërmetet që kanë ndodhur në të kaluarën e largët a të afërt. Sipas Institutit Sizmologjik në Ministrinë e Zhvillimit Ekonomik (MZHE), çdo ditë ndodhin tërmete, por me një intensitet shumë të ulët, kurse tërmete që ndodhin mbi 3.5 shkallë të Riterit, paraqiten në tabelën e mëposhtme.

Tabela 4.1 Të dhëna historike të rasteve të tërmeteve (Instituti Sizmologjik, 2012).

Emri i regjionit	Data	Magnituda
Prizren	16 qershor 1456	6.0
Pejë	11 nëntor 1662	6.0
Ferizaj	26 shkurt 1755	6.1
Kopaonik	18 maj 1980	5.7
Gjilan	24 prill 2002	5.7
Istog	10 mars 2010	5.2
Ferizaj-Viti	10 gusht 1921	6.1

Në tabelën 4.1. jepen të dhëna historike të rasteve të tërmeteve, por për këto ngjarje të fatkeqësive natyrore duhet të kemi të dhëna në kohë reale.

Nga Instituti Sizmologjik kemi të dhëna kryesisht mujore “Buletini Mujor”, siç paraqiten në tabelat e mëposhtme.

Tabela 4.2 Instituti Sizmologjik i Kosovës (ISK), “Buletini mujor” (Mars 2010) (Instituti Sizmologjik, 2012).

Data	Koha	Vendi	M	Lat	Long	Depth
10. 3. 2010	13:38:03.1	Kosovë	5.2	42.76	20.62	5.22 km
	13:44:41.8	Kosovë	1.8	42.75	20.63	5.00 km
	14:13:20.9	Kosovë	1.9	42.69	20.58	18.83 km
	15:32:43.2	Kosovë	2.2	42.65	20.95	Unknown
11. 3. 2010	04:33:54.9	Kosovë	1.5	42.73	20.58	16.06 km

	08:59:19.4	Kosovë	3.3	42.74	20.60	Unknown
12. 3. 2010	14:28:16.1	Kosovë	1.3	42.47	20.67	Unknown
13. 3. 2010	01:54:00.3	Kosovë	1.8	42.72	20.57	6.28 km
	13:53:44.5	Kosovë	1.5	42.76	20.61	15.28 km
15. 3. 2010	21:37:40.5	Kosovë	2.1	42.44	20.90	14.04 km
18. 3. 2010	02:48:10.9	Kosovë	1.6	42.71	20.60	9.99 km
26. 3. 2010	07:46:10.8	Kosovë	1.8	42.71	20.58	17.18 km
	07:46:33.8	Kosovë	1.8	42.71	20.57	17.10 km

Tabela 4.3 Instituti Sizmologjik i Kosovës (ISK), “Buletini mujor” (Nëntor 2013) (Instituti Sizmologjik, 2012).

DATA	KOHA	LAT	LONG	DEPTH	M	VENDI
1.11.2013	14:11:17.8.	42.644	21.02	2 km	1.6	Kosovë
1.11.2013	14:50:25.2.	43.085	20.89	Unknown	1.5	Kosovë
1.11.2013	18:42:07.3.	42.475	20.41	16 km	1.6	Kosovë
16.11.2013	07:42:34.4.	42.869	21.025	Unknown	2.9	Kosovë
16.11.2013	07:44:01.0.	42.871	21.005	1 km	2.4	Kosovë
17.11.2013	08:14:14.0.	42.864	21.011	10 km	3.7	Kosovë
17.11.2013	14:27:52.0.	42.102	20.656	8 km	2.3	Kosovë
17.11.2013	14:48:19.3.	42.837	20.991	Unknown	1.6	Kosovë
17.11.2013	23:58:53.2.	42.859	21.001	4 km	2.8	Kosovë
18.11.2013	11:23:39.7.	42.765	21.055	Unknown	1.5	Kosovë

18.11.2013	13:15:03.3.	42.863	21.014	4 km	4.8	Kosovë
18.11.2013	13:28:05.8.	42.901	21.023	1 km	1.5	Kosovë
18.11.2013	13:30:04.2.	42.883	21.031	2 km	1.6	Kosovë
18.11.2013	13:35:50.7.	42.849	21.001	1 km	2	Kosovë
18.11.2013	13:42:25.2.	42.877	21.027	10 km	3.5	Kosovë
18.11.2013	14:42:25.2.	43.209	20.698	Unknown	1.8	Kosovë
18.11.2013	15:24:27.6.	42.905	21.064	6 km	1.8	Kosovë
18.11.2013	16:49:24.0.	42.857	21.016	Unknown	1.7	Kosovë
18.11.2013	20:31:07.5.	42.838	20.998	Unknown	1.5	Kosovë
18.11.2013	23:28:34.3.	42.386	21.561	Unknown	1.5	Kosovë
19.11.2013	01:37:14.5.	42.86	21.052	Unknown	1.6	Kosovë
19.11.2013	03:39:49.9.	42.865	21.061	5 km	2.2	Kosovë
19.11.2013	06:31:00.6.	42.879	21.062	Unknown	1.6	Kosovë
19.11.2013	16:51:09.0.	42.812	21.02	Unknown	1.9	Kosove
19.11.2013	20:35:06.5.	42.795	21.081	Unknown	1.5	Kosovë
19.11.2013	21:14:16.0.	42.853	21.054	Unknown	2	Kosovë

Këto raportime të të dhënave nga ISK bëhen mbi baza mujore, e që, në fakt, nevoja për informacione është e nevojshme të jenë në kohë reale. Në të njëjtën kohë, duhet të përmendim se territori i Kosovës është prekur shumë herë me pasoja të rënda edhe nga tërmetet e ndodhura në vendet fqinje, p. sh., tërmetet e ndodhura në territorin e Shqipërisë, në Mal të Zi, Maqedoni etj.

4.2.2 Rasti i zjarreve pyjore

Pyjet mbulojnë një pjesë të konsiderueshme të Kosovës. Sipas disa burimeve deri në 43 për qind të territorit mbulohet me pyje e shkurre dhe këto zona janë veçanërisht të prira ndaj zjarreve pyjore, posaçërisht në fund të pranverës dhe gjatë muajve të thatë të verës (Faulkner, 2011).

Zjarret në pyje dhe kullota shpesh janë ngjarje të një shkalle më të vogël, në krahasim me ato që kanë ndikim dramatik mbi jetën dhe pronën. Përfshirja bëjnë zjarret që mbulojnë sipërfaqe të mëdha shkurresh ose pyje ku përfshihen edhe zona të banuara, pasi këto të fundit mund të shkaktojnë shkatërrim në një zonë të gjerë.

Si rezultat, zonat specifike të rrezikut janë lehtësisht të identifikueshme, ndërkohë që parandalimi, lehtësimi, mbrojtja dhe parapërgatitja janë të mundura dhe mund të kenë efekte të ndjeshme për uljen e rrezikut. Zjarret që përfshijnë sipërfaqe të tëra, janë shumë vështirë për t'u kontrolluar, qofshin zjarre të shkaktuara nga njerëzit ose zjarre natyrore. Drejtimi dhe forca e erës janë faktorë të veçantë në paparashikueshmërinë e tyre. Në vendin tonë rreziku më i madh i zjarreve të pyjeve është evident në fund të pranverës dhe gjatë periudhës së verave të thata.

Që prej vitit 2000 numri i zjarreve pyjore është në rritje. Zjarrfikësit dhe ekipet e tjera relevante operacionale kanë kryer 2000-3000 intervenime në secilin vit pasues. Në vijim paraqesim të dhënat e vitit 2006, nga Shërbimi i Zjarrfikësve të Kosovës (SHZK) (Faulkner, 2011).

Tabela 4.4 Të dhënat e vitit nga Shërbimi i Zjarrfikësve të Kosovës (SHZK) (2006)

Komuna	Data e aksidentit	Nr. i zjarreve	Shërbime	Aksidente rrugore	Vërshime	Të vdekur	Të lënduar
Prishtina	2006	911	276	78	62	1	12
Gjilani	2006	712	213	35	41	1	0
Prizreni	2006	575	274	78	26	7	8
Mitrovica	2006	545	265	18	21	1	4
Peja	2006	653	148	15	22	2	2

4.2.3 Analizimi i të dhënave i Institutit Hidrometeorologjik - Rasti i vërshimeve

Në vitin 1973 qyteti i Ferizajt ishte goditur nga një vërshim i ashpër, i cili dëmtoi pjesën jugore të qytetit. Vërshime pat edhe në komunën e Pejës gjatë vitit 2013. Janë vetëm disa nga shembujt e që vendi ynë është ballafaquar me situata emergjente të vërshimeve (Armenv&Hachim&George, 2011).

Nëse i marrim të dhënat e stacioneve meteorologjike në Pejë, Prishtinë dhe Ferizaj të vitit 2012, vlerat mesatare mujore të elementeve dhe dukurive meteorologjike janë të paraqitura si më poshtë (Faulkner, 2011).

Tabela 4.5 Të dhënat e Stacionit Meteorologjik në **Pejë** (Institut Hidrometeorologjik, 2014).

2012	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	m. vjet
Tmax.	2.3	0.2	14.1	17.2	21.1	28.9	32.1	32.6	22.6	18.1	16	6.4	17.7
Tmin.	-4.9	-7.5	1.3	5.1	9.8	14.2	17.5	16.3	11.8	7.9	5.3	0.6	6.4
Tmes.	-1.5	4.1	7.6	11.8	16.3	23.1	25.3	26.0	16.8	11.6	8.5	4.1	12.8
Lagështia %	89	86	83	79	73	72	59	53	72.1	86	83	87.3	76.8
Shty. atm	943.6	948.3	940.2	946.3	951.2	956.8	961.7	963.1	958.3	955.4	952.8	953.2	952.5
Era/ms.	1.2	1.8	1.6	1.9	1.3	0.8	1.2	0.9	0.8	1.4	1.3	1.1	1.3
Reshje	121.6	41	13.2	58.4	96.3	5.8	40.6	0.3	24.7	34.5	65.4	70.3	572.1

Tabela 4.6 Të dhënat e Stacionit Meteorologjik në **Prishtinë** (Instituti Hidrometeorologjik, 2014).

2012	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	m. vjet
Tmax.	2.3	0.0	13.6	16.8	20.7	28.5	31.7	31.8	27.6	21.6	14.1	3.0	17.6
Tmin.	-5.2	-7.9	1.1	4.9	9.4	13.7	16.6	15.0	12.1	7.5	4.7	-3.0	5.7
Tmes.	-1.7	3.7	7.1	11.1	15.5	22.2	24.9	24.2	20.0	13.8	8.6	-0.6	12.4
Lagështia %	83	81	62	68	66	57	54	47	58	68	79	86	67.4
Shty. atm	945.7	945.5	950.4	939.4	944.5	947.2	946.8	948.5	947.5	945.8	947.4	943.2	945.9
Era/ms.	1.8	2.2	2.2	2.6	1.9	1.2	1.6	1.3	1.4	1.5	1.5	1.7	1.7
Reshje	105.7	36.1	12.8	51.5	102	6.2	53.3	3.9	13.7	60.4	29.6	65.9	541.1

Tabela 4.7 Të dhënat e Stacionit Meteorologjik në **Ferizaj** (Instituti Hidrometeorologjik, 2014).

2012	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	m. vjet
Tmax.	1.9	0.2	12.6	16.4	20.0	27.9	31.9	31.0	26.2	20.9	12.7	2.4	17.0
Tmin.	-5.7	-7.4	0.6	4.2	8.7	11.9	15.2	14.6	11.7	7.0	-4.4	-3.1	4.4
Tmes.	-2.2	-3.1	6.8	10.5	14.4	21.1	24.2	23.8	18.7	13.2	7.8	-0.6	11.2
Lagështia %	87	83	66	71	75	65	60	53	68	78	87	88	73.4
Shty. atm.	946.0	945.5	951.0	939.5	944.3	947.5	947.1	48.9	942.7	945.9	947.6	943.6	870.8
Era/ms.	1.2	2.4	1.4	1.9	1.4	1.2	1.8	0.9	1.4	1.1	1.8	1.8	1.52
Reshje	67.4	41.7	17.1	64.5	132.5	3.8	14.1	48.2	26.6	27.9	30.7	50.8	525.3

Nga këto të dhëna përfshihen nivelet e reshjeve, sasia e borës në malet e larta, rrjedha e lumenjve, niveli i ujërave nëntokësor dhe niveli i mundësive akumuluese të liqeneve. Të gjitha këto variable duhet të monitorohen vazhdimisht dhe të raportohen nga qendra tek QOAME, që ka përgjegjësi për monitorimin kombëtar të përmbytjeve.

Për këtë regjistruesi automatik i të dhënave në stacionin matës, duhet të jetë i lidhur me rrjetin telefonik (mobil) GPRS.

E gjithë literatura mbi praktikën më të mirë sa i përket menaxhimit të përmbytjeve konfirmon dy parime thelbësore:

- Të dhënat hidrometeorologjike duhet që në mënyrë konstante të monitorohen, në mënyrë që të zbulohet fillimi i një situate të thatësisë;
- Duhet të ekzistojnë nivele të qarta të alarmimit (varësisht nga statusi i të dhënave hidrometeorologjike) në bazë të të cilave mund të ndërmerren veprime të ndryshme për zbutjen e kësaj situate.

Nga pikëpamja menaxhuese, ky lloj i përmbytjeve shpesh mund të jetë i parashikuar që më parë (10 deri në 30 ditë), duke e lejuar zbatimin e masave lehtësuese. Por, përmbytjet mund të vazhdojnë për shumë ditë apo javë, e që mund të bllokojnë infrastrukturën komunikuese (energjinë, telekomunikacionin dhe transportin) dhe si pasojë e kësaj të ketë nevojë për kërkesa të mëdha sociale në lidhje me infrastrukturën (evakuimi, shpëtimi dhe shërbimet mjekësore dhe strehimi i të zhvendosurve) (Faulkner, 2011).

4.2.4 Rasti i orteikut të borës në Restelicë të Sharrit në vitin 2011 bëri që shumë njerëz edhe të vdesin (Forumi për siguri, 2013). Në Kosovë shumica e reshjeve dimërore bëhet nga bora, e cila akumulohet kryesisht në malet e larta. Ky akumulim përfaqëson një formë të magazinimit/akumulimit të ujit dhe për këtë arsye monitorimi i kësaj variable në të vërtetë është shumë i rëndësishëm dhe është një mekanizëm parashikimi afatgjatë. Shtresat e borës mund të maten çdo javë me përdorimin e një matësi shumë të thjeshtë, të monitoruar nga vëzhguesi.

Për shembull, një nivel i ulët i shtresës së borës gjatë muajve Dhjetor deri në Mars do të rezultojë në:

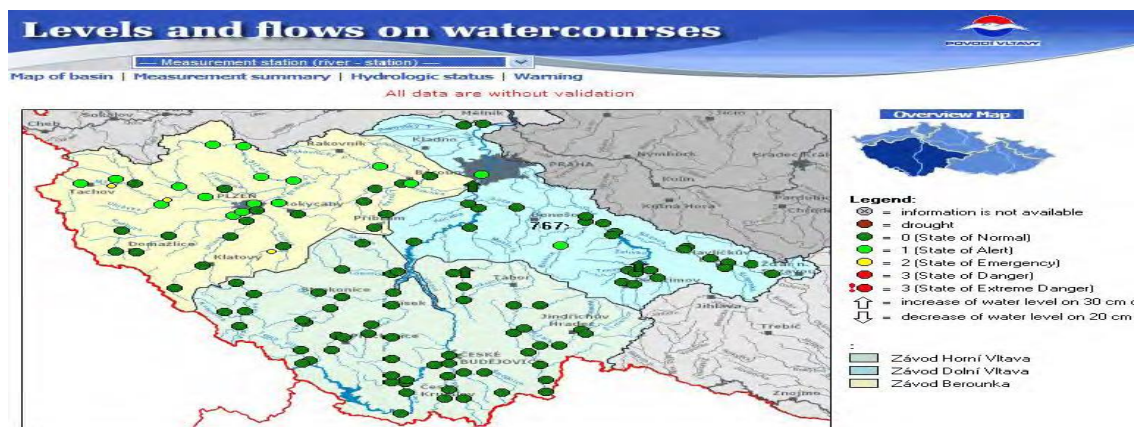
- a) Prurje më të ulëta se normalja në Dhjetor deri Mars;
- b) Sasi të mëdha të borës së shkrirë gjatë Prillit dhe Majit.

Prandaj, thellësitë e borës më të larta se mesatarja janë një tregues i mirë i mundësisë së zhvillimit të përmytjeve gjatë muajve pranverorë.

Instituti Hidrometeorologjik i Kosovës (IHK) duhet të merret më seriozisht me instalimin e matësve të thellësisë së borës në malet e grykëderdhjeve të lumenjve kryesorë. Ata do të duhej të lexoheshin vetëm në baza javore nga një observues i trajnuar dhe, zakonisht, vetëm për muajt nga Shkurti deri në Prill.

Një nga shembujt më të mirë të një sistemi të tillë është ai i Institutit Hidrometeorologjik Çek. Ky sistem online paraqet të dhëna reale nga rrjeti hidrometeorologjik në faqen e internetit të Institutit (shih figurën 4.1). Është e mundur për të kontrolluar statusin e çdo stacioni të vetëm, duke përdorur faqen e internetit. Ky lloj i burimit të informacionit online, në kohë reale, mund të përdoret nga shumë menaxherë të ujit për të marrë vendime afatshkurtra operative.

Figura 4.1 Sistemi Informativ Hidrometeorologjik Çek (Faulkner, 2011).



Sistemi Informativ Çek i ujit, jo vetëm që siguron informacion në kohë reale, por ofron një arkiv të trendeve të prurjeve, të cilët janë të rëndësishëm në atë se tregon se si përmytjet ose thatësira mund të zhvillohen. Analiza e trendeve është shumë e rëndësishme në parashikimin e periudhave të ardhshme të përmytjeve. E rëndësishme është, po ashtu, të tregohet statusi i prurjeve në raport me nivele të ndryshme të emergjencës së përmytjes (ose thatësirës), kështu që është e lehtë për të identifikuar nëse prurja e lumit është në një gjendje të normalitetit, alarmimit, emergjencës etj. Pasi sistemet GPRS të llojit telemetrik të futen në rrjetin hidrometeorologjik të Kosovës, procesi i këtyre të dhënave, për të siguruar informata në kohë reale e të sakta, duhet të bëhet prioritet i IHM-së.

Mund të përfundojmë se paraqitja e rregullt e të dhënave hidrometeorologjike është një komponentë e rëndësishme e gatishmërisë ndaj përmytjeve.

Aktualisht IHMK operon me një numër të pajisjeve “automatike” matëse të reshjeve (përafërsisht 12). Këto pajisje regjistrojnë të dhënat vetëm në një regjistruar të dhënash që, pastaj, duhet të shkarkohen në terren. Ky proces nuk është i volitshëm për paralajmërim të emergjencës së përmytjes dhe ne kemi propozuar avancimin e pajisjeve aktuale në sistemin e “gjendjes telemetrike” me anë të përdorimit të sinjalit GPRS. Transmetimet automatike të të dhënave mbi reshjet në një qendër kontrolli, patjetër, do të ofronin një mundësi të paralajmërimit të hershëm mbi konditat e mundshme për përmytje (Faulkner, 2011).

4.3 Analizimi i të dhënave në rastin e emergjencave dhe fatkeqësive natyrore

Katastrofat e mëdha kanë shënuar dhe formësuar historinë e njerëzimit. Ato nuk janë ngjarje vetëm të shekullit të fundit. Që nga kohët parahistorike e deri më sot shumë ngjarje kanë lënë gjurmë, si në ndikimin fizik mbi planetin ashtu edhe në kujtesën e njerëzve.

Një nga dallimet e rëndësishme në përcaktimin e madhësisë së një katastrofe është nëse diskutojmë për numrin e viktimave. Shumica e studimeve mbi katastrofat janë të fokusuara, kryesisht, në dy shekujt e fundit, e sidomos në dekadat e fundit.

Duke analizuar me kujdes tri dekadat e fundit në Evropë (1980-2014), të paraqitura në tabelën 4.8, vihet re fakti se trendi i katastrofave natyrore është në rritje.

Tabela 4.8 Katastrofat natyrore në tri dekadat e fundit (1980-2014) (<http://www.emdat.be>)

Vitet	Fat. natyrore	Nr. rasteve	Humbje të njerëzve	Të prekurit
1980-2014	Tërmetet	105	33159	3305601
1980-2014	Vala e të ftohtit	153	5778	1386133
1980-2014	Vala e të nxehtit	60	134397	2120
1980-2014	Stuhitë	414	2347	8759900
1980-2014	Përmytjet	497	3356	13193747
1980-2014	Thatësira	40	2	10488769
1980-2014	Zjarret	98	502	1295600
1980-2014	Epidemitë	45	440	189556

Nëse i analizojmë dëmet e shkaktuara nga katastrofat për të njëjtën periudhë, katastrofat natyrore shkaktojnë shumë më tepër dëme sesa ato me natyrë njerëzore.

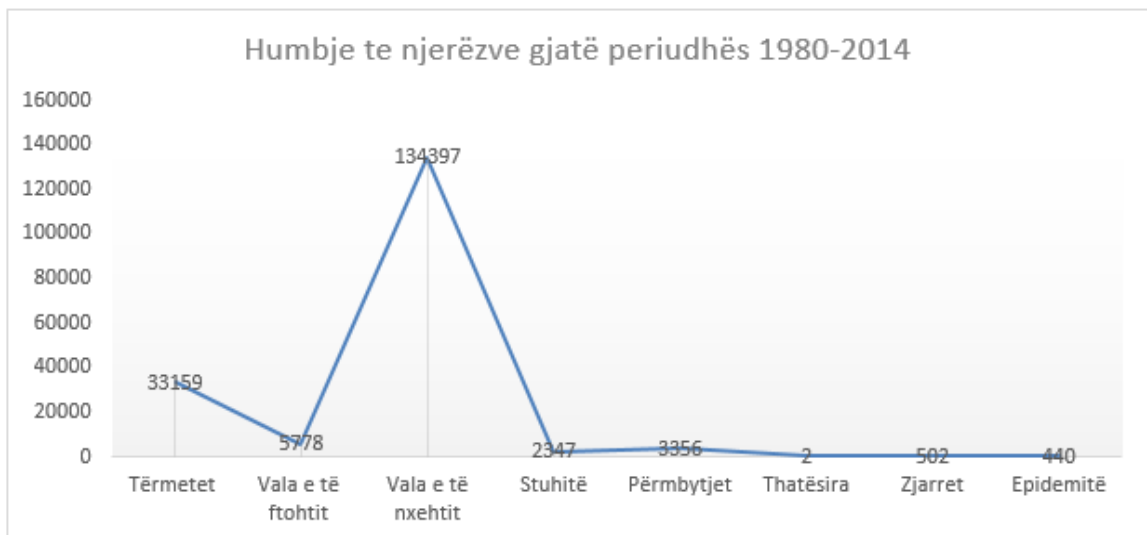
Tabela 4.9 paraqet humbjet nga katastrofat me natyrë njerëzore për periudhën 1980-2014.

Tabela 4.9 Katastrofat natyrore – fatkeqësitë teknologjike (www.emdat.be)

Vitet	Fat. teknologjike	Nr. rasteve	Humbje të njerëzve	Të prekurit
1980-2014	Fat. teknologjike	812	26445	308533

Vihen re se dëmet më të mëdha shkaktohen nga ngjarje natyrore me natyrë: meteorologjike, hidrologjike dhe gjeofizike (tërmetet, vala e të ftohtit, vala e të nxehtit, stuhitë, përmbytjet, thatësira, zjarret, epidemitë etj.) (Lito, 2013), siç shihet në grafikun 4.1

Grafiku 4.1 Përmbledhje e katastrofave natyrore 1980-2014



Nisur nga kjo, për qëllim analize, është shqyrtuar gjendja aktuale lidhur me kapacitetet e komunave për intervenim në raste emergjente, qoftë në kuadrin ligjor, qoftë kapacitetet teknike dhe njerëzore, që disponojnë institucionet relevante dhe përgatitja e vendit tonë për menaxhimin e emergjencave. Kur kësaj i shtohet edhe mungesa e analizave të kësaj problematike, qoftë nga institucionet e qoftë nga mekanizmat tjerë joqeveritarë, e konsiderojmë të nevojshme vështrimin e kapaciteteve institucionale dhe të nevojave të institucioneve relevante për reagime në situata emergjente (Forumi për siguri, 2013). Qëllimi i teknologjisë së informacionit është të lehtësojë shkëmbimin e informacionit dhe të komunikimit në mesin e organizatave të shumta dhe agjencive (Kristina, 2004).

Sistemet e informacionit janë një fazë e rëndësishme në planifikimin dhe përgatitjen për fatkeqësitë (Sriganesh&Iyer). Në një situatë emergjente, informacioni përkatës në lidhje me elementët e përfshirë është i nevojshëm. Për më tepër, procesi i menaxhimit të emergjencave është dinamik dhe përfshin disa hapa të caktuar (Tanasescu&Gugliotta, 2006:956).

4.4 Burimi i të dhënave kryesore të QOAME-së

Qendra Operative e Agjencisë së Menaxhimit të Emergjencave disponon me sasi të mëdha të dhënash. Mund të jetë njëri nga sektorët, i cili, për vetë funksionet që kryen dhe platformat që zotëron, ruan më shumë informacion se çdo tip tjetër industrie (Juliana, 2014).

Pra, Qendra operative e tillë ka shumë burime informacioni, por janë disa kategori të cilat ne do t'i trajtojmë në zgjidhjet vijuese:

1. Të dhënat e Agjencisë së Regjistrimit Civil (ARC) përmbajnë informacione nga regjistrimi i popullsisë, duke përfshirë shënimet bazike, si: emri, mbiemri, mbiemri i vajzërisë, vendi i lindjes, komuna e lindjes, shteti i lindjes, data e lindjes, gjinia, gjatësia, grupi i gjakut, statusi martesor, komuna, vendi, rruga, nr. i ndërtesës, nr. i apartamentit, fotografia e personit, shenjat e gishtërinjve, nënshkrimi elektronik.

2. Të dhënat që lidhen me Institutin e Mjekësisë së Kosovës (IMK), janë kryesisht të dhëna që identifikojnë shënimet fillestare të një pacienti, duke përfshirë pjesën historike të trajtimeve mjekësore, ku shënimet fillestare do të jenë: numri rendor i pacientëve, institucioni shëndetësor, departamenti-reparti, njësia, emri, mbiemri, mbiemri i vajzërisë, emri i prindit, vendi i lindjes, profesioni, datëlindja, gjinia, grupi i gjakut, Rh-faktori, alergjitë, vendbanimi, adresa, telefoni kontaktues, e-maili, statusi bashkëshortor, statusi i sigurimit shëndetësor,

medikamentet e rekomanduara, data e fillimit të trajtimit, data e përfundimit të trajtimit, emri dhe vendi i institucionit, gjendja gjatë vizitës - hospitalizimit, emri dhe mbiemri i mjekut.

3. Të dhënat që lidhen me Institutin Sizmologjik të Kosovës (ISK), janë kryesisht informacione nga ndodhja e tërmeteve dhe përmbledhen si shënime bazike: numri rendor, emri i stacioneve sizmike, data e lëkundjes sizmike, koha lokale e lëkundjes sizmike, latituda sizmike, longituda sizmike, magnituda e lëkundjes sizmike, thellësia sizmike, data fillestare e lëkundjes sizmike, data përfundimtare e lëkundjes sizmike, pamja satelitore e lëkundjes sizmike.

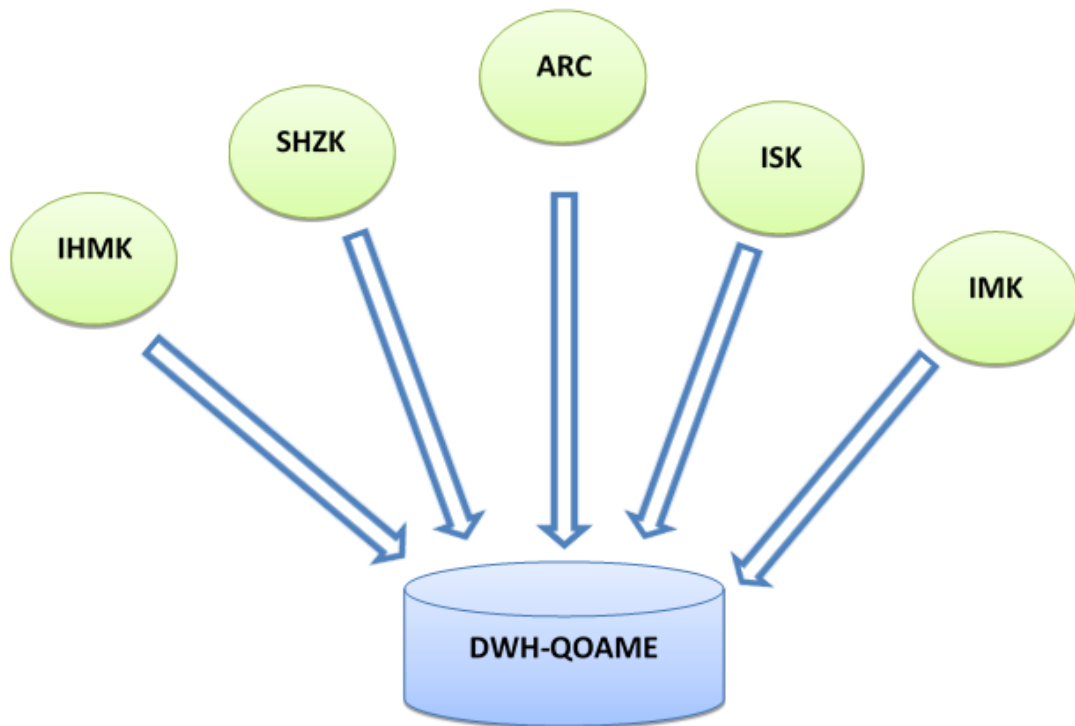
4. Ndërsa të dhënat nga Shërbimi i Zjarrfikësve të Kosovës (SHZK), janë kryesisht informata që përmbledhen nga rastet e zjarreve dhe përfshihen: numri i raportit, data e aksidentit, vendi i aksidentit, komuna, vendndodhja, lloji i zjarreve, sinjali alarmues, numri i telefonit, intervenimet teknike, koha e nisjes, koha e arritjes, koha lokale, koha e kthimit, vërejtjet.

5. Të dhënat që lidhen me Institutin Hidrometeorologjik të Kosovës (IHMK), që përmbajnë informacione nga regjistrimi i pajisjeve matëse, duke përfshirë shënimet bazike, si: numri rendor, emri i stacionit matës, numri i stacioneve matëse, temperatura maksimale, temperatura minimale, temperatura mesatare, lagështia e ajrit, shtypja atmosferike, shpejtësia e erës, reshjet atmosferike.

Këto që i përmendëm më lart (Juliana, 2014), konsiderohen burime të brendshme, pasi QOAME ka edhe burime të jashtme, siç mund të jenë të dhënat nga shërbimet e sigurisë kombëtare.

Në figurën e mëposhtme tregohen burimet që shfrytëzohen për ndërtimin e DWH-së, e që ndihmon në analizimin dhe raportimin në procesin e vendimmarrjes në QOAME.

Figura 4.2 Diagrami i kalimit të informacionit drejt DWH-QOAME.



Duke u nisur nga këto sisteme, QOAME ka interes grumbullimin dhe integrimin e informacionit, e me këtë lind nevoja për një integrim të të dhënave të gjeneruara nga platformat e ARC-së, ISK-së, IMK-së, IHMK-së dhe SHZK-së.

4.5 Integrimi i të dhënave dhe menaxhimi i katastrofave

Integrimi i të dhënave është proces i kombinimit të dhënave që ndodhen në burime të ndryshme dhe sigurimi i një pamjeje të unifikuar i këtyre të dhënave për përdoruesin. Ky proces zhytet në një sërë situatash, qoftë të bizneseve (kur dy kompani të ngjashme duhet të bashkojnë databazat e tyre), qoftë shkencore (kombinimi i rezultateve të kërkimit nga ato të ndryshme sizmike, mjekësore, meteorologjike, të zjarrit dhe emergjencave). Problemi i integritetit të të dhënave bëhet më kritik, ndërsa sasia e të dhënave që duhet të ndahet rritet. Me investigimin e të dhënave nënkuptojmë procesin e futjes në një sistem të dhënash, i cili vjen nga disa burime heterogjene.

Shkallëzimi është një çështje e rëndësishme në gëlltitjen e të dhënave, meqenëse rrjedha e informacionit mund të jetë shumë e lartë gjatë kohës së një ngjarjeje kritike (Hristidis& Chen, 2010:1701). Fusha e integritetit të të dhënave ose të integritetit të informacionit ka bërë progres të madh në themelet teorike dhe në zhvillimin e algoritmeve dhe mjeteve (Peng&Zhang, 2010:316–327). Integrimi i të dhënave është problem i kombinimit të të dhënave, të cilat gjenden në burime të ndryshme. Si rezultat i përparimit në teknologjinë e informacionit ka, gjithashtu, DWH për t'i siguruar përdoruesit një pamje të unifikuar të këtyre të dhënave. Problemi i sistemeve të ndërtimit për integrimin e të dhënave është i rëndësishëm në aplikimet e sotme. Duke marrë parasysh zhvillimet në teknologjinë e informacionit, menaxhimi i emergjencave kërkon të dhëna modelimi nga burime të ndryshme, duke koordinuar aktivitetet midis institucioneve dhe agjencive qeveritare.

Sot, më shumë se kurrë, është nevoja për të krijuar një sistem institucional për t'i koordinuar të gjitha burimet dhe aftësitë që janë në dispozicion e për t'u përgatitur për të reaguar në një

mënyrë të koordinuar dhe të planifikuar, në rast të katastrofave natyrore dhe incidenteve të tjera. Gjatë viteve të fundit, një zhvillim i një teknologjie të shumëllojshme informacioni dhe komunikimi, kanë propozuar alternativa për t'i menaxhuar katastrofat natyrore në nivel kombëtar e rajonal.

Teknologjia e informacionit është bërë një bazë për zhvillimin e shumë sistemeve të informacionit sot. Në shumicën e fushave të jetës është bërë shumë e vështirë të mbahen komunikimet, koordinimet ose komunikimet ndërkombëtare pa përdorimin e gjerë të teknologjisë së informacionit (Krasniqi, 2014). Gjatë çdo faze të menaxhimit të katastrofave, të fatkeqësive, mund të dëshirojmë qasjen në të dhënat nga departamentet: Agjencia e Regjistrimit Civil (ARC), Instituti i Mjekësisë së Kosovës (IMK), Instituti Sizmologjik i Kosovës (ISK), Shërbimi i Zjarrfikësve të Kosovës (SHZK) dhe Instituti Hidrometeorologjik i Kosovës (IHMK). Integrimi i këtyre sistemeve të ndryshme është i rëndësishëm dhe kjo mund të realizohet (*Sam&Lance*, 2009).

Integrimi i të dhënave përfshin kombinimin e të dhënave nga disa burime të ndryshme, të cilat janë ruajtur duke përdorur teknologji të ndryshme dhe siguron një paraqitje të unifikuar të të dhënave. Integrimi i të dhënave bëhet gjithnjë e më i rëndësishëm në rastet e bashkimit të sistemeve të shumë agjencive apo konsolidimin e aplikacioneve brenda një agjencie, që të ofrojë një pamje të unifikuar të të dhënave të aseteve të agjencive për menaxhimin e emergjencave. Përfitimet e një depoje të të dhënave i mundësojnë një biznesi t'i performojë analizat, bazuar në të dhënat në depo. Kjo nuk do të jetë e mundur të bëhet me të dhënat e vlefshme vetëm në sistemin e burimit. Arsyeja është se sistemet e burimeve nuk mund t'i

përmbajnë të dhënat përkatëse. Edhe pse të dhënat kanë emër të njëjtë, ato mund t'u referohen subjekteve të ndryshme (Lenzerini, 2002).

4.6 Konkluzion

Shtetet nuk janë të sigurt nga fatkeqësitë natyrore dhe nga dora e njeriut, pavarësisht përpjekjeve të jashtëzakonshme për t'i zvogëluar fatkeqësitë. Mënyrat tradicionale shpesh nuk janë të mjaftueshme për t'u marrë me fatkeqësitë natyrore. Prandaj, duhet të eksplorohehen mënyra të reja, me qëllim që të zvogëlohen katastrofat. Në bazë të këtyre fakteve, është treguar një përmbledhje gjithëpërfshirëse e shumë çështjeve, problemeve, sfidave dhe kërcënimeve që rimodelojnë menaxhimin e fatkeqësive natyrore.

Me qëllim që të adresohen këto çështje:

- Së pari, teknologjia DWH eliminon teprinë dhe papajtueshmërinë e të dhënave, duke arritur ndarjen e të dhënave pa letër në një organizatë dhe duke rritur besimin në cilësinë e të dhënave.
- Së dyti, një DWH përmirëson efektshmërinë e detyrave operacionale, meqë ka premtimin që t'i mundësojë organizatave të gjenerojnë raporte dhe analiza ad-hoc për vendimmarrje më të mirë.
- Së treti, një DWH i kupton dhe i plotëson më mirë nevojat e klientëve, rrit në maksimum rezultatet e fushatave, ndërton besnikërinë dhe, së fundi, përmirëson rezultatet finale.

KAPITULLI V :METODOLOGJIA E GRUMBULLIMIT TË TË DHËNAVE DHE SINKRONIZIMI NË SISTEMET ETL NË KOHË REALE

5.1 Hyrje

Me përfundimin e ndërtimit të modelit data mart, duhet të zhvillohen edhe të gjitha proceset që do të bëjnë të mundur grumbullimin e informacionet nga burimet origjinale, transformimin dhe ngarkimin e tij në destinacion (Juliana, 2014). Të tri këto procese llogariten të jenë proceset më komplekse gjatë ndërtimit të sistemeve ETL (Johnson&Jones, 2015). Sipas arkitekturës specifike, destinacioni i të dhënave do të jetë një data mart i ndërtuar për një proces të menaxhimit të emergjencave dhe të fatkeqësive natyrore.

Duhet të përcaktohet nëse informacioni sapo të nxirret nga burimi të transformohet me mjetet e duhura dhe më pas të ngarkohet në strukturën e databazës destinacion apo fillimisht do të ngarkohet ashtu si është në databazë në tabela Stage, për t'u transformuar nga procedura të vetë databazës destinacion. Të dyja mundësitë janë metoda të njohura në literaturë, si ETL, ELT. Në rastin tonë do të përdorim metodat ETL. Proceset (ETL) janë në vetvete zgjidhje software-ike, të cilat janë përgjegjëse për tërheqjen e të dhënave nga burime të ndryshme, për pastrimin e tyre, përshtatjen dhe ngarkimin në databaza.

Andaj, në ditët e sotme agjencitë e menaxhimit të emergjencave gjenerojnë sasi të mëdha të dhënash, me një ritëm i cili mund të arrijë me lehtësi shumë megabyte ose gigabyte në ditë. Dhe, ndërsa koha kalon, një DWH e menaxhimit të emergjencave mund të rritet në disa terabyte ose edhe petabyte. Madhësia e DWH-së tregon se çdo kërkim që ekzekutohet

kundrejt të dhënave të saj, zakonisht, hap sasi të mëdha regjistrimesh, e, po ashtu, bën veprime, si: bashkime, radhitje, grupime dhe llogaritje.

Për të optimizuar këto akseset, DWH përdor struktura të dhënash të brendshme të përcaktuara, si, për shembull, indeksset ose particionet, të cilat janë, po ashtu, të mëdha në sasi dhe kanë një nivel shumë të madh kompleksiteti.

Këto fakte tregojnë se është shumë vështirë të përditësohen në mënyrë të efektshme DWH-të në kohë reale, meqë përhapja e të dhënave transaksionale në kohë reale me shumë mundësi do të mbingarkojë serverin. Dhe, duke pasur parasysh shpeshtësinë e përditësimit dhe volumin, do të përfshijë veprimtari jashtëzakonisht të madhe mbi strukturat e DWH-së dhe do të degradojë dramatikisht performancën.

Me pak fjalë, DWH-ja në kohë reale synon uljen e kohës në komunikimet e rasteve të fatkeqësive natyrore dhe emergjencave. Në këso rastesh, duhet marrë vendime dhe për të arritur një kohëzgjatjeje zero, midis shkakut dhe efektit për atë vendim, duhet mbyllur hapësirën midis sistemeve reaktive inteligjente dhe proceseve të sistemeve. Synimi ynë është transformimi i një DWH-je standard, duke përdorur ngarkimin në grup gjatë dritareve të përditësimit (gjatë të cilave nuk lejohet aksesit analitik), një mjedis analitik me kohëzgjatjeje thuajse zero, duke siguruar të dhëna aktuale, me qëllim që të mundësohet (thuajse) përhapja në kohë reale e informacionit të ri në AME.

Kërkesat e AME-së për këtë lloj mjedisi analitik, prezantojnë një set të ri me marrëveshje të nivelit të shërbimit që shkojnë përtej asaj që është tipike në një DWH tradicionale. Problemi më i madh është si të bëhet i mundur integrimi i vazhdueshëm i të dhënave dhe të sigurohemi që minimizon impaktin negativ në veçori të ndryshme të mëdha të sistemit, si disponueshmëria dhe koha e përgjigjes së sistemeve OLTP dhe OLAP.

Është dhënë një diskutim në thellësi i këtyre veçorive nga këndvështrimi analitik (për të bërë të mundur analizën konsistente në kohë). Kombinimi i sistemeve shumë të disponueshme me aktivitet të vendimmarrjes, lejon përhapjen e informacionit në kohë reale për DWH. Dhe, e gjitha bashkë është baza për mjediset analitike me kohëzgjatjeje zero.

DWH në kohë reale siguron akses në një pamje të saktë, të integruar, të konsoliduar të informacioneve të AME-së dhe ndihmon në dhënien e informacioneve thuajse në kohë reale për përdoruesit e vet. Kjo kërkon teknika ETL të efektshme, që bëjnë të mundur integrimin e vazhdueshëm të të dhënave (Santos&Bernardino, 2008).

5.1.1 Nxjerrja e të dhënave

Nxjerrja e të dhënave nga një sërë sistem burimesh të ndryshme, shpesh është aspekti më sfidues i ETL-së, pasi përcakton fazën për mënyrën si do të vazhdojnë proceset në vijim. Në përgjithësi, synimi i fazës së ekstraktimit është t'i konvertojë të dhënat në një format të vetëm, i cili është i përshtatshëm për përpunim transformimi.

Çdo sistem i veçantë mund të përdorë, gjithashtu, një format/organizatë të ndryshme të dhënash. Formatet e burimeve të të dhënave të përbashkëta janë databazat relacionale dhe fajllat e sheshta, por mund të përfshijnë struktura databazash jorelacionale, si, për shembull, Sistemi i Menaxhimit të Informacionit (SMI) ose struktura të tjera të dhënash, si Metoda e Aksesit të Ruajtjes Virtuale (MARV) ose Metoda e Aksesit Sekuencial të Indeksuar (MASI) (ETL, 2015). Në të shumtën e rasteve, të dhënat në sistemin burim janë shumë komplekse, kështu që përcaktimi se cilat të dhëna janë përkatëse është shumë i vështirë. Ndërtimi dhe krijimi i proceseve të ekstraktimit është përpjekje programimi që kërkon shumë kohë. Për t'i mbajtur të dhënat të përditësuara në një DWH, ato duhet të ekstrahohen disa herë në mënyrë

periodike. Janë dy metoda logjike për ekstraktimin: ekstraktimi i plotë dhe ekstraktimi progresiv.

1. Ekstraktimi i plotë - Në këtë lloj ekstraktimi të dhënat nga sistemi burim ekstraktohen plotësisht. Derisa ekstraktimi i plotë i ekstrakton të gjitha të dhënat nga sistemet burim, nuk ka nevojë t'i ndiqni ndryshimet e bëra në sistemin burim, duke krahasuar me ekstraktimin e mëparshëm.

2. Ekstraktimi progresiv - Në këtë lloj ekstraktimi vetëm ndryshimet e bëra në sistemin burim do të ekstraktohen në krahasim me ekstraktimin e mëparshëm. Kapja e ndryshimit të të dhënave (CDC) është një mekanizëm që përdor ekstraktimin progresiv. Shumica e mjeteve ETL nuk e përdorin mekanizimin CDC dhe në vend të kësaj i ekstraktojnë të gjitha tabelat nga sistemet burim në zonën Stage dhe i krahasojnë ato tabela me të dhëna ose tabela të ekstraktuara nga ekstraktimet e mëparshme për t'i dalluar ndryshimet. Ky krahasim rezulton në një ndikim të madh të performancës në proceset ETL të DWH-së. Janë dy metoda fizike të ekstraktimit: Ekstraktimi online dhe ekstraktimi offline.

- Ekstraktimi online - Procesi i ekstraktimit ETL lidhet me sistemin burim për të ekstraktuar tabelat burim ose për t'i ruajtur ato në një format të konfiguruar tashmë në sisteme ndërmjetëse, p. sh., tabela log.
- Ekstraktimi offline - Të dhënat e ekstraktuara jashtë zonës Stage të sistemeve burim. Procesi i lëvizjes fizike të të dhënave nga një sistem në një sistem tjetër, gjithashtu, është pjesë e procesit të ekstraktimit, i cili quhet transport dhe përfshin:

1. Lëvizjen e të dhënave nga sistemi burim në DWH
2. Lëvizjen e të dhënave nga databaza e stazhimit në DWH
3. Lëvizjen e të dhënave nga DWH në data mart

Transportimi mund të bëhet duke përdorur mekanizma fajllash të sheshta, mekanizëm operacionesh të shpërndara ose duke përdorur hapësirat e tabelave të transportueshme (Kakish&Kraft 2012:1-12).

5.1.2 Transformimi

Faza e transformimit aplikon një sërë rregullash ose funksionesh në të dhënat e ekstraktuara nga burimi për t'i nxjerrë të dhënat për ngarkim në vendin final. Disa burime të të dhënash kërkojnë shumë pak ose aspak manipulim të dhënash. Rastet e tjera kërkojnë një ose më shumë prej llojeve të mëposhtme të transformimit:

1. Zgjedhja e vetëm disa kolonave për të ngarkuar;
2. Përkthimi i vlerave të koduara (p.sh., nëse sistemi burim ruan 1 për mashkull dhe 2 për femër, por magazina ruan M për mashkull dhe F për femër);
3. Vlerat e enkodimit formë e lirë (p. sh., hartim "Mashkull" në "1");
4. Nxjerrja e një vlere të sapollogaritur.
5. Renditja;
6. Bashkimi i të dhënave nga disa burime (p.sh., kërkim, bashkim) dhe dyfishim i të dhënave;
7. Grumbullimi;
8. Gjenerimi i vlerave surrogate-kyçe;
9. Zhvendosja ose lëvizja (kthimi i disa kolonave në disa rreshta ose anasjelltas);
10. Ndarja e një kolone në disa kolona;
11. Ndarja e kolonave të përsëritura në një tabelë të veçantë të detajuar;

12. Kërkimi dhe vlerësimi i të dhënave përkatëse nga tabelat ose fajllat referenciale për dimensionet që ndryshojnë ngadalë;

13. Aplikimi i çdo forme të thjeshtë ose komplekse të verifikimit të të dhënave (ETL, 2015).

5.1.3 Ngarkimi i të dhënave

Kjo fazë ngarkon të dhënat në vendin final, zakonisht në DWH. Në varësi të kërkesave të organizatës, ky proces ndryshon gjerësisht. Disa DWH mund t'i mbishkruajnë informacionet ekzistuese me informacione grumbulluese.

Përditësimet e shpeshta me të dhënat e ekstraktuara bëhen çdo orë, ditë, javë ose muaj. Koha dhe shtrirja për t'i zëvendësuar ose pritur përditësimet janë zgjedhje ndërtimi strategjik, e që varen nga koha në dispozicion dhe nevoja e agjencioni i menaxhimit të emergjencave . Faza e ngarkimit bashkëvepron me databazën, kufizimet e përcaktuara në skemën e databazës, si dhe shkakton aktivizimin e procesit të ngarkimit të të dhënave (për shembull, unikaliteti, integriteti referencial, fushat e detyrueshme) me qëllim që do t'i kontribuojnë performancës së përgjithshme të datawarehouse (DWH) në procesin ETL.

Mekanizmat për të ngarkuar një DWH përfshijnë:

1. SQL loader: i përdorur përgjithësisht për të ngarkuar fajlla të sheshta në datawarehouse (DWH).
2. Tabelat e jashtme: ky mekanizëm bën të mundur që të dhënat e jashtme të përdoren si tabelë virtuale, e cila mund të kërkohet dhe bashkohet para se të ngarkohet në sistemin e synuar.

3. Oracle Call Interface (OCI) dhe Application Programming Interface (API) me rrugë të drejtpërdrejtë: janë metodat e përdorura kur procesi i transformimit bëhet jashtë databazës.

4. Eksport/import: ky mekanizëm përdoret nëse ka transformime komplekse dhe të dhënat mund të ngarkohen në sistemin e DWH-së, të synuar ashtu siç janë.

DWH-ja tipike me bazë ETL përdor zonën Stage, integrimin, si dhe shtresa për të strehuar funksionet e saj kyçe. Shtresa e zonës Stage ose databaza Stage ruan të dhënat e papërpunuara, të nxjerra nga secili prej sistemeve të ndryshme të të dhënave burim. Shtresa e integritit integron setet e ndryshme të të dhënave, duke i transformuar të dhënat nga shtresa e stage, e shpesh duke i ruajtur këto të dhëna të transformuara në një databazë depo të dhënash operationale (ODS). Pastaj, të dhënat e integruara lëvizin në një databazë tjetër, shpesh të quajtur databaza e DWH-së, ku të dhënat radhiten në grupe hierarkike të quajtura dimensionale dhe në fakte e fakte agregate.

ETL luan një rol të rëndësishëm në arkitekturën e DWH-së meqenëse këto procese ETL lëvizin të dhënat nga sistemet transaksionale ose burim në vendet e zonës stage të të dhënave dhe nga vendet e zonës stage të DWH-së (Kakish&Kraft 2012:1-12).

- Rasti studimor- Instituti i Mjekësisë së Kosovës (IMK)

5.2 Mjetet e përdorura për ETL–Talend Open Studio for Big Data

Për ETL është përdorur mjeti Talend Open Studio for Big Data, që përbëhet nga komponentët e mëposhtëm:

Datawarehouse (DWH) mund të konsiderohet vegla më me peshë në marrjen e vendimeve strategjike. Megjithatë, në një ambient shkencor, ende nuk janë përdorur si duhet të gjithë komponentët e saj. Synimi i këtij kapitulli është të dizajnojë dhe të implementojë procese

ETL me anë të integrimit të të dhënave të databazave heterogjene që përdorin Talend Open Studio (TOS).

Përdorimi i Talend Open Studio and Big Data(TOS) për dizajnimin dhe implementimin gjysmë-autonom të detyrave ETL në Java, mundëson një mënyrë të shpejtë të adaptimit me burime të reja të dhënash.

Me këtë jemi në gjendje të menaxhojmë integrimin e burimeve të ndryshme të të dhënave dhe t'i vendosim në një DWH qendrore, duke përdorur Talend Open Studio and Big Data.

Është e rëndësishme që të sigurohemi që kapacitetet e ETL-së të jenë të barabarta me ato të çdo produkti tjetër në treg, sepse kjo do të lehtësojë implementimin e projekteve DWH me kosto më të ulët. (Azwa & Abdul).

Figura 5.1 Komponentët e Talend Open Studio for Big Data .



Duke ndjekur edhe tiparin e ndarjes të sistemeve DWH, proceset ETL duhet të bëhen mbi një databazë të ndryshme dhe të pavarur nga burimet dhe DWH. Aplikimi i teknikave ETL, me anë të komponentit Talend Open Studio for Big Data, do të bëhet sipas këtyre hapave:

Hapi 1: Shpjegohen mundësitë e tërheqjes së të dhënave nga burimet

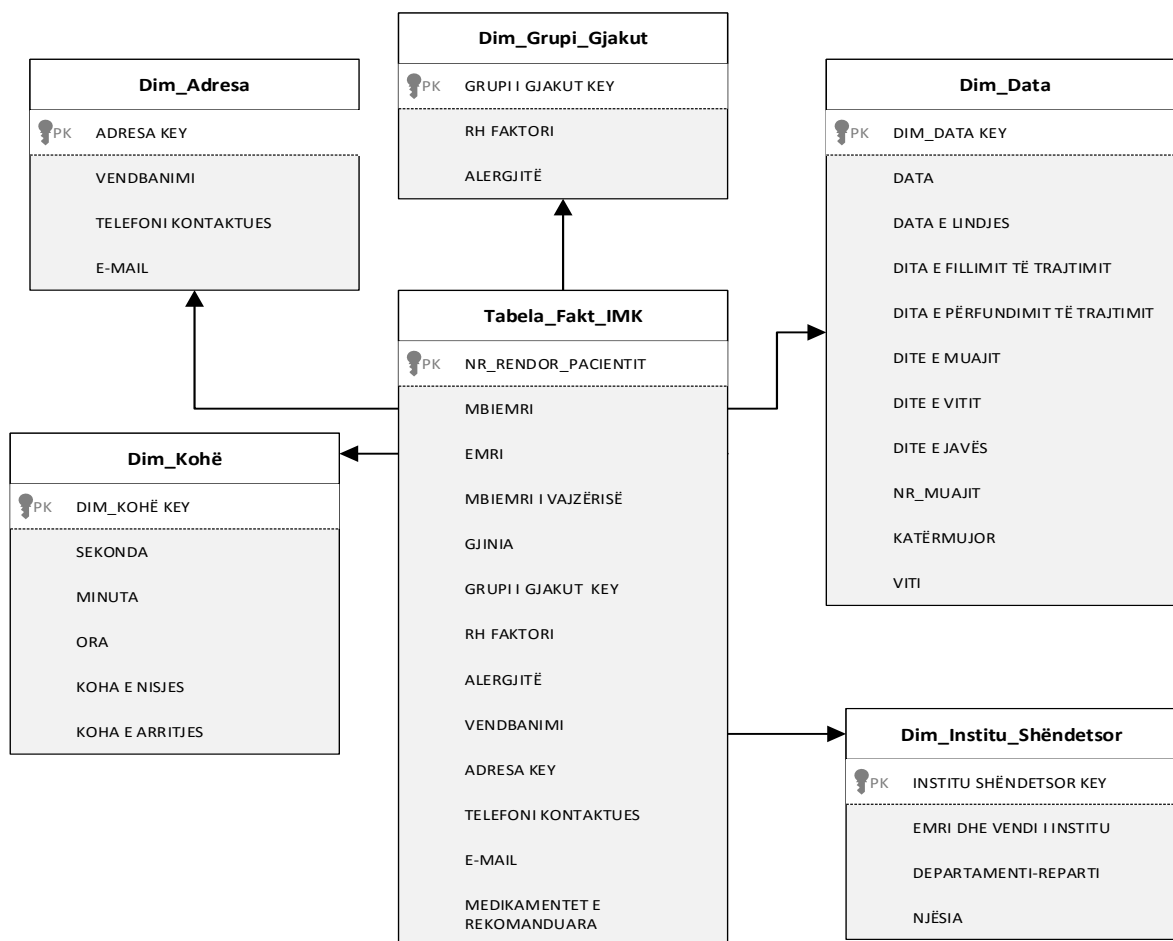
Identifikimi i burimeve është i qartë, meqë fajllët e gjeneruar nga platformat do të jenë burim i informacionit. Nëse fokusi ynë është te fajllët që gjeneron Instituti i Mjekësisë së Kosovës (IMK), janë të dhëna të pastrukturuara. Krijohet një tabelë faktesh Tabela_FAKT_IMK, e cila përmban të dhënat, si: nr. rendor i pacientëve, mbiemri, emri, mbiemri i vajzërisë, gjinia, grupi i gjakut, Rh faktori, alergjitë, vendbanimi, adresa, telefoni kontaktues, e-maili, medikamentet e rekomanduara.

Pesë tabela dimensionit lidhen me Tabela_FAKT_IMK dhe ato janë: Dim_Adresa, Dim_grupi_gjakut, Dim_Data, Dim_Kohe dhe Dim_institu_shendetsor. Dim_Adresa përbëhet nga të dhënat për vendbanimin, si: adresa, vendbanimi, telefoni kontaktues, e-maili. Ndërsa, Dim_grupi_gjakut përbëhet nga të dhënat për grupin e gjakut, llojin e grupit të gjakut dhe alergjitë.

Dim_Data përfshin të gjitha llojet e datave, duke përmblodhur: data e fillimit të trajtimit, data e përfundimit të trajtimit (këto janë për rastet e trajtimit mjekësor), ndërsa data, data e lindjes, ditë e muajit, ditë e vitit, ditë e javës, nr. i muajit dhe viti tregojnë të dhënat e datave të ndryshme sipas nevojës.

Ndërsa, Dim_Kohe përmblodh pjesën kohore të matur sipas: sekonda, minuta, ora, koha e nisjes, koha e arritjes. Kurse, Dim_institu_shendetsor përmban të dhëna të caktuara: emri dhe vendi i institucionit, departamenti-reparti, njësia.

Figura 5.2 Tregon Entity Relationship Diagram (ERD) të sistemeve target.



Nga tabela e mëposhtme paraqiten burimet e të gjitha dimensioneve, si: Dim_Adresa, Dim_grupi_gjakut, Dim_Data, Dim_Kohe dhe Dim_institu_shendetsor, e që janë EXCELFILE, por dhe burimi tjetër tabela FAKT_IMK, që është ndërtuar në ORACLE.

Tabela 5.1 Burimet DBMS dhe atributet

DBMS	TABELA	ATRIBUTET
ORACLE	FAKT_IMK	NR_RENDOR_PACIENTËVE
		MBIEMRI
		EMRI
		MBIEMRI I VAJZËRISË
		GJINIA
		VENDI I LINDJES
		PROFESIONI
EXCELFILE	DIM_KOHË	DIM_KOHË KEYS
		SEKONDA
		MINUTA
		ORA
		KOHA E NISJES
		KOHA E ARRITJES
EXCELFILE	DIM_GRUPI I GJAKUT	GRUPI I GJAKUT KEY RH FAKTORI ALERGJITË
EXCELFILE	DIM_ADRESA	ADRESA KEY
		VENDBANIMI
		TELEFONI KONTAKTUES
		E-MAIL
EXCELFILE	DIM_DATA	DIM DATA KEY
		DATA
		DATA E LINDJES
		DITA E FILLIMIT TË TRAJTIMIT

Kjo do të thotë se çdo tip rekordi në fajllin burim ka një numër të caktuar fushash të ndryshme nga tipat e tjerë. Nevojitet të ndërtohet një proces parësimi mbi rekordet e IMK-së. Ai proces do të realizojë strukturimin e të dhënave.

Ndërkohë, burimet e tjera të informacionit, si, p.sh, Agjencia e Regjistrimit Civil (ARC) prodhon informacion të strukturuar, ku çdo rekord i fajllit të prodhuar ka të njëjtën përmbajtje. Për data mart-in tonë na intereson informacioni i IMK-së. Procesi i tërheqjes së të dhënave nga burimi përfshin:

- Procesin fillestar të nxjerrjes së informacionit,
- Procesin e rifreskimit të vazhdueshëm.

Procesi fillestar ndodh kur nuk ka asnjë informacion në destinacion, ndërsa freskimi është ai proces që sigurohet se të dhënat janë të fundit dhe nxjerr vetëm atë pjesë të informacionit që është i shtuar apo i ndryshuar nga versioni paraardhës. Për përcaktimin e të dhënave të ndryshuara ka disa metoda, por ajo që ne do e përdorim për dimensionet tona është si më poshtë: Pas çdo nxjerrje informacioni nga një burim ruhet në zonën Stage një kopje e të dhënave të nxjerra (në hash-file). Gjatë ekzekutimit pasardhës procesi merr gjithë tabelën burim dhe e krahason me versionin paraardhës të ruajtur në zonën Stage.

Hapi 2: Zhvillimi i hapave të strukturimit të informacionit

Zhvillimi i strukturimit të informacionit përfshihet në këta hapa:

- a) Në këtë rast, do e ndajmë informacionin në modele të veçanta, në bazë të rregullave të secilit tip informacioni. Ndarja do të bëhet në bazë të rregullave që i ndërtojmë për diferencimin e çdo tipi të burimit të informacionit.

b) Hapi i dytë është strukturimi dhe ruajtja e këtyre të dhënave në mënyrë të përshtatshme, për të mundësuar manipulimin e tyre të mëtejshëm. Në rastin tonë, të dhënat do të ruhen në tabela Stage, sipas arkitekturës së zgjedhur. Në logjikën e transformimit do të ndërtohen rregullat e parësimit të informacionit të pastrukturuar, për ta transformuar në një informacion të strukturuar. (Juliana, 2014).

c) Hapi i tretë është përdorimi i operacioneve tJoin dhe tMap në Talend Open Studio. Është shumë i leverdishëm që t'i fusni skemat e të dhënave në *Repository*. Repository ose depoja ka një listë të strukturave (skemave) input dhe output që mund të referohen dhe mirëmbahen globalisht në sistem.

Për këtë janë përdorur dy komplete të dhënash si inpute, Dim_Kohe dhe Dim_Data. Këto të dhëna përmbajnë të dhëna nga instituti seizmik, të depozituara në skema të ndryshme. Sfidë fillestare është të kombinohen të gjitha fushat (skemat) për një institut të caktuar në një shënim (rresht) në strukturën e të dhënave output (një fajllë ose tabelë databazë). Ky operacion performohet nga bashkimi i fushave të një strukture të dhënash input me ato të një tjetër strukture të dhënash.

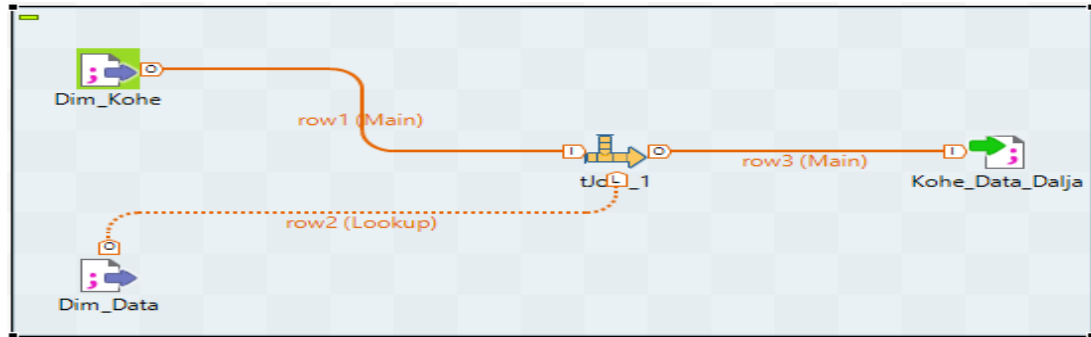
Përdorimi i Talend Open System për të përgatitur kompletin tuaj të të dhënave do të pranojë një varietetet të madh operacionesh manipulimi të të dhënave që nuk gjenden në veglat e analizimit të të dhënave (data mining).

Komponenti tJoin i Talend Open System

Për shumë punë me data mining duhen kombinuar burime të shumëfishta të dhënash për një entitet të caktuar (për shembull, klient), për të prodhuar një rresht të vetëm në një strukturë të dhënash që përmban të gjitha fushat e burimit të inputit të të dhënave. Bashkimi i dy fushave konsiston në kombinimin e fushave (kolonave) të një burimi të dhënash me fushat e

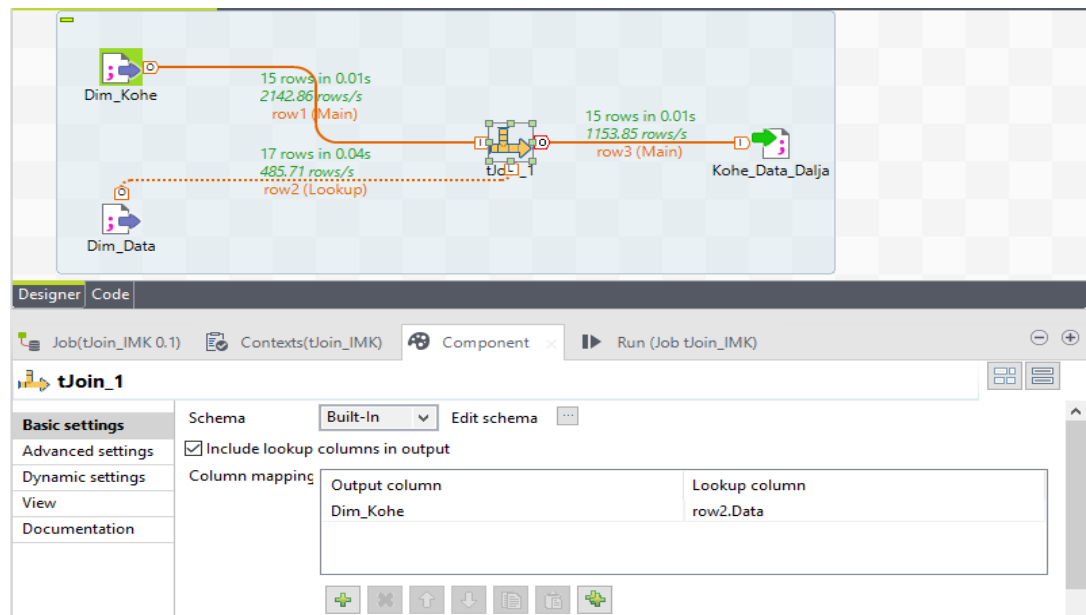
një tjetër burimi të dhënash sa herë që një fushë kryesore përputhet mes tyre. Bashkimi arrihet përmes komponentit tJoin në Talend Open Studio.

Figura 5.3 Tregon një punë të thjeshtë bashkimi të dizajnuar për të kombinuar dy burime të dhënash.



Çdo burim të dhënash duhet të konfigurohet për të pasur akses në kompletin e caktuar të të dhënave sipas informacionit të kërkuar në panelin Basic Settings. Për të parë kushtet e bashkimit, klikoni dy herë te komponenti tJoin për të hapur kutinë e dialogut të treguar në figurën 5.4

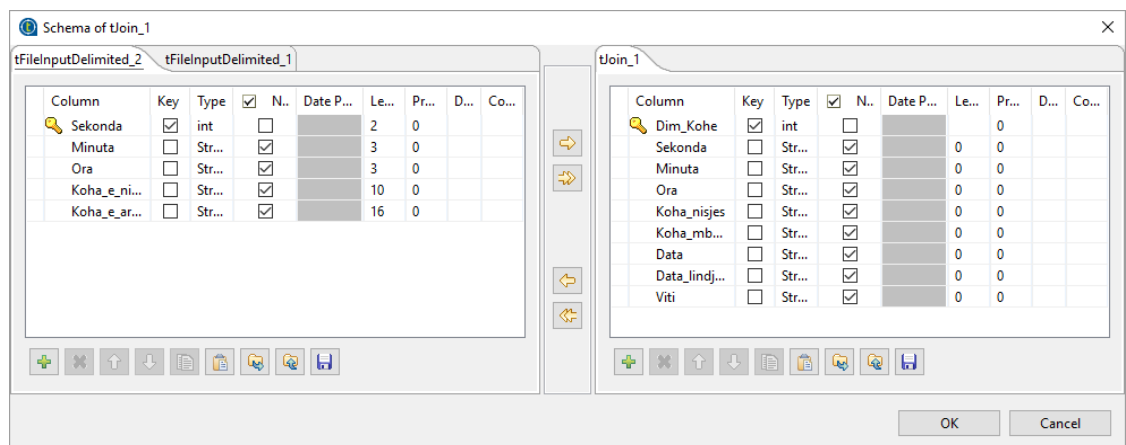
Figura 5.4 Komponenti tJoin - kutia e dialogut



Kutia e dialogut e tJoin tregon opsionet për të specifikuar një bashkim të brendshëm dhe për të edituar skemën e këtij bashkimi.

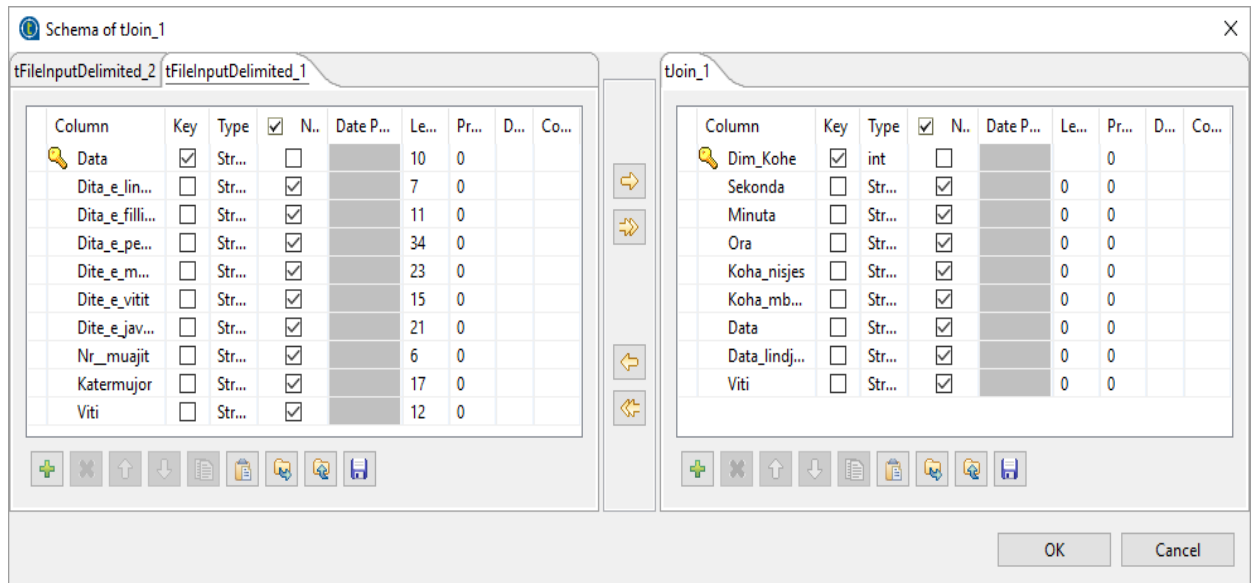
Tipi inner join specifikohet duke klikuar opsionin inner join në ekranin e treguar në figurën 5.4. Një bashkim i majtë specifikohet duke lënë hapur kutinë e bashkimit të brendshëm dhe duke lidhur shigjetën nga “left” data source (Dim_Kohe) në komponentin tJoin, së pari, pastaj duke lidhur burimin tjetër të të dhënave (që do të trajtohet si burim të dhënash “i djathtë”). Burimi i djathtë i të dhënave do të trajtohet si një look-up data source ose burim të dhënash për t’u kombinuar. Një bashkim i djathtë nuk është specifikuar qartazi. Nëse deshirojmë të performojmë një bashkim për zonat B dhe C në figurën 5.5 (që është, në fakt, një bashkim logjik i djathtë), vetëm lidhni, së pari, burimin Dim_Data. Fusha e përbashkët në secilën prej burimeve të të dhënave (çelësi) është specifikuar në definicionin e skemës në depo (Repository). Për të ndryshuar skemën automatike të prezantimit grafik ose default mapping të kolonave input, klikoni në kutinë Edit schema në kutinë e dialogut tJoin, që të shfaqni skemën e prezantimit grafik të treguar në figurën 5.5 (selektoni View Schema në fillim, pastaj klikoni OK).

Figura 5.5 Skema e fushës tJoin për burimin e të dhënave Dim_Data.



Andaj, figura 5.6 tregon ekranin e skemës për burimin e të dhënave Dim_Kohe

Figura 5.6 Skema tJoin për burimin e të dhënave Dim_Kohe.



Qëllimi i operacionit të bashkimit është t'i shtohen fushat nga burimi i të dhënave Dim_Kohe burimit të të dhënave Dim_Data, duke rezultuar në një shënim për një klient të dhënë (Kohe_Data_Dalje) me të gjitha fushat e përfshira.

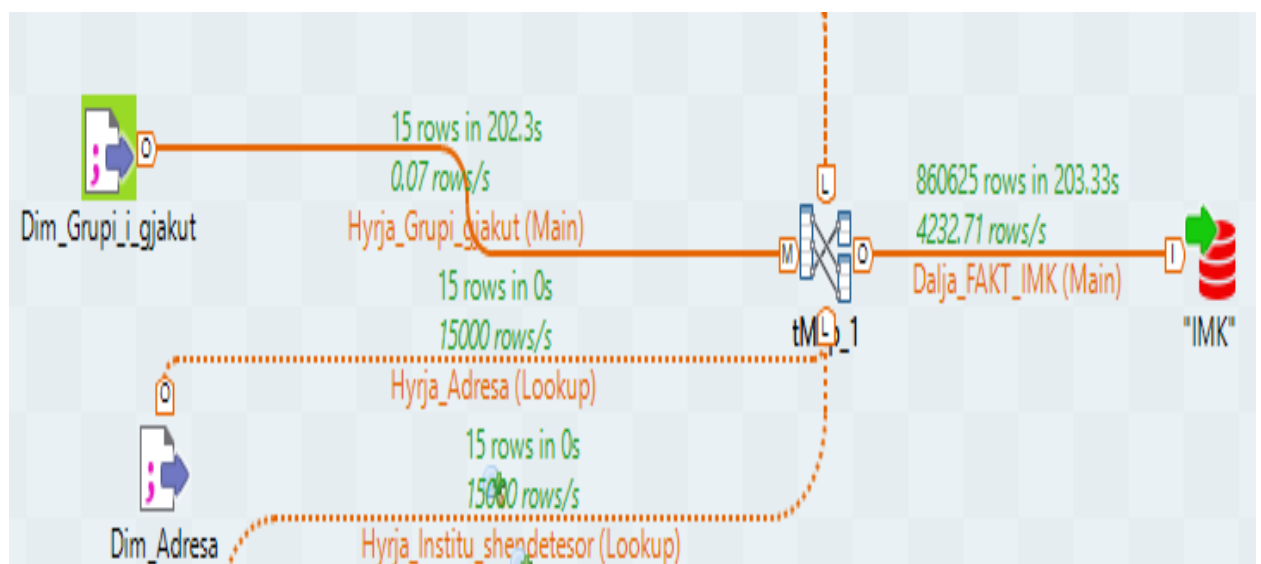
Paneli i majtë tregon fushat për dy burimet input të të dhënave (të selektueshme, duke klikuar në tabin e duhur). Paneli i djathtë shfaq fushat e selektuara nga të dy fajllët. Automatikisht, të gjitha fushat e secilit fajllë janë të selektuara. Andaj, me këtë mund të largohen nga secila prej fushave të strukturës së të dhënave output, duke e theksuar dhe duke klikuar shigjetën e majtë në mes të pamjes. Komponentët output të treguar në figurën 5.5 dhe 5.6 (Kohe_Data_Dalje) përbëhen nga një fajllë, që shërben për transferimin e të dhënave nga një aplikacion te tjetri (comma-delimited text file).

Duke klikuar dy herë mbi komponentin, do të mundësohet konfigurimi i komponentit për të output rrjedhën e bashkuar të të dhënave në dosjen që përzgjidhet.

Komponenti tMap i Talend Open System

Nga figura tregohet komponenti tMap për të prezantuar grafikisht se cilat fusha input lidhen me secilën fushë në strukturën e të dhënave output. Fushat mund të ripozicionohen në shënimet e të dhënave me këtë komponent.

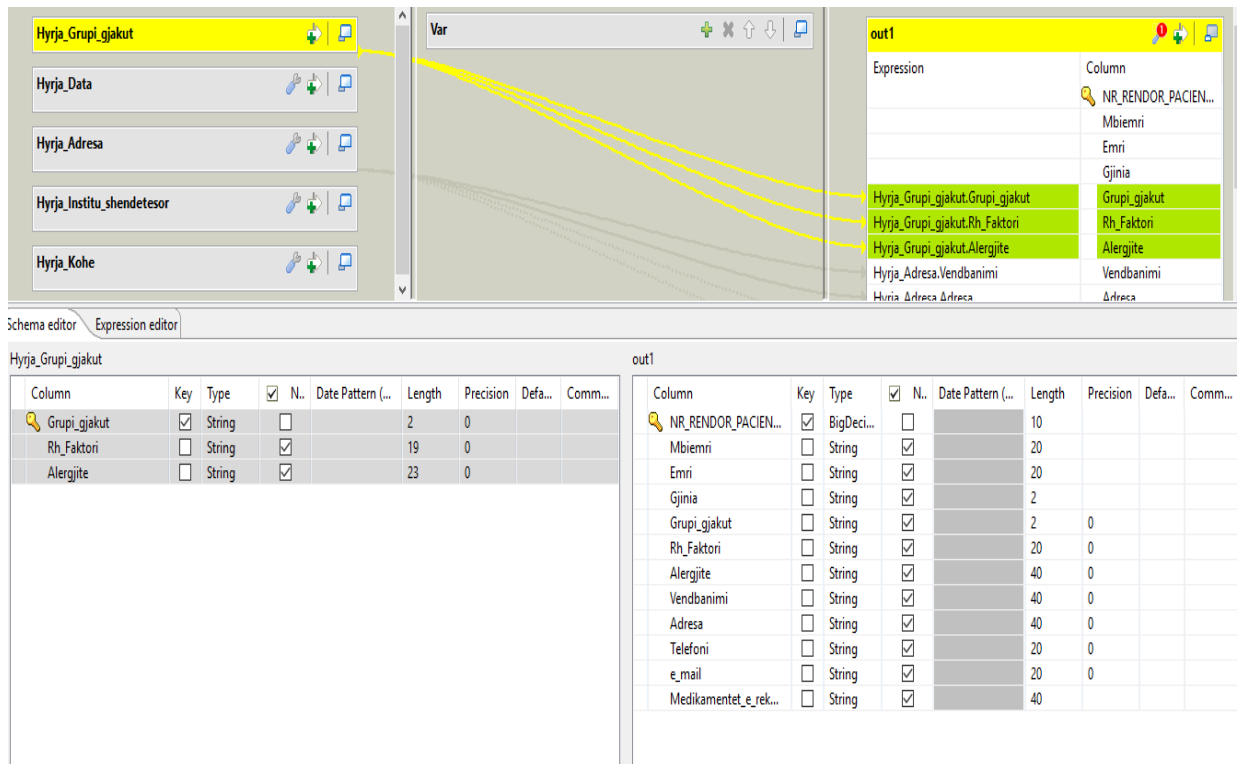
Figura 5.7 Komponenti tMap i Talend open studio



Kështu, figura 5.7 tregon komponentin tMap i Talend për prezantimin grafik të fushave në një strukturë të dhënash output.

Klikimi i dyfishtë në komponentin tMap_1 shfaq ekranin e konfigurimit tMap si në figurën 5.8, si më poshtë.

Figura 5.8 Konfigurimi tMap



Panelet e sipërme tregojnë fushat e input dhe output të skemave të lidhura me vija, të cilat tregojnë prezantimin grafik të dëshiruar. Andaj, fusha si: Hyrje_Grupi_gjakut (Grupi_gjakut, Rh_faktori dhe Alergjite) në panelin e majtë është prezantuar grafiksht me variablën me të njëjtin emër, si edhe fusha out1 në strukturën e të dhënave output. Tabelat tMap që ndodhen mes paneleve të epërme dhe të poshtme përfshijnë një tab që shfaqë Expression Editor. Kombinimi i komponentëve tJoin dhe tMap mundëson të performohen shumë operacione komplekse të manipulimit të të dhënave që nuk mbështeten nga vegla për data mining. Kombinimi i veglave të data mining dhe Talend Open Studio mundëson procesimin në të

gjitha hapat e projektit data mining, duke ndjekur një nga modelet standarde të procesimit.(Robert, 2009)

Hapi 3: Specifikojmë transformimet që do të ndodhin

Pasi të nxirren të dhënat nga burimet, ato i nënshtrohen proceseve të transformimeve. Veprimet transformuese në rastin tonë bëhen mbi të dhënat e nxjerra nga burimet apo mbi objektet e përkohshme nga zona Stage. Po kështu, edhe vetë transformimet mund të krijojnë rezultate të përkohshme që ruhen në zonën Stage.

Proceset më kryesore transformuese që do të ndodhin janë:

- Konvertimi dhe normalizimi që punon në të dy anët e informacionit për t'i kthyer të dhënat uniforme;
- Procesi i lidhjes që realizon bashkimin logjik të fushave ekuivalente nga burime të ndryshme;
- Përzgjedhja që ul numrin e fushave dhe rekordeve të burimit.

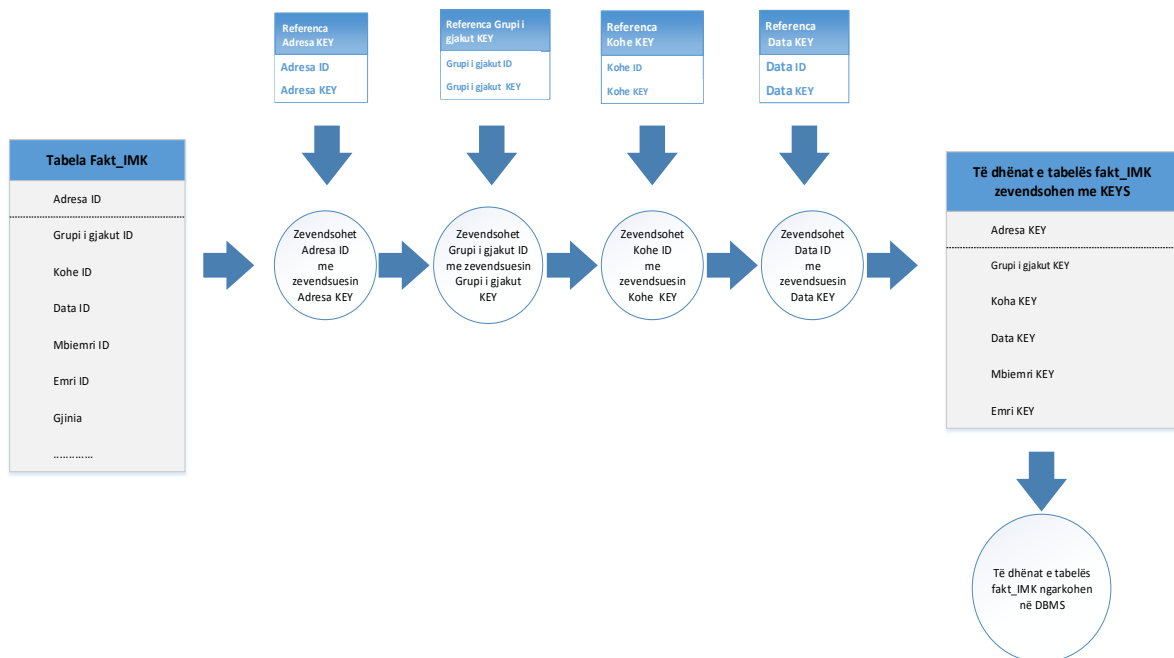
Hapi 4: Ngarkimi i të dhënave

Ngarkimi i informacionit në tabelat dimension dhe fakt kërkon kujdes të veçantë. Ky proces kujdeset për të gjithë hapat që nevojiten për të administruar dimensionet që ndryshojnë pak, të njohura me termin SCD dhe shkruajnë në disk informacionin e plotë dimensional me të gjithë çelësat e dimensioneve SK dhe NK.

Gjatë procesit të ngarkimit të tabelës Fakt IMK shfrytëzohen referencat e krijuara. Të gjitha tabelat e referencës që nevojiten duhet të ruhen në memorie, që të mund të kapen shumë lehtë sa herë procesohet një rekord i tabelës Fakt me informacionin e çelësit natyral të dimensionit.

Prandaj, këto tabela reference janë të veçanta nga tabelat origjinale Dimension të data mart-it. Figura e mëposhtme tregon proceset që ndjek shënimet bazike të pacientit, para ngarkimit në një tabelë Fakt. (Juliana, 2014)

Figura 5.9 Diagrama e ngarkimit të tabelës Fakt IMK (Johnson, 2013)

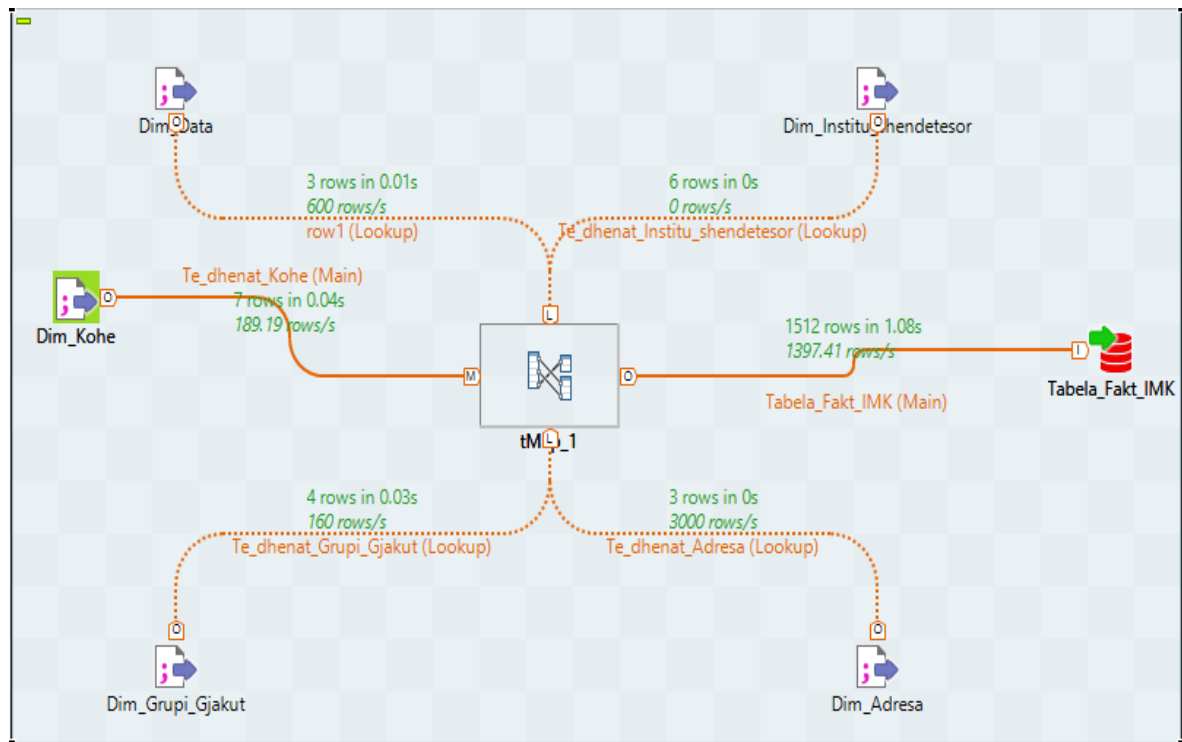


Qëllimi është të tregojë se Talend Open Studio for Big Data (TOS) është në gjendje t'i bëjë të njëjtat detyra si çdo vegël komerciale ETL.

Për të vërtetuar këtë, ne do të përdorim të dhëna të caktuara të marra për këtë nga dy DBMS të ndryshëm, që janë Oracle dhe EXCELFIL, për të testuar procesin ETL në një mjedis akademik. Krahës kësaj, disa fajlle të shënimeve, si: Dim_Adresa, Dim_grupi_gjakut, Dim_Data, Dim_Kohe dhe Dim_institu_shendetsor, teksti në formatin Microsoft Excel është shtuar si e dhënë burim për t'u integruar me DBMS. Oracle është zgjedhur si databazë

kryesore, sepse shumica e kompanive përdorin Oracle në aplikacionet e tyre. Për rastin e ngarkimit të të dhënave në tabelën Fakt IMK është ndërtuar job-i i mëposhtëm në Talend Open Studio for Big Data, siç edhe paraqitet në figurën e mëposhtme.

Figura 5.10 Job-i Talend për ngarkimin e tabelës Fakt IMK



Ndërkohë, në Oracle, një tabelë e njohur si FAKT_IMK, krijohet për të magazinuar të dhënat si të Institutit Mjekësor të Kosovës, duke përfshirë: nr. rendor i pacientëve, mbiemri, emri, mbiemri i vajzërisë, gjinia, grupi i gjakut, Rh faktori, alergjitë, vendbanimi, adresa, telefoni kontaktues, e-maili, medikamentet e rekomanduara. Ndërsa, fajllët në microsoft Excel përmbajnë të dhënat për kohën, adresën, datën, grupin e gjakut dhe institucionin shëndetësor.

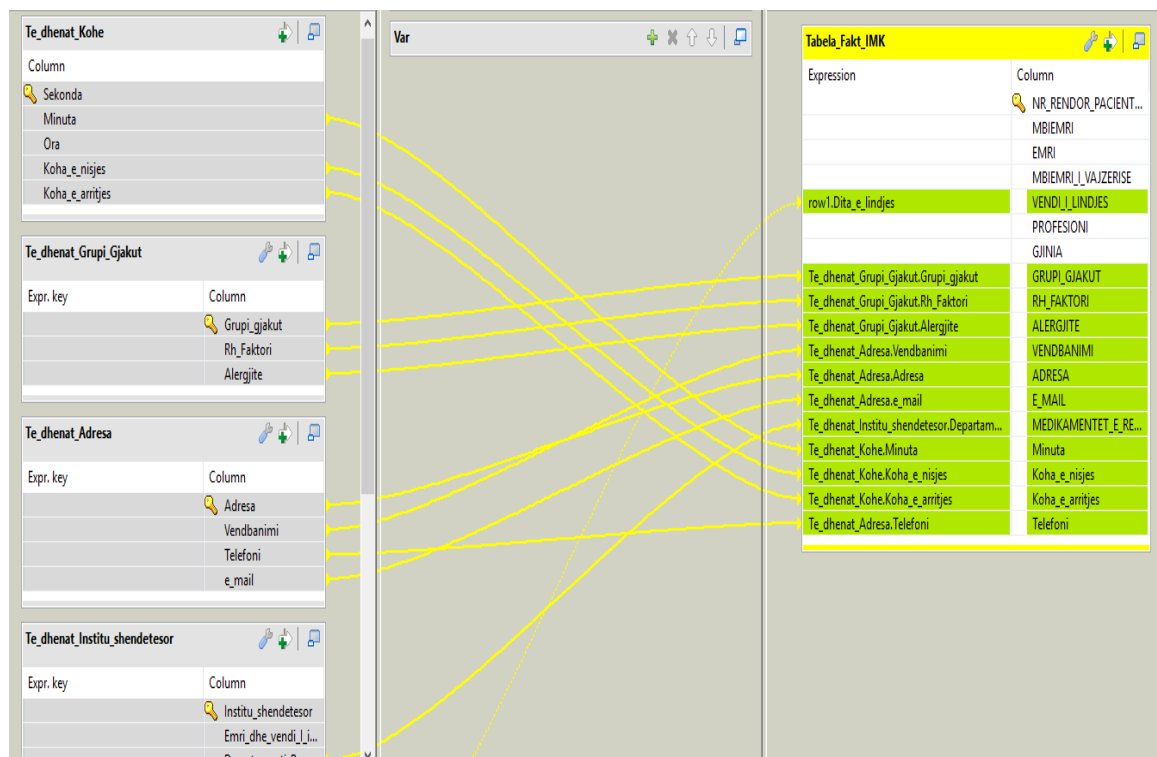
(Azwa & Abdul)

Gjatë procesit të ngarkimit mund të bëjmë edhe shtimin e shumë elementëve të informacioneve, siç është rasti i fushave që na paraqiten pas planifikimit, si: mbiemri, emri, gjinia, që duhet për ta mbajtur në tabelë kohën e ngarkimit të rekordit në databazën, si në figurën më poshtë.

Këta elementë të shtuar nuk gjenden në të dhënat burim, por gjenerohen nga sistemi me anë të funksioneve standarde të Talend Open Studio for Big Data (TOS).

Fushat që përmendëm shërbejnë për menaxhimin e ngarkimit të informacionit, apo për investigime në rast problemesh. (Juliana, 2014)

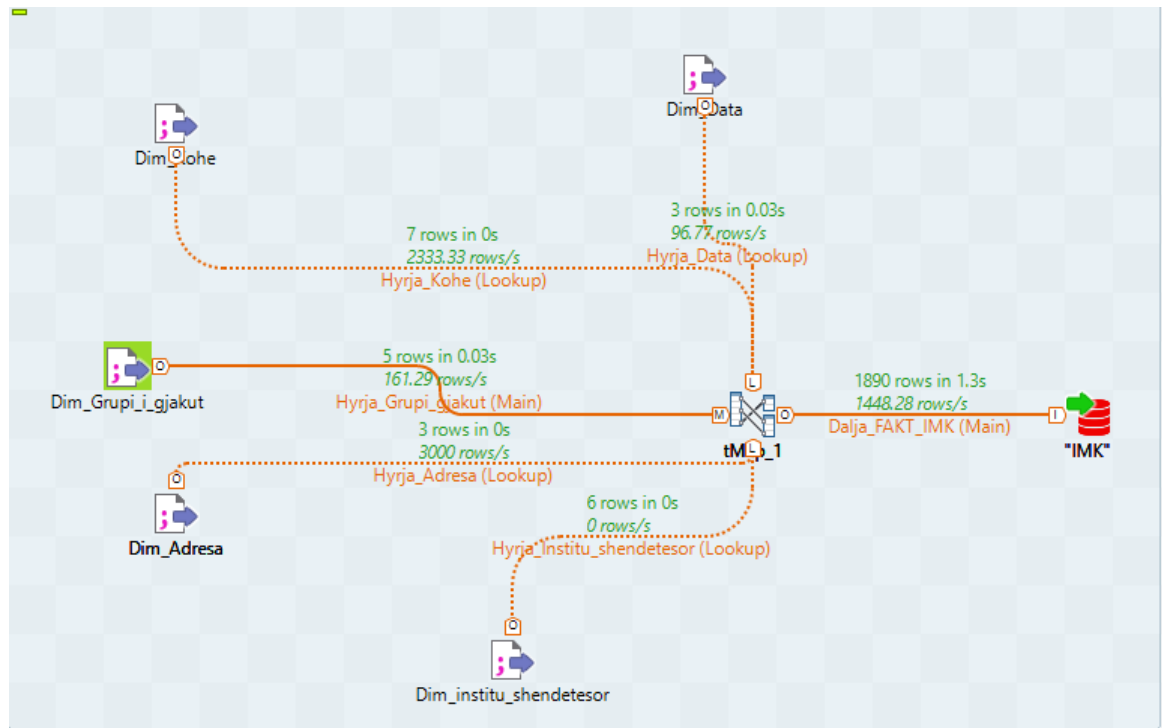
Figura 5.11 Job-i TOS për shtim informacioni



Kjo do të thotë, çdo tip rekordi në fajllin burim ka një numër të caktuar fushash të ndryshme nga tipat e tjerë. Nevojitet të ndërtohet një proces parësimi mbi rekordet e IMK-së. Ai proces

do të realizojë strukturimin dhe shtimin e të dhënave, siç është paraqitur edhe në figurën e mëposhtme.

Figura 5.12 Job-i TOS për shtim dhe ngarkim të informacionit



Hapi 5: Monitorimi i ekzekutimit dhe automatizimi i procesit në Talend Open Studio for Big Data (TOS)

Një ndër sfidat më të mëdha, me të cilën përballen zhvilluesit e proceseve integruese dhe stafi i operacioneve IT, i ngarkuar me kontrollimin e ekzekutimit të tyre, është të qenët i aftë për të kontrolluar dhe monitoruar ekzekutimin e këtyre proceseve delikate.

Në fakt, dështimi në menaxhimin e këtyre proceseve duhet të përfshihet në procesin e integritetit të të dhënave. Për më tepër, përtej lajmërimit për gabimet, shpesh është kritike që të monitorohet e tërë gjendja e procesit të integritetit dhe të shihet për ndonjë përkeqësim në performancë.

Activity Monitoring Console monitoron punën (sukseset, dështimet, paralajmërimet etj.), kohën e ekzekutimit dhe volumin e të dhënave nëpërmjet një konsole të vetme që vepron e pavarur.

Talend Open Studio for Big Data ofron veçori të ndryshme informuese gjatë ekzekutimit, si, për shembull, statistika dhe gjurmë, duke lehtësuar monitorimin e punës dhe punën e gjetjes dhe zgjidhjes së defekteve (debugging).

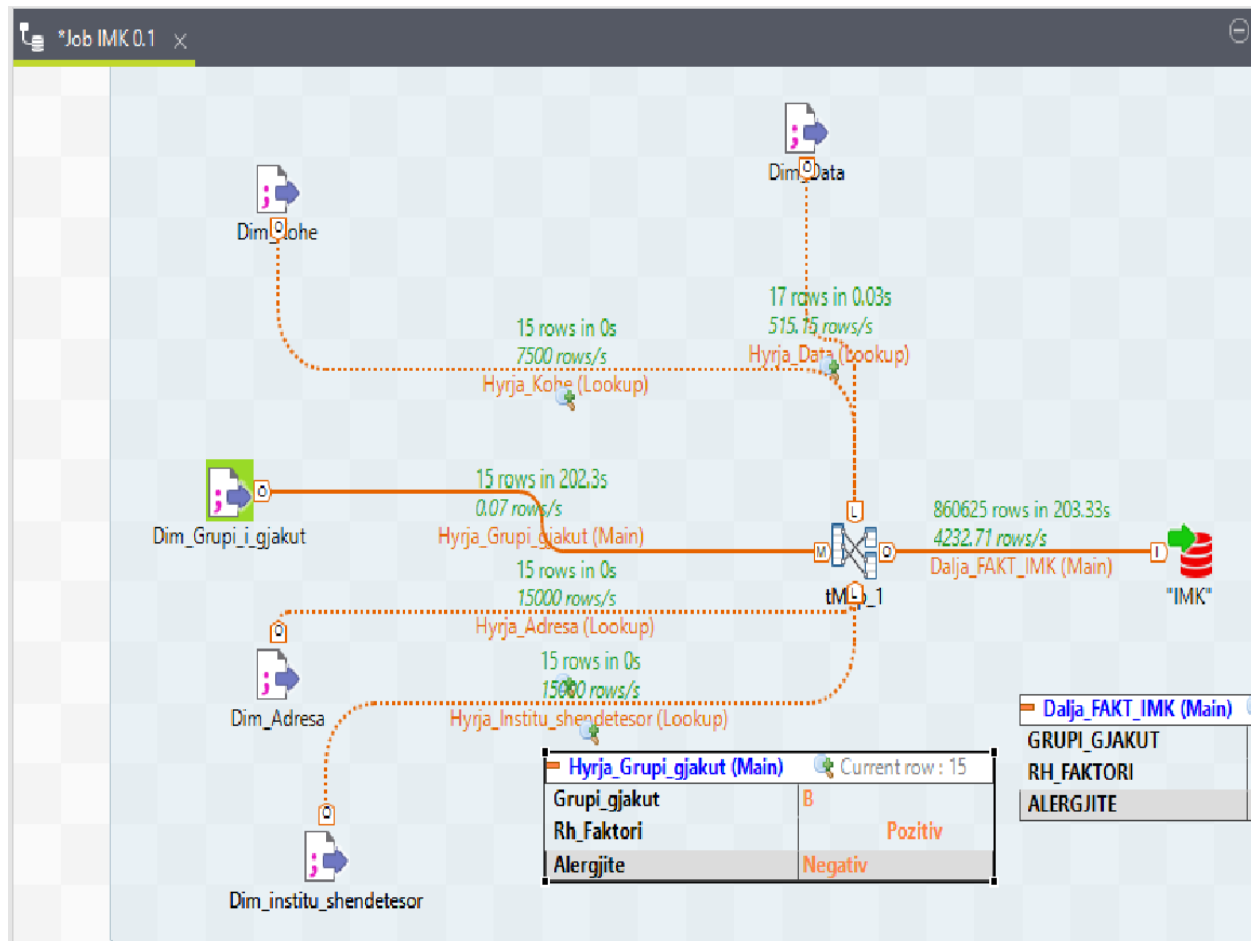
Andaj, veçoria e gjurmimit bën të mundur të monitorohen proceset e të dhënave kur janë duke bërë ekzekutimin e Job-it në Talend Open Studio for Big Data.

Kjo veçori u jep një pamje rresht pas rreshti të sjelljes së komponentit dhe u shfaq rezultate dinamike përkaj linkut të rreshtit në hapësirën e dizajnit.

Me këtë mund të aktivizohet a çaktivizohet Traces ose të vendoset se cilat kolona të procesuara të shfaqen në tabelën e përcaktuar që shfaqet në hapësirën e dizajnit kur të nisni Job-in në fjalë. Për të aktivizuar modalitetin Traces në një Job:

- Klikojmë pamjen Run.
 - Më pas klikojmë tabin Debug Run për të pasur akses në modalitetet debug dhe traces.
 - Klikojmë shigjetën që tregon poshtë butonit Java Debug dhe zgjidh optionin Traces Debug.
- Shfaqet një ikonë nën secilën prej rrjedhave të Job-it, që tregon se është aktivizuar procesi i monitorimit. Zgjedhim Traces Debug për të ekzekutuar Job-in në modalitetin Traces, e që e kemi pasqyruar në figurën 5.13.

Figura 5.13 Monitorimi i ekzekutimit të Job-it në Talend Open Studio for Big Data



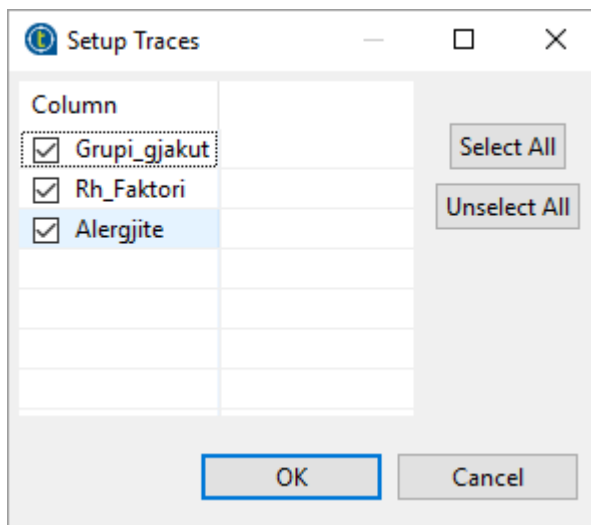
Ky tipar lejon monitorimin e gjithë përbërësve të një Job-i, pa qenë nevoja të kalohet në modalitetin debug, domethënë pa pasur nevojë të keni shumë njohuri për Java.

Funksioni Traces shfaq përmbajtjen e rreshtave të proceduar në një tabelë, ku zgjidhet ana e tastit të djathtë së ikonës Traces nën rrjedhën adekuate.

Që të zgjidhet se cilat nga kolonat e të dhënave të proceduara duhet të shfaqen në tabelën e gjurmimit, japim sa vijon:

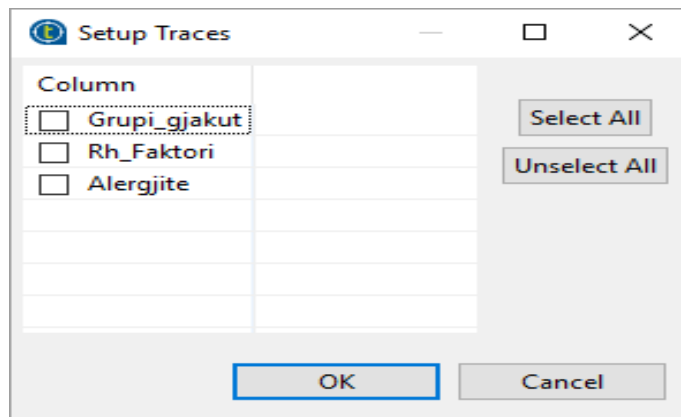
- Më pas klikojmë me të djathtën në ikonën Traces për rrjedhën që të duhet, pastaj zgjidhet nga lista Setup Traces. Ku shfaqet kutia dialoguese si në figurën 5.14

Figura 5.14 Setup Traces, lista përzgjedhëse



Më pas, në kutinë e dialogut hiqen shenjëzimet që u korrespondojnë kolonave që më pas të mos na paraqiten në Traces table, siç paraqitet në figurën 5.15

Figura 5.15 Setup Traces, lista jopërzgjedhëse



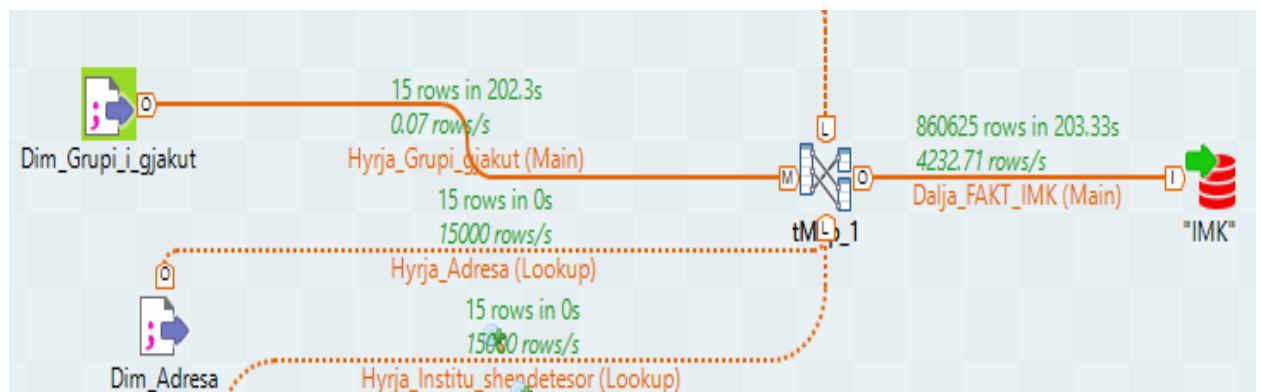
Monitorimi i procesimit të të dhënave nis kur ekzekutohet Job-i dhe përfundon në fund të ekzekutimit.

Andaj, për ta bërë më të lehtë ekzekutimin e Job-ave merren parasysh këta tiparë:

- **Statistics**, ky tipar shfaq normën e performancës së procesimit.
- **Exec time**, ky tipar shfaq kohën e ekzekutimit në konsolë në fund të ekzekutimit.
- **Save Job before execution**, ky tipar lejon të ruhet Job-i automatikisht para ekzekutimit.
- **Clear before run**, ky tipar fshin të gjitha rezultatet e ekzekutimit të mëparshëm përpara se të riekzekutojë Job-in.
- **JVM Setting**, ky tipar u lejon të definoni parametrat e JVM-së suaj sipas nevojave tuaja, për shembull, parametrat e përdorur për të shfaqur karaktere speciale.

Pra, Tipari Statistics shfaq secilën normë të performancës së komponentëve nën linçet e rrjedhës në hapësirën e dizajnit, siç edhe paraqitet në figurën 5.16

Figura 5.16 Paraqitja e tipareve Statistics



Kështu, tipari Statistics tregon numrin e rreshtave të proceduar dhe kohën e procedimit në sekonda, duke u lejuar të dalloni menjëherë çdo proces bllokues (bottleneck) në rrjedhën e procesimit.

Për linqe ‘trigger’ si OnComponentOK, OnComponentError, OnSubjobOK, OnSubjobError dhe If, opsioni Statistics shfaq gjendjen e ‘trigger’ gjatë ekzekutimit të Job-it, ku paraqitet Error dhe True ose False. Në pamjen Run klikohet në tabin e Advanced Settings dhe selektoni kutinë e shenjzimit Statistics për të aktivizuar tiparin e statistikave dhe hiqni shenjën nga kutia për ta bllokuar atë.

Përlllogaritja nis vetëm kur nis ekzekutimi i punës dhe ndalet në fund të tij. Më pas klikohet butoni Clear nga pamjet Basic ose Debug Run për të hequr shfaqjen e statistikave të përlllogaritura. Selektoni kutinë e shenjzimit Clear before Run për të resetuar tiparin Status para çdo ekzekutimi. Tipari i statistikave e ngadalëson ekzekutimin e Job-it, sepse Job-i duhet t’i dërgojë këto statistika të dhënash te hapësira e dizajnit që të shfaqen.

Për të shfaqur kohën totale të ekzekutimit të Job-it pas ekzekutimit të saj, selektoni tabin Advanced Settings në pamjen Run dhe shenjzoni kutinë Exec time para se të nisni Job-in. Mund ta pastroni edhe hapësirën e dizajnit para çdo ekzekutimi të Job-it, duke selektuar kutinë Clear before Run.

5.3 Problematikat e paraqitura në procesin e modelimit të data mart-it IMK

Është e qartë se përpara se të ndërtohej logjika e proceseve ETL, kërkohet të bëhet studimi i burimeve për të zhvilluar mënyrën e ngarkimit të informacionit në data mart. Proceset ETL janë në avancim të vazhdueshëm, për arsye të ndryshimeve apo të përmirësimeve që mund t’i ketë si në sistemet burim apo edhe në skemën e DWH-së.

Në ditët e sotme, shumë burime të të dhënave të Agjencisë së menaxhimit të emergjencave i grumbullojnë dhe i analizojnë të dhënat. Për këtë arsye modelimi i një data mart-i për këto të

dhëna është shumë interesant për t'u zhvilluar. Ky data mart mund të zgjerohet për të përfshirë në të edhe informacione shtesë që do të ndihmonin më tej në analizë. (Juliana 2014).

Andaj, problemet të cilat janë hasur gjatë procesit të modelit data mart IMK janë të paraqitura si më poshtë:

- Ngarkimi fillestar i një tabele të re parashtrohet një sërë sfidash unike. Sfidat kryesore është të menaxhohet volumi i jashtëzakonshëm i të dhënave. Kështu, shumë vegla ETL (dhe disa databaza) ofrojnë funksionalitetin update else insert. Ky funksionalitet është shumë i leverdishëm dhe thjeshton logjikën e rrjedhjes së të dhënave, por është shumë i ngadaltë. Proceset ETL, që kërkojnë përditësime për të dhënat ekzistuese, duhet të përfshijnë logjikën e cila ndan shënimet që ekzistojnë nga ato të reja. Kur kemi të bëjmë me një sasi të madhe të dhënash, ato do të ngarkohen përnjëherë në një datawarehouse (DWH).

Fatkeqësisht, shumë vegla për ngarkim të madh të dhënash ose bulk-load nuk mund të përditësojnë shënimet ekzistuese. Duke ndarë edhe përditësimet nga shtesat, së pari, duhet të procesohen përditësimet dhe, pastaj, të ngarkohen përnjëherë me balancën e shënimeve për të pasur performancë optimale të ngarkimit të të dhënave.

Andaj, përdorimi i një vegle të tillë, si rasti SQL INSERT për të ngarkuar të dhëna, ul shumë koston e punës dhe rrit në mënyrë drastike performancën e ngarkimit të të dhënave.

Por, kur veprohet në ngarkimin paralel, me këtë rast, ngarkohet një volum i madh i të dhënave, që duhet ndarë në të dhëna fizike dhe në segmente logjike. Nëse, p. sh., bëjmë ngarkimin e të dhënave në pesë viteve, më mirë do të ishte t'i bëni pesë fajlla me të dhëna me nga një vit veç e veç. Disa vegla ETL mundësojnë të bëhet ndarja e të dhënave, duke u bazuar në një varietet të dhënash. Pasi të dhënat të ndahen në segmente të barabarta, fillohet me procesin ETL për t'i ngarkuar të gjitha segmentet në formën paralelisht.

Minimizimi i përditësimeve fizike

Përditësimi i shënimeve në një tabelë kërkon kosto të jashtëzakonshme në DBMS, pjesa më e madhe e së cilës shkaktohet nga popullimi i rollback log nga databaza. Në shumë raste është më mirë të fshihen shënimet që do të përditësohen dhe më pas të ngarkohen versionet e reja të këtyre shënimeve së bashku me të dhënat-shënimet që hyjnë në DWH për herë të parë. Duke qenë se raporti i shënimeve që përditësohen përballë numrit të rreshtave ekzistues luan rol kryesor në përzgjedhjen e teknikës optimale, zakonisht, kërkohen edhe testime që do nxjerrin edhe strategjinë e duhur të ngarkimit të të dhënave për çdo situatë.

Ngarkimi në rritje është proces që ndodh periodikisht për të mbajtur depozitën e të dhënave të sinkronizuara me sistemet e veta burimore. Proceset e tilla mund të veprojnë në intervale të caktuara ose vazhdimisht. Në momentin e shkrimit të këtij materiali, intervali kohor për ngarkimin e të dhënave në vendin e depozitimit DWH është një herë në ditë, por nuk ka rregulla të përcaktuara strikte në lidhje me intervalet e tilla.

Përdoruesit, zakonisht, pëlqejnë përditësime ditore, sepse i lënë të dhënat statike në DWH gjatë ditës, duke parandaluar kështu të dhënat e ndryshueshme, që mund të shkaktojnë ndryshim të vazhdueshëm të të dhënave dhe inkonsistencë ditore.

Rutina e ETL-së, që ngarkon të dhëna në rritje, është zakonisht rezultat i proceseve që në fillim ngarkonin të dhënat historike në DWH.

Është një praktikë e preferuar që të mbahen të ndara këto dy procese. Skuadra ETL duhet të vendosë datën e fillimit dhe të fundit të procesit të nxjerrjes së të dhënave, kështu që rutina ETL të ketë fleksibilitetin që të ngarkojë segmente të vogla në rritje ose të dhënat historike në tërësi. (Kimball&Caserta, 2004).

5.3.1 Rezultatet e përfituara nga analizat për rastin e data mart-it IMK

Analizat për rastin e data mart-it IMK, ku data mart-i IMK është ndërtuar të funksionojë në shkëmbimin dhe integrimin e të dhënave nga shumë dimensione (DIM_KOHE, DIM_GRUPI_GJAKUT, DIM_ADRESA, DIM_DATA, DIM_INSTITU_SHENDETSOR) në një fakt (FAKT_IMK) të vetëm.

Kështu, modeli i data mart-it të mësipërm ofron performancë të lartë dhe të shpejtë për gjenerim të rezultateve dhe të informatave të kërkuara, që mundëson prodhimin e analizave brenda afatit të kërkuar. Si ky model i shpjeguar data mart-i, mund të ndërtohen edhe model të tjerë të data mart-ëve, që shërbejnë për gjenerimin e informacioneve të të dhënave në internet dhe në pajisje të vogla mobile.

Duke marrë parasysh këtë data mart IMK, përfitojmë analiza të larmishme për formën e komunikimit në Institutin e Mjekësisë së Kosovës. Nëse merren si shkas kërkesat e personelit mjekësor për informacionet që lidhen me pacientët, mund të merren lehtësisht përgjigje në kërkesat e parashtruara, duke u bazuar në tabelat dimension.

Këtu llogaritet edhe koha e procedurave që realizojnë nxjerrjen, ngarkimin dhe transformimin e të dhënave. Andaj, tabela e mëposhtme jep rezultatin e përftuar nga këto procedura.

Tabela 5.2 Krahاسimi i ekzekutimit të job-eve

Emri i Job-it	Frekuenca	Kohezgjatja
DIM_DATA	Rows/s=600;ByHour=0;ByMinute=0;BySecond=0.1	00:00:01
DIM_KOHE	Rows/s=189.19;ByHour=0;ByMinute=0;BySecond=0.4	00:00:04
DIM_GRUPI_GJAKUT	Rows/s=160;ByHour=0;ByMinute=0;BySecond=0.3	00:00:03
DIM_ADRESA	Rows/s=3000;ByHour=0;ByMinute=0;BySecond=0	00:00:00
DIM_INS_SHENDET	Rows/s=6;ByHour=0;ByMinute=0;BySecond=0	00:00:00
FAKT_IMK	Rows/s=1397.41;ByHour=0;ByMinute=1;BySecond=06	00:01:06

Nga rezultatet e mësipërme duket që DIM_DATA, nga 600 rreshta/s barten për 0.1 sekondë, DIM_KOHE, nga 189 rreshta/s barten për 0.4 sekondë, DIM_GRUPI_GJAKUT, nga 160 rreshta/s barten për zero sekondë, DIM_INS_SHENDET, nga 6 rreshta/s barten për zero sekondë. Kështu, të gjitha të dhënat nga dimensionet barten për në tabelën FAKT_IMK me 1397 rreshta/s me një kohe totale prej 1.6 min.

Kështu që, për rastin e data mart-it IMK varet nga rifreskimi i datawarehouse (DWH) në një interval të caktuar (ose në vazhdimësi). Dështimi i realizimit të kësaj përgjegjësie vë në pikëpyetje besueshmërinë dhe varshmërinë e të dhënave në data mart-it IMK. Për këtë arsye, datawarehouse (DWH) nuk mund të ketë sukses pa mekanizmin efikas dhe të vazhdueshëm të furnizimit me të dhëna.

Skuadra ETL duhet të monitorojë dhe të vlerësojë punët ETL për t'u siguruar që po operojnë në mënyrë efikase dhe se datawarehouse (DWH) po ngarkohet në mënyrë efikase. Monitorimi i detyrave ETL merr në konsideratë shumë aspekte të procesit.

Burimet e jashtme në raport me qëllimin e sistemit ETL, si administrimi i pajisjeve harduerike dhe infrastrukturës dhe përdorimit, si dhe ambienti burimor dhe target, luajnë rol vendimtar në efikasitetin e sistemit ETL. Këtu do të paraqesim disa tregues të performancës së ETL-së që do t'u tregojnë se sa mirë apo jo janë duke u ekzekutuar proceset tuaja. Treguesit janë pjesë e metadata-ve operationale, që duhet të vendosen në një depo me qëllim që të analizohen me kalimin e kohës nga skuadra ETL.

Matja e treguesve të performancës së ETL-së

Ekzistojnë matje specifike që ndihmojnë të kuptohet nëse performanca është ajo e duhura. Siç pritet, sistemi ETL ka treguesit e vet të performancës. Treguesit ETL janë specifike në lidhje me lëvizjet fizike dhe menaxhimin e të dhënave. Ata janë një hap më poshtë se treguesit tipikë të performancës.

Kur themi poshtë themi se ata nuk matin sistemin operativ ose burimet harduerike, por vetë procesin ETL. Matja kryesore e performancës së efikasitet të ETL-së është koha që i duhet për të proceduar të dhënat. Në vijim janë matjet specifike që rezultojnë të dobishme në matjen e performancës.

Kohëzgjatja në sekonda. Kjo përlllogaritje është baza e të gjitha përlllogaritjeve të tjera. Kohëzgjatja është dallimi mes kohës së fillimit dhe kohës së përfundimit të një procesi ETL në sekonda. Për shembull, nëse një proces nis në orën 4:00 a.m. dhe kompletet në 4:15 a.m., kohëzgjatja është 900 sekonda.

Rreshtat e proceduar për sekondë. Kjo është ekuivalentja e rreshtave të ngarkuar për sekondë, me përjashtim të rasteve ku të dhënat burimore janë më të mëdha se ato target, si

në rastin e ngarkesës totale. Një shembull i përlogaritjes rresht për sekondë është 1.000.000 rreshta të proceduar në 15 minuta ($1000000 / (15 * 60) = 1111.11$ rresht/sek).

Rreshta të lexuar për sekondë. Numri i rreshtave të rezultatit të SQL-së që merr të dhënat nga sistemi burimor i pjesëtuar me kohëzgjatjen në sekonda. Të dhënat më pas kalojnë përmes proceseve transformuese në ETL, ku numri i rreshtave mund të rritet ose ulet varësisht nga procesi.

Rreshta të shkruar për sekondë. Numri i rreshtave në tabelën targ pasi janë transformuar, pjesëtuar me kohëzgjatjen në sekonda. Në rastet ku kemi tabela me target të shumëfishtë, është shuma e të gjithë rreshtave të vendosur në të gjitha tabelat pjesëtuar me kohëzgjatjen në sekonda. (Kimball&Caserta, 2004).

5.4 Sinkronizimi i të dhënave

Sinkronizimi i të dhënave, zakonisht, kërkohet për t'i mbajtur dy ose më shumë sisteme të sinkronizuara dhe të përditësuara. Një arsye e zakonshme për sinkronizimin e të dhënave është migrimi i sistemit (si rezultat i konsolidimit të biznesit, zvogëlimit etj.), ku kërkohet vënia në punë e dy sistemeve paralel për një periudhë kohe, ndonjëherë disa vjet.

Gjatë kësaj periudhe transaksionet që kapen nga një sistem duhet të update-ohen në sistemin tjetër, gjithashtu. Një arsye tjetër e zakonshme për sinkronizimin e të dhënave është shkrirja ose blerja (M&A), ku sistemet e përdorura nga institutet e menaxhimit të emergjencave vihet sipër tjetrës dhe të dyja duhet të jenë të update-uara.

CDC bën të mundur kapjen e ndryshimeve, ndërsa ndodhin dhe i përpunon, duke përdorur mjete të tilla, si EAI, ETL ose programet e ndërtuara vetë. Duke inkorporuar CDC sinkronizimi i të dhënave mund të bëhet i efektshëm dhe në kohë reale (Efficient and Real-

time Data Integration, 2009). Një DWH siguron informacione për përpunim analitik, vendimmarrje dhe mjete të shfrytëzimit të të dhënave.

Një DWH grumbullon të dhënat nga shumë sisteme me burim operacional heterogjen (OLTP - On-Line Transaction Processing) dhe ruan të dhënat e përmbledhura të integruara të AME-së në një depo qendrore të përdorur për aplikimet analitike (OLAP - On-Line Analytical Processing), me kërkesa të ndryshme përdoruesi. Vendi i të dhënave të një DWH, zakonisht, ruan historinë e plotë të një AME-je. Procesi i zakonshëm për marrjen e një informacioni të vendimmarrjes bazohet në përdorimin e mjeteve OLAP. Këto mjete e kanë burimin e tyre të të dhënave të bazuar në një vend të dhënash DWH, në të cilin update-ohen regjistrimet nga mjetet ETL (ekstraktim, transformim dhe ngarkim).

Proceset ETL janë përgjegjëse për identifikimin dhe ekstraktimin e të dhënave përkatëse nga sistemet e burimit OLTP, duke i personalizuar dhe integruar këto të dhëna në një format të përbashkët, duke i pastruar të dhënat dhe duke i konfirmuar në një format të integruar të përshtatshëm për update-imin e vendit të të dhënave të DWH-së dhe, së fundi, ngarkimin final të të dhënave të formatuara në databazën e tyre.

Tradicionalisht, është pranuar gjerësisht se databazat e DWH-së përditësohen rregullisht, zakonisht në ditë, javë ose muaj, duke treguar se të dhënat e saj nuk janë asnjëherë të update-uara, pasi rekordet OLTP, të ruajtura midis atyre update-imeve, nuk janë përfshirë në vendin e të dhënave.

Kjo tregon se rekordet operationale më të fundit nuk janë përfshirë në vendin e të dhënave, duke u përjashtuar kështu nga rezultatet e siguruara nga mjetet OLAP. Deri kohëve të fundit përdorimi i të dhënave të update-uara periodikisht nuk ishte problem. Sidoqoftë, me sipërmarrjet, si, biznes, shkëmbime burse, telekomunikacione online dhe sisteme

shëndetësore, për shembull, informacionet përkatëse duhet të jepen sa më shpejt që të jetë e mundur për punonjësit e njoftimeve ose sistemet e vendimeve të cilët mbështeten në një mënyrë thuajse në kohë reale, sipas të dhënave të reja dhe më të fundit të kapura nga sistemi i informacionit të QOAME.

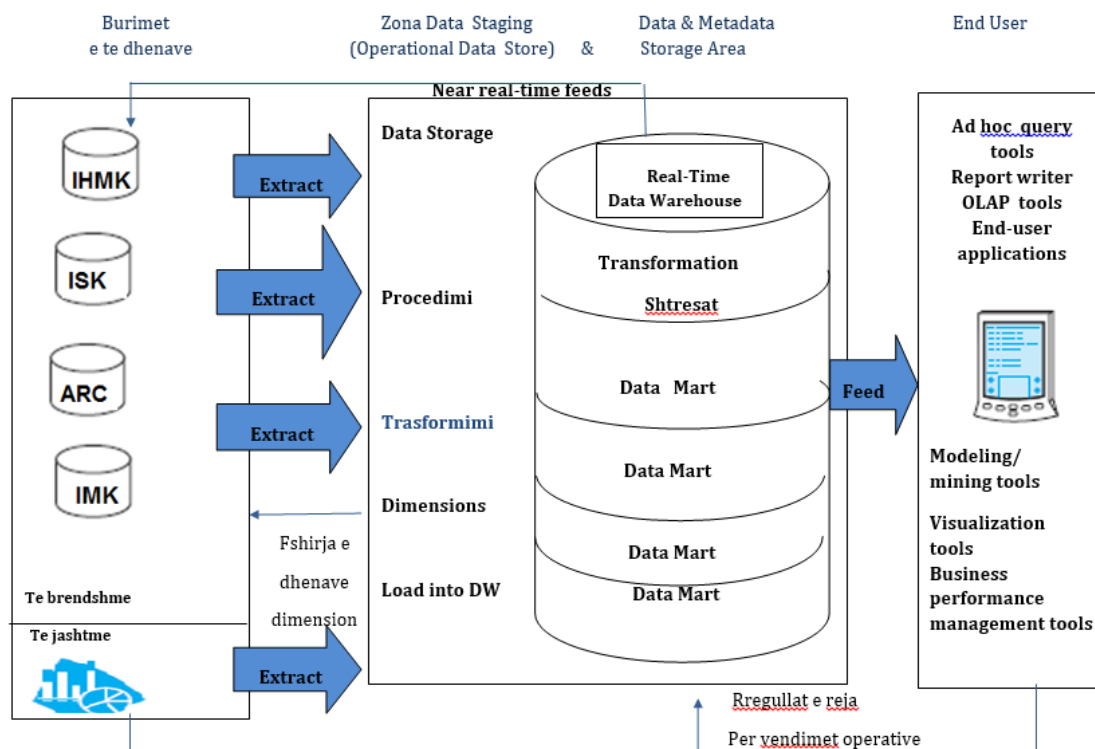
Kjo e bën mbështetjen DWH-së në kohë reale (RTDW) një çështje kritike për këto lloje aplikimesh. Kërkesa për të dhëna të freskëta në DWH-së ka qenë gjithmonë një dëshirë e fortë. Rifreskimi i DWH-së (integrimi i të dhënave të reja) kryhet tradicionalisht në një metodë off-line. Kjo do të thotë se teksa po kryhen proceset për update-imin e të dhënave, përdoruesit dhe aplikacionet OLAP nuk mund t'i përdorin të dhënat.

Ky set aktivitetesh, zakonisht, bëhet në një dritare ngarkimi me kohë të përcaktuar, për të shmangur mbingarkimin e sistemeve të burimit OLTP operacional me ngarkesë ekstra. Dhe, përsëri, përdoruesit po kërkojnë nivele më të larta rifreskimi, meqenëse gjithnjë e më shumë agjencia e fatkeqësive natyrore operon në një orar 24x7. Magazinimi aktiv i të dhënave i referohet një trendi të ri, ku DWH përditësohet sa më shpesh të jetë e mundur, për shkak të kërkesave të mëdha të përdoruesve për të dhëna të freskëta.

5.5 Arkitektura e data mart-ëve logjike dhe i DWH-së në kohë reale

Arkitektura e data mart-ëve logjike dhe DWH-së në kohë reale është praktike vetëm për DWH me përmasa mesatare ose kur përdoren teknologji DWH me performancë të lartë, si, për shembull, sistemi Teradata. Siç mund të shihet në figurën 5.17, kjo arkitekturë ka karakteristikat e mëposhtme unike.

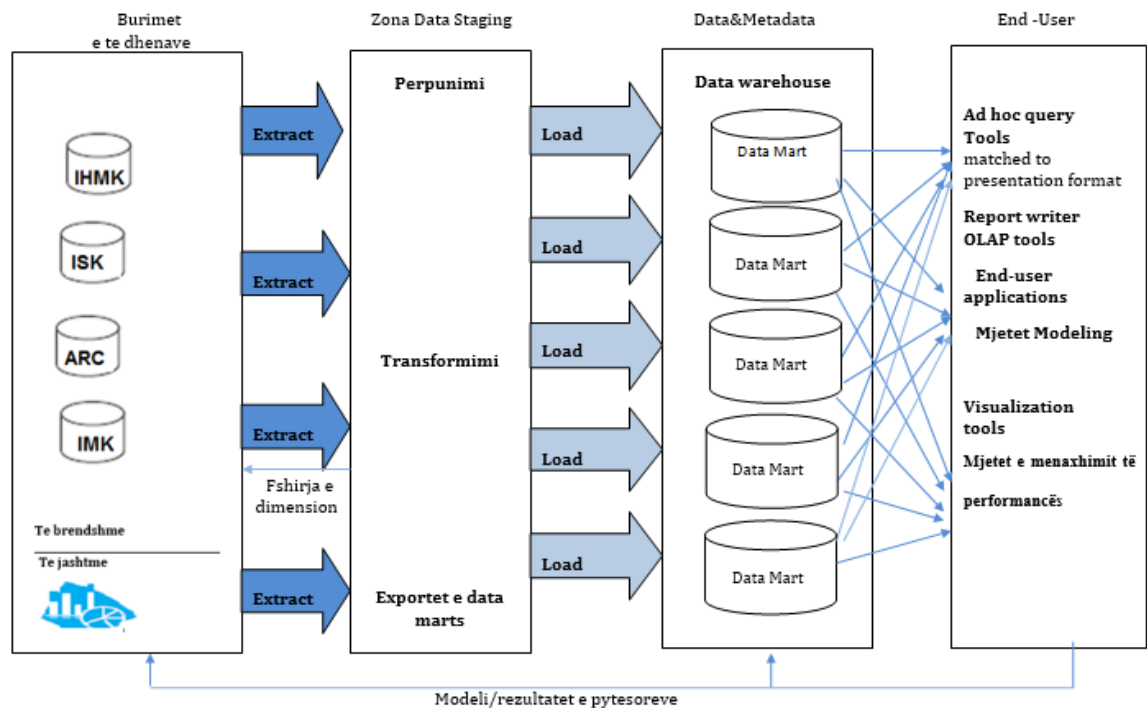
Figura 5.17 Data mart-ët logjikë dhe arkitektura në kohë reale e datawarehouse (Hoffer&Ramesh, 2011).



1. Data mart-ët logjikë nuk janë databaza fizikisht të veçantë, por pamje relative të ndryshme të një DWH relative fizike, paksa të denormalizuar.
2. Të dhënat lëvizin në DWH në vend ose në një fushë të veçantë skenimi, për të përdorur fuqinë kompjuterike me performancë të lartë të teknologjisë së magazinës për të kryer hapat e pastrimit dhe transformimit.
3. Data mart-ët e rinj mund të krijohen shpejt, sepse asnjë databazë fizike ose teknologji databaze nuk ka nevojë të krijohet ose merret dhe nuk ka nevojë të shkruhen rutina ngarkimi.
4. Data mart-ët janë gjithmonë të update-uar, sepse të dhënat në një pamje krijohen kur pamja është referencuar, pamjet mund të materializohen nëse një përdorues ka një seri kërkesash dhe analizash që duhet të shlyejnë të njëjtin ilustrim të data mart-it.

Rolet e ndryshme, qoftë logjik apo fizik te data mart-ët dhe DWH-ja janë përmbledhur në figurën 5.18

Figura 5.18 Arkitektura datawarehouse për data mart-ët e pavarur (Hoffer&Ramesh, 2011).



Megjithëse me shtrirje të kufizuar, një data mart mund të mos jetë i vogël. Prandaj, teknologjia e shkallëzimit shpesh është kritike. Një barrë dhe kosto e madhe vihet mbi përdoruesit kur ata vetë duhet t'i integrojnë të dhënat nëpër data mart të veçantë fizik (nëse kjo është e mundur). Ndërsa shtohen data-mart-ët, një DWH mund të ndërtohet në faza. Mënyra më e lehtë që të ndodhë kjo është që të ndiqet arkitektura e data mart-ëve logjike dhe e DWH-së në kohë reale (Hoffer & Ramesh, 2011).

5.6 Datawarehouse në kohë reale dhe teknikat ETL

Një DWH, disa vite më parë, zakonisht, përditësohej çdo ditë ose çdo javë. Në dy deri tre vitet e kaluar ka pasur gjithnjë e më shumë kërkesa për të rritur shpeshtësinë e update-imeve. Përdoruesit duan që të dhënat në DWH të update-ohen çdo dy minutë ose edhe në kohë reale. Një DWH në kohë reale është një DWH që update-ohet nga ETL në momentin që ndodh transaksioni në sistemin burim.

Për shembull, mund të vendoset shkaku në tabelën e transaksionit të ISK në sistemin burim, në mënyrë që kurdoherë të futet një transaksion në databazë. Shkaku qëllon dhe e dërgon rekordin e ri në DWH si mesazh. DWH ka një dëgjues aktiv që kap mesazhin në momentin që vjen, e pastron, e bën DQ, e transformon dhe e futë në tabelën e fakteve menjëherë. Këtu diskutohet për një diferencë kohore prej dy sekondash, midis momentit që ka ndodhur një fatkeqësi natyrore në një zonë, që të dhënat janë në dispozicion në tabelën e fakteve.

Qasja tjetër e implementimit të një DWH-je në kohë reale është të modifikohet aplikimi burim për të shkruar në vendin e skenimit të DWH-së, menjëherë pasi i shkruan të dhënat në databazën e saj të brendshme. Në databazën e skenimit vendosni shkaku që do të qëllojnë çdo herë që futet një regjistrim i ri dhe këta shkaktare përditësojnë DWH-në.

Qasjet thuajse në kohë reale mund të implementohen, duke përdorur një mini-grup me dy deri pesë minuta shpeshtësi, i cili i nxjerrë të dhënat nga vendi i skenës në vend se duke përdorur shkaku. Ky mini-grup bën, gjithashtu, punën normale ETL, duke transformuar të dhënat dhe duke i ngarkuar në databazën dimensionale të DWH-së. Mini-grupi mund të nxjerrë, gjithashtu, të dhënat direkt nga sistemi burim, duke eliminuar nevojën e modifikimit të sistemit burim për të përditësuar vendin e skenimit (Rainardi, 2008).

DWH-të në kohë reale vijnë nga prirjet e fundit të globalizimit të biznesit, veprimtarive 24/7, volumeve të të dhënave gjithnjë e në rritje, presioneve të konkurrencës, rritjes së kërkesave të menaxhimit të emergjencave dhe rritjes së kërkesave nga vendimmarrësit për të dhëna në kohë reale.

Këto prirje të fundit kërkojnë që menaxhimi i emergjencave të kenë akses në të dhënat më të update-uara për qëllime analize dhe statistike, e cila bën të nevojshme një kërkesë për ndërtimin e DWH-së në kohë reale dhe të ETL-së.

Teknikat për të arritur një DWH në kohë reale përfshijnë teknikën Change Data Capture (CDC) dhe integrimin e kapjes së ndryshimit të të dhënave me proceset ETL ekzistuese për të rritur në maksimum performancën e ETL-së dhe për të arritur ETL në kohë reale.

Integrimi i CDC-së me mjetet ETL ekzistuese siguron një qasje të integruar për të reduktuar sasinë e informacionit të transferuar, duke minimizuar ndërkohë kërkesat e burimeve dhe maksimizuar shpejtësinë dhe efikasitetin.

Në kontrast, migrimi i të dhënave në DWH, duke përdorur mjetet tradicionale ETL, ka një problem në kohëzgjatjen me volumet e mëdha të seteve të të dhënave, pasi proceset ETL konsumojnë burime të konsiderueshme të CPU-së dhe kohë për sete të mëdha të dhënash (Kakish&Kraft 2012:1-12).

Një rast i zakonshëm për përdorimin e CDC-së përfshin procesin e lëvizjes së informacionit në DWH. Tradicionalisht, përditësimet DWH përpunohen me një mjet ETL (ekstrakt, transformim dhe ngarkim).

ETL është një program softueri që ekstraktin të dhënat nga sistemi burim, i transformon dhe pastaj i ngarkon në DWH. Këto procese kërkojnë që sistemi a sistemet operacionale të vihen offline për një periudhë kohe të caktuar.

Kjo periudhë kohe quhet "Batch Window", e matur zakonisht në orë dhe ndonjëherë në ditë, gjatë së cilës sistemi është i zënë me lëvizjen e të dhënave dhe nuk mund të kryejë funksione operacionale dhe funksione të tjera kritike për misionin. Duke pasur parasysh kufizimin e kësaj qasjeje "shumicë", shumica e agjencive të emergjencave përdorin IT që t'i përditësojnë DWH-të e tyre vetëm mbi bazë ditore dhe shpesh javore.

Duke pasur parasysh nevojën e shumë agjencive që menaxhojnë me fatkeqësitë natyrore për informacione të minutës së fundit, ata kanë filluar të kërkojnë mënyra të tjera për të update-uar DWH-të e tyre në kohë reale, duke reduktuar ndjeshëm kohëzgjatjen.

Mjetet EAI merren ndonjëherë në shqyrtim për të arritur këtë synim. Sidoqoftë, CDC siguron një qasje të re për t'i lëvizur informacionet në një DWH dhe mund të punojë, njëkohësisht, me mjetet ETL ose EAI.

CDC i çon ndryshimet në një mjet ETL ose EAI në grup ose në kohë reale, duke lejuar përmirësimin në mënyrë dramatike të efikasitetit të të gjithë procesit, reduktimin ose eliminimin totalisht të 'batch windows', sigurimin e informacionit në kohëzgjatje të ulët dhe reduktimin e kostove përkatëse, përfshirë ciklet e CPU-së, memorien, 'bandwidth' të rrjetit dhe burimet njerëzore (Efficient and Real-time Data Integration, 2009).

Gjithnjë e më tepër ka nevojë për të mbështetur dhe marrë vendime në menaxhimin e fatkeqësive natyrore thuajse në kohë reale bazuar në vetë të dhënat operacionale. Arkitekturat tipike ETL janë të orientuara në përditësim grupi dhe shkaktojnë një boshllëk të madh në aktualitetin e informacionit në DWH.

Ne vëmë në dyshim efektivitetin e performancës së arkitekturave tipike ETL grupi dhe përditësimet me bazë thuajse në kohë reale mbi të dhënat operacionale dhe ngremë pyetje

për të adresuar studimin në kohë të menjëhershme, me qëllim që të adresojmë procesin e vendimmarrjes në AME.

Ngremë çështjen se procesi ETL aktual ka nevojë të largohet nga rifreskimet periodike në update-ime të vazhdueshme. Sidoqoftë, update-imi online i DWH-së ngre sfida të reja të sinkronizimit të të dhënave dhe caktimit të burimeve. "Për t'u përballur me kërkesat në kohë reale, DWH duhet të jetë në gjendje të mundësojë integrimin e vazhdueshëm të të dhënave, me qëllim që të merret me të dhënat më të fundit të Agjencisë së menaxhimit të emergjencave se problemet e sinkronizimit të të dhënave lindin kur shënimet janë të përbëra nga të dhëna të nxjerra nga disa burime të dhënash që po përditësohen pa dallim. Sfidat e burimit vijnë kur ka kërkesa burimesh në konflikt të kërkesave të analizave afatgjata në praninë e përditësimeve në ecuri e sipër.

Në mjetet ETL tradicionale, ngarkimi bëhet rregullisht gjatë kohës joproductive dhe gjatë kësaj kohe askush nuk mund të hapë t'i dhënat në një DWH.

Ndarja midis kërimit dhe përditësimit qartë thjeshtëson shumë aspekte të implementimit të DWH-së, por ka disavantazh të madh, sepse DWH nuk përditësohet vazhdimisht. Mjetet tradicionale ETL nuk janë në gjendje të merren mjaft me këto futje ose update-ime të vazhdueshme pa kohë joproductive në DWH.

Në DWH në kohë reale ngarkimi bëhet vazhdimisht, ndryshe nga një bazë periodike në qasjet tradicionale. Një qasje në arkitekturën e përgjithshme të një DWH-je në kohë reale përbëhet nga elementët e mëposhtëm:

- a) Sistemet e të dhënave që postojnë sistemet e prodhimit të cilat popullojnë DWH-në;
- b) Një zonë përpunimi të dhënash (DPA) e ndërmjetme ku bëhet pastrimi dhe transformimi i të dhënave;

c) Datawarehouse (Kakish&Kraft, 2012:1-12).

Ndërtimi i një zgjidhjeje ETL DWH në kohë reale kërkon klasifikimin e disa objektivave të AME-së shpesh të lëvizshme, kuptimin e një seti të shumëllojshëm teknologjish, pasjen e një ndërgjegjësimi të disa qasjeve pragmatike, të cilat janë përdorur me sukses nga të tjerët dhe zhvillimin e fleksibilitetit dhe kreativitetit inxhinierik.

Kjo fushë mbetet e re, me teknologji të reja, metodologji emergjente dhe fjalorë të rinj. Qartësisht, kjo situatë mund të jetë një dëshmi telashesh, por DWH në kohë reale u ofron, gjithashtu, adoptuesve të hershëm potencial të madh për të arritur një avantazh konkurrues. Ky kapitull propozon një proces me katër hapa për të drejtuar profesionistin e DWH-së përmes zgjedhjes së një arkitekture teknike dhe metodologjie ETL të përshtatshme në kohë reale. Ne kemi treguar katër hapat kryesorë të sistemit ETL në një DWH: nxjerrje, garantim i cilësisë, konfirmim dhe strukturim i të dhënave, si një seri skemash dimensionale gati për t'u konsumuar nga përdoruesit finalë.

5.7 Sfidat dhe mundësitë e DWH-së në kohë reale

DWH në kohë reale përbën një numër unik sfidash dhe mundësish për inxhinierin ETL. Nga këndvështrimi i arkitekturës teknike ka potencialin të ndryshojë qasjen big-beng të nevojshme gjatë grupit të natës ETL në një rrjedhë ngjashëm ETL përgjatë ditës. Kërkesat e disponueshmërisë së sistemit mund të përshkallëzohen, ndërsa biznesi mbështetet në disponueshmërinë e ulët të transaksioneve të biznesit në DWH.

Nëse AME zgjedh qasjet e menaxhimit të dimensioneve në kohë reale të treguara në këtë kapitull, disponueshmëria bëhet një avantazh strategjik. Nga këndvështrimi i arkitekturës së të dhënave, DWH në kohë reale sfidon qëndrimin e DWH-së si sistem i matjeve periodike të

veçanta - një ofrues i pamjeve të biznesit, duke përkrahur kështu një sistem më gjithëpërfshirës dhe informacione të përkohshme të vazhdueshme.

Ky kalim ndodh pak, për shembull, nëse shpeshtësia e ngarkimit të fakteve rritet nga një herë në ditë në çdo 15 minuta, por në mënyrë më dramatike nëse ngarkimi i fakteve dhe regjistrimeve të dimensioneve ndodh vazhdimisht. DWH mund të hapë, pastaj, një regjistrim të transaksioneve të biznesit dhe të kontekstit të tyre dimensional në të gjitha pikat në kohë. Dimensionet që ndryshojnë ngadalë bëhen shpejt dimensione në ndryshim dhe qëndrimi i DWH-së bëhet më operacional nga funksioni. Në fakt, nëse DWH në kohë reale mbështet përputhjen dhe sinkronizimin e dimensioneve në kohë reale, atëherë ajo evoluohet në një shtrirje logjike të vet sistemeve operative.

5.8 Rishikimi i DWH-së në kohë reale

Qasja në kohë reale e DWH-së mund të ndjekë një linjë të qartë të asaj që është quajtur fillimisht ODS. Motivimet e ODS-ve fillestare ishin të ngjashme me magazinat moderne të të dhënave në kohë reale, por implementimi i DWH-së në kohë reale pasqyron një gjeneratë të re hadruerësh, softuerësh dhe teknikash. Pjesët e mëposhtme i zhvillojnë këto ide më me hollësi.

Gjenerata 1 – Operational Data Store

Është një ndërtim i DWH-së i gjeneratës së parë, i synuar për të mbështetur raportimin e kohëzgjatjes së ulët përmes krijimit të një konstrukti arkitekturor të qartë dhe aplikimi të veçantë nga DWH. ODS është sistem gjysmë operacional dhe gjysmë vendim mbështetës,

duke u përpjekur të krijojë një ekuilibër midis nevojës për të mbështetur njëkohësisht përditësimet e shpeshta dhe kërkimet e shpeshta.

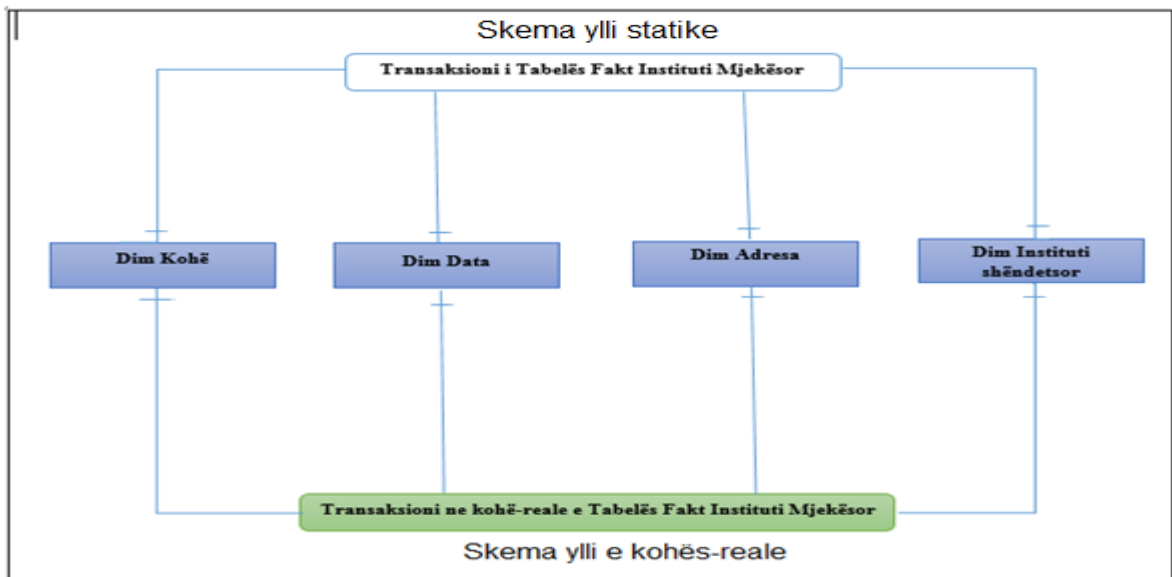
Arkitekturat më të hershme ODS-në e përshkruanin si një vend ku të dhënat integroheshin dhe dërgoheshin në një DWH “downstream”, duke vepruar kështu si një lloj zgjatimi i shtresës ETL të DWH-së. Arkitekturat e mëvonshme e përshkruajnë si bashkues të të dhënave të integruara nga shtresa ETL e DWH-së dhe e kategorizojnë si Tipi 1 deri 4 dhe ODS të brendshme ose të jashtme, në varësi të vendit ku qëndron brenda arkitekturës së përgjithshme dhe urgjencës me të cilën duhet t’i ngarkojë të dhënat nga fusha operationale.

Gjenerata 2 - Particionimi në kohë reale

Përdorimi i particionit logjik dhe fizik në kohë reale, siç është përshkruar fillimisht nga Ralph Kimball, është një zgjidhje pragmatike e disponueshme për të dhënat analitike në kohë reale nga një DWH.

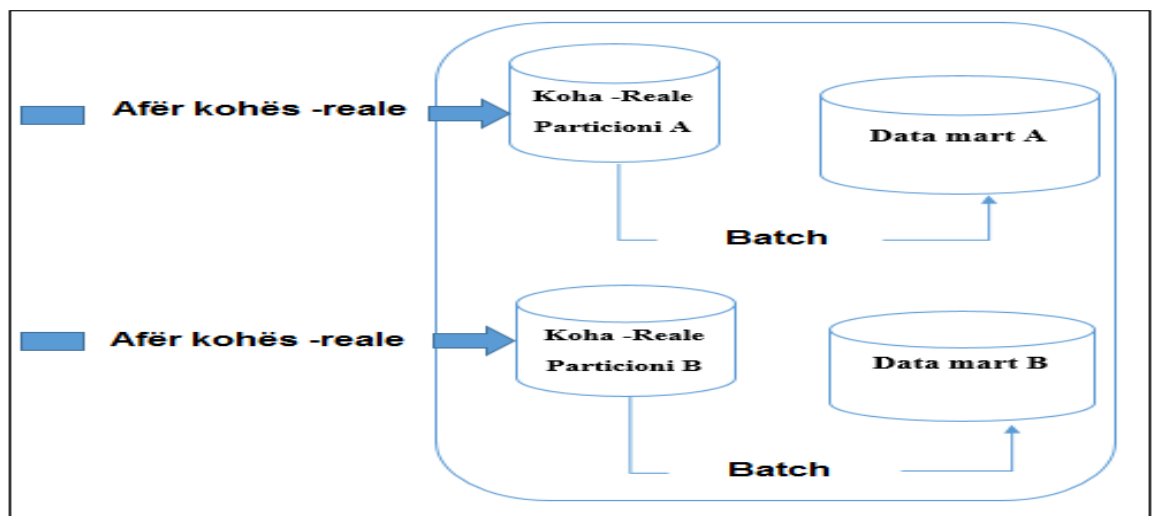
Duke përdorur këtë qasje krijohet një tabelë faktesh në kohë reale, struktura dhe dimensionaliteti i së cilës përshtatet me tabelën përkatëse të fakteve në DWH statike (ngarkesa e natës). Kjo tabelë faktesh në kohë reale përmban vetëm faktet e ditës aktuale (ato që nuk janë ngarkuar akoma në tabelën e DWH-së statike). Figura 5.19 tregon dy skema yll të lidhura me tabelat e fakteve në kohë reale dhe statike të pikës së institutit mjekësor, duke ndarë një set të përbashkët dimensionesh.

Figura 5.19 Lidhja midis formës statike dhe skemës yll në kohë reale (Kimball&Caserta, 2004).



Çdo natë përmbajtja e tabelës së partitionit në kohë reale shkruhet në tabelën e fakteve statike dhe partitioni në kohë reale pastaj bartet, gati për të marrë transaksionet e ditës tjetër. Figura 5.20 jep një ide se si punon procesi.

Figura 5 20 Lidhja logjike për partitionet e data mart-ëve në kohë-reale (Kimball&Caserta, 2004).



Në thelb, kjo qasje sjell dobitë e raportimit të të dhënave në kohë reale të ODS-së në vetë DWH-në, duke eliminuar shumë prej mbingarkesës arkitekturore ODS-në në proces.

Faktet rrjedhin në tabelën e fakteve në kohë reale përgjatë ditës dhe kërkimet e përdoruesit kundrejt tabelës në kohë reale ndalohen ose ndërpriten nga ky proces ngarkimi.

Indeksi në tabelën e fakteve në kohë reale është minimal ose nuk ekziston, për të minimizuar përpjekjet e ngarkimit të të dhënave dhe ndikimin e tij në kohët e përgjigjes së kërkimit. Performanca arrihet duke kufizuar sasinë e të dhënave në tabelë (vetëm një ditë) dhe duke futur të gjithë tabelën e fakteve në kohë reale në memorie.

Në mënyrë opsionale mund të krijohet një pamje që i kombinon faktet në tabelën e fakteve në kohë reale dhe statike (unione), duke siguruar një skemë virtuale yll për të thjeshtuar kërkimet që kërkojnë shikime të masave historike që shtrihen deri në moment.

Nëse vetëm rekordet e fakteve rrjedhin-futen në particionin në kohë reale, nevojitet një politikë për t'u marrë me ndryshimet në dimensionet që ndodhin midis dy ngarkesave të mëdha natën.

Për shembull, rekordet e institutit mjekësor të krijuar gjatë ditës, për të cilat keni fakte që mund të ketë nevojë të rikthehen në një seri rekordesh gjenerike të pacientit në dimensionin e institutit shëndetësor për t'u update-uat në rekordet e pacientit, më përshkruese në mbrëmje, kur një ngarkesë e plotë e të dhënave të rekordit të pacientit dhe të ndryshuara ngarkohet në dimensionin statik të institutit mjekësor.

Ndryshe, DWH në kohë reale mund të zgjedhë të marrë më shumë pamje të shpeshta të imazheve dimensionale që ndryshojnë për të braktisur konceptin pikë-në-kohë së bashku dhe për të kapur të gjitha ndryshimet dimensionale që ndodhin.

Më vonë, ky kapitull tregon disa prej çështjeve që lidhen me zgjedhjen e një politike të përshtatshme për t'u marrë me ndryshimet dimensionale, disa qasje pragmatike për të rrjedhur

të dhënat në particione në kohë reale përgjatë ditës së punës dhe avantazhet e disavantazhet e këtyre qasjeve.

5.9 Përcaktimi i ETL-së në kohë reale

Në shumë raste ofrimi i DWH-së në kohë reale kërkon një qasje krejt të ndryshme nga metodat ETL të përdorura në DWH me orientim në grup. Thjesht, përdorimi i grupeve ETL të zakonshme mbi një program më të shpeshtë përgjatë ditës mund të mos jetë praktik, qoftë për sistemet OLTP apo për DWH. Anasjelltas, përfshirë DWH-në në logjikën e kryerjes të transaksionit të sistemit OLTP nuk mund të punojnë, gjithashtu.

Sistemi OLTP nuk e ka luksin e të priturit në kryerjen e transaksionit të ngarkimit të DWH-së para se të vazhdojë me transaksionin tjetër, as nuk ka ndonjë kyçje ose logjikë kryerjeje me dy faza praktike nëpër sisteme me struktura të ndryshme dhe nivele të ndryshme.

Në vend të kësaj, u referohet, thjesht, të lëvizni në transaksione të reja në një particion në kohë reale të veçantë (e përkufizuar më tej në këtë kapitull) të DWH-së brenda një afati kohor të pranueshëm për menaxhimin e fatkeqësive natyrore, duke ofruar mbështetje analitike për vendimet e veprimtarive ditë për ditë. Për momentin, kjo procedurë është përkufizimi ynë praktik i ETL-së në kohë reale.

5.9.1 Qasjet ETL në kohë reale

Përmes njëfarë riciklimi kreativ të teknologjive dhe mjeteve ETL të themeluara, një paletë e maturuar dhe e gjerë teknologjish, është e disponueshme për të adresuar kërkesat e DWH-së. Pjesët në vijim diskutojnë këto teknologji. ETL-ja e zakonshme me qasje është jashtëzakonisht e efektshme në adresimin e kërkesave të raportimit në grup çdo ditë, javë dhe muaj. Transaksionet e reja ose të ndryshuara (rekordet e fakteve) lëvizin në masë dhe dimensionet kapen si pamje në një pikë kohore për çdo ngarkesë.

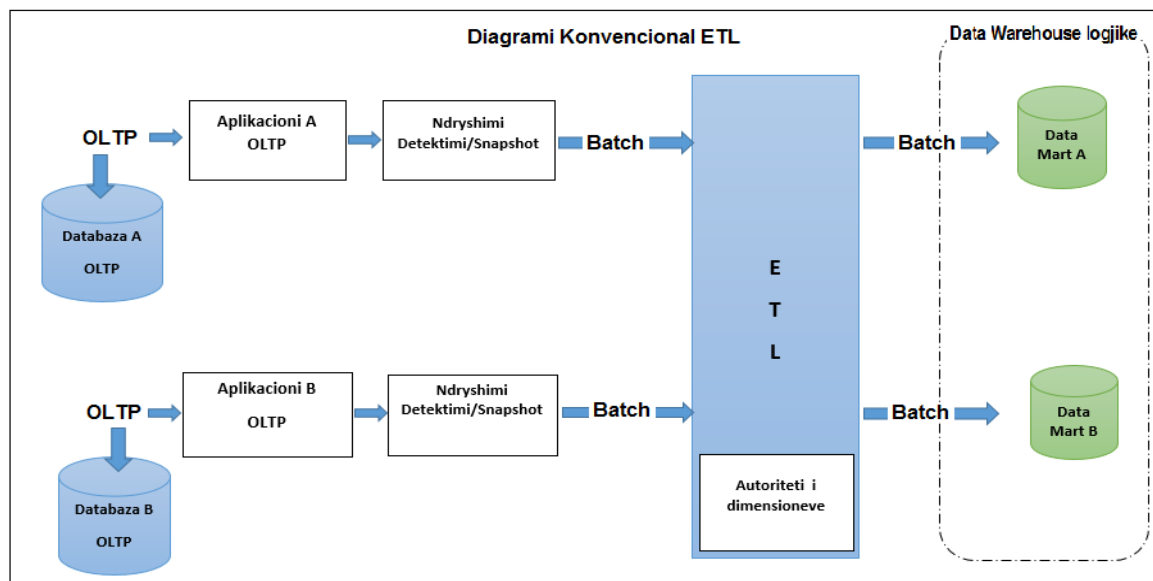
Prandaj, ndryshimet në dimensionet që ndodhin midis përpunimeve në grup nuk janë të disponueshme në DWH.

Andaj, ETL nuk është një teknikë e përshtatshme për integrimin e të dhënave ose aplikimit për AME-në, që kanë nevojë për raportim me kohëzgjatje të shkurtër ose për organizatat që kanë nevojë për kapje ndryshimi dimensional më të detajuar.

Por, ETL-ja e zakonshme është një metodë e thjeshtë, direkte dhe e provuar për organizatat, të cilat kanë kërkesa në kohëzgjatje më rastësore dhe sfida integrimi më komplekse.

Figura 5.21 tregon procesin ETL tradicional.

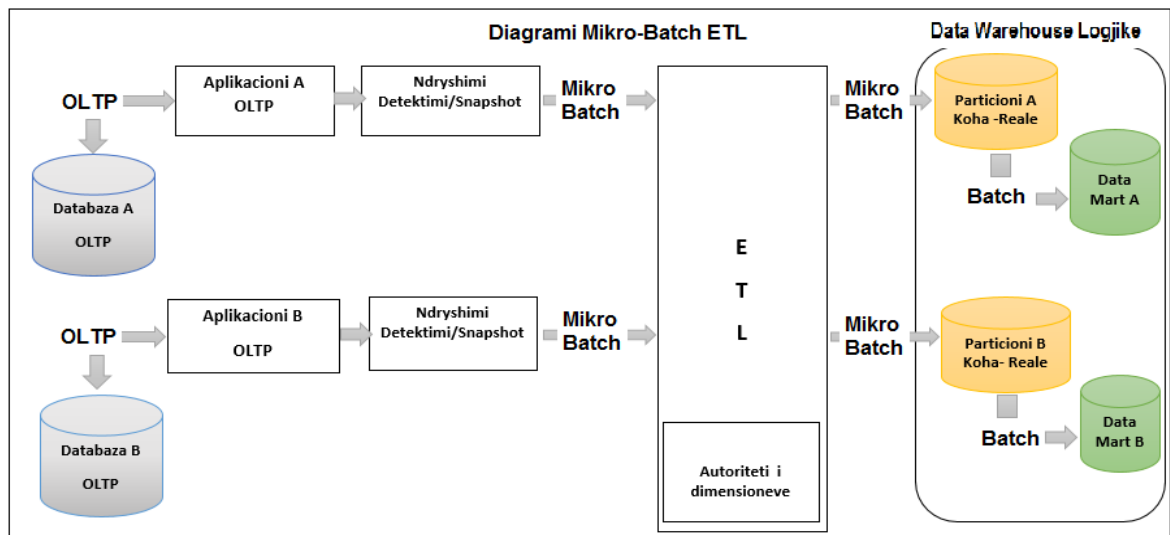
Figura 5.21 Diagrami Konvencional ETL (Kimball&Caserta, 2004).



ETL mikro-batch është shumë i ngjashëm me ETL-në tradicionale, përveç se shpeshtësia e grupeve rritet, ndoshta deri në shpeshtësinë prej çdo ore. Këta mikro-batch të shpeshtë vihen në punë përmes një procesi ETL-je tradicional tjetër dhe ushqejnë direkt particionet në kohë

reale të data mart-ëve. Njëherë në ditë partitionet në kohë reale kopjohen në data mart statikë dhe bashkohen. Figura 5.22 tregon një diagram të ETL-së mikro-batch.

Figura 5.22 Diagrami mikro-batch ETL (Kimball&Caserta, 2004).



Sistemet menaxhere të dimensioneve gjenerojnë imazhe dimensionale të reja në dimensionet në ndryshim Tipi 2 ose 3, por shkak i rritjes në shpeshtësinë e punimit, dimensionet që ndryshojnë përgjatë ditës mund të ndryshojnë shpejt dhe të shkojnë thellë.

Një alternativë është të injorohen ndryshimet në dimensionet që ndodhin gjatë ditës dhe në vend të kësaj të gjenerohen regjistrime dimensionesh vetëm për disa raste, duke përdorur vlerat standarde në të gjitha kolonat.

Ky kompromis mund të mjaftojë për organizatat që gjenerojnë pak regjistrime të reja mbi një ditë të caktuar dhe janë tolerante për kohëzgjatjen e kontekstit dimensional nga mbrëmja e mëparshme, por pakëson qartë disa prej dobive të raportimit në kohë reale.

Nëse është e pashmangshme, zgjidhja e vetme praktike për t'u marrë me dimensionet që ndryshojnë shpesh është përdorimi i matur i një minidimensioni, ku krijohen dimensione të veçanta për shumicën e attributeve që ndryshojnë shpesh një dimension të madh dhe duke

pakësuar kështu numrin e rasteve dimensionale të reja që kanë nevojë të krijohen nga procesi ETL.

ETL mikro-grup kërkon një kontroll pune gjithëpërfshirës, programim, vartësi dhe metodë për minimizim të gabimeve, një që është e fortë aq sa të punojë e pavarur për shumicën e kohës dhe në gjendje të ekzekutojë strategjitë e publikimit në DWH përballë shumicës së çështjeve të ngarkimit të të dhënave.

5.10 Enterprise Application Integration (EAI)

Në fundin e sipërm të spektrit të kompleksitetit shtrihet Enterprise Application Integration (EAI), të quajtur ndonjëherë integrim funksional. EAI përshkruan setin e teknologjive që mbështetin integrimin e vërtetë të aplikimit, duke lejuar sistemet operacionale individuale të ndërveprojnë në mënyra të reja dhe të ndryshme nga ato që u caktuan fillimisht.

EAI, zakonisht, detyron ndërtimin e një seti komponentësh adaptorë dhe brokerë që lëvizin transaksionet e menaxhimit të fatkeqësive natyrore, në formën e mesazheve, nëpër sisteme të ndryshme në rrjetin e integrit, duke i izoluar të gjitha sistemet nga njohuria ose vartësitë në sistemet e tjera në rrjetin e integrit.

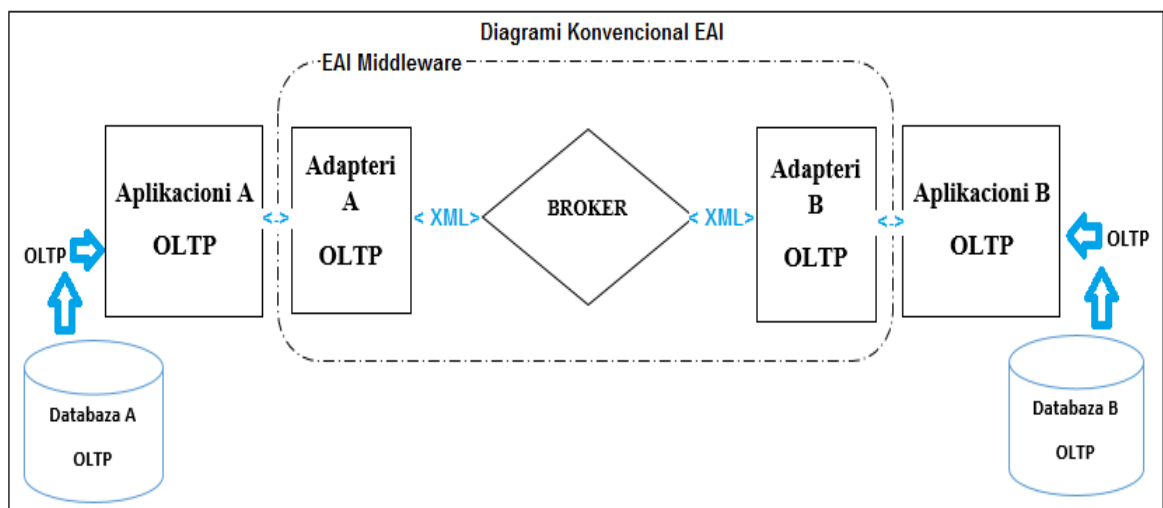
Adaptorët specifikë të aplikimit janë përgjegjës për t'u përballur në gjithë logjikën që duhet për të krijuar dhe ekzekutuar mesazhe dhe brokerët janë përgjegjës për përcjelljen e mesazheve në mënyrë të përshtatshme, bazuar në rregullat e publikimit dhe regjistrimit. Adaptorët dhe brokerët komunikojnë përmes mesazheve që varen nga aplikacioni, shpesh të pasqyruara në XML.

Kur ndodh një eveniment i konsiderueshëm aplikimi, si updat-imi i një rekordi të pacientëve, lëshohet një shkak dhe adaptor i aplikacionit krijon një mesazh të ri.

Adaptori është, po ashtu, përgjegjës për fillimin e transaksioneve përkatëse kur merr një mesazh që përmban informacione që ka zgjedhur të marrë, për shembull, një rekord i pacientit të ri nga sistemi menaxheri i dimensionit instituti mjekësor.

Mesazhet rrugë-broker midis adaptorëve, bazuar në një set publikimesh dhe rregullat e regjistrimit. Radhët e mesazheve shpesh vendosen midis aplikacioneve dhe adaptorëve të tyre dhe midis adaptorëve dhe brokerëve, për të dhënë një fushë stazhimi për mesazhet asinkrone dhe për të mbështetur garancitë e dërgimit dhe qëndrueshmërinë e transaksioneve nëpër rrjetin e integritetit. Në figurën 5.23 aplikacionet A dhe B punojnë në mënyrë të pavarur, por janë në gjendje të shkëmbejnë të dhëna dhe të ndërveprojnë përmes mesazheve EAI.

Figura 5.23 Diagrami Konvencional EAI (Kimball&Caserta, 2004).



Për shembull, ndryshimet në regjistrimet e një pacienti në aplikacionin A mund të lëshojnë një shkak nga adaptori i aplikacionit A, i cili krijon dhe dërgon një mesazh XML të ndryshimit të një brokeri.

Nëse aplikacioni B është regjistruar në mesazhet pacient-ndryshim nga aplikacioni A, brokeri përcjell mesazhin në adaptorin e aplikacionit B, i cili mund të zbatojë, pastaj, të gjitha ose një nënsset të ndryshimit të regjistrimit të pacientit në aplikacionin B.

Aplikacionet A dhe B nuk kanë nevojë të dinë asgjë për njëri-tjetrin, adaptorët e tyre përkatës janë përgjegjës për kapjen, interpretimin dhe aplikimin e mesazheve për aplikacionin e tyre.

Ky koncept është i fuqishëm, sepse lejon rrjetet EAI të shtrihen në mënyrë solide, prezantimi i aplikacioneve të reja në rrjetin e integritetit kërkon vetëm krijimin e një adaptori të ri dhe rregullave të reja të publikimit/regjistrimit të brokerit.

Çdo adaptor duhet të jetë në gjendje të ekzekutojë transaksione komplekse në sistemin e tij host dhe të merret me shumë probleme konkurrence që mund të shfaqen kurdoherë që aplikacionet e pavarura operojnë mbi të dhëna logjike të përbashkëta.

Pavarësisht nga qasja e integritetit, duhet të merren në shqyrtim disa çështje të caktuara diku në arkitekturë dhe adaptorët EAI e bëjnë këtë në pozicionin optimal, sa më pranë aplikacionit që të jetë e mundur.

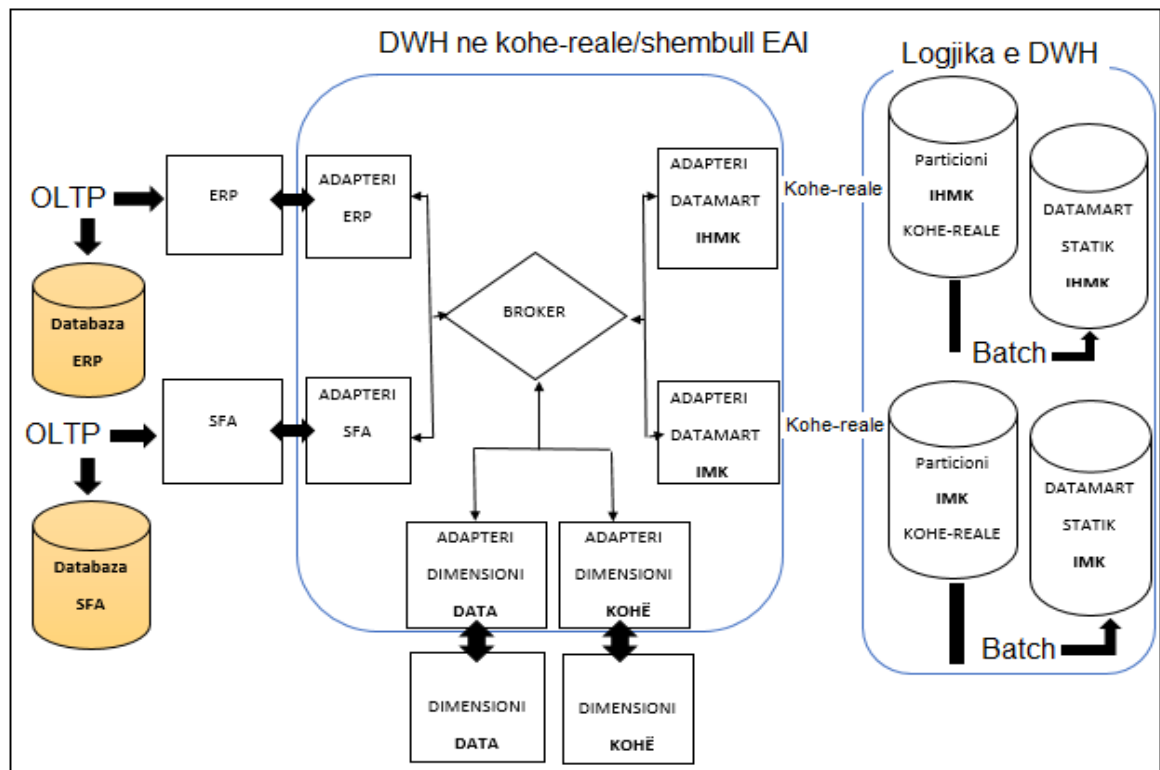
Teknologjitë EAI mund të jenë mjete të fuqishme për DWH-në në kohë reale, sepse ato mbështetin aftësinë për të sinkronizuar të dhëna të rëndësishme, si informacionet e pacientit nëpër aplikacione dhe ato ofrojnë një mjet të efektshëm për shpërndarjen e aseteve të informacioneve të nxjerra nga DWH, si, për shembull, vlerat e segmentimit të pacientit nëpër sipërmarrje.

Arkitektura e DWH EAI në kohë reale modularizon bllokun ETL monolitik, duke tërhequr sistemin a sistemet menaxhere dimensionale si komponentë arkitekturore të veçantë, secili me adaptorët e vet dhe duke e vënë përgjegjësinë për shumicën e bërthamave të transformimit

dhe ngarkimit të particioneve në kohë reale të data mart-ëve në shumicën e adaptorëve të data mart-ëve.

Figura 5.24 paraqet një diagram të DWH-së në EAI në kohë reale.

Figura 5.24 Diagrami i DWH-së në kohë reale i EAI-së



Një skenar tipik real mund të përfshijë implementimin e adaptorëve për një set sistemesh OLTP, si, për shembull, një planifikim i burimit të sipërmarrjeve, ERP dhe SFA, sistemet menaxhere të dimensionit KOHË dhe DATA (të cilat kryejnë pastrim dhe deduplikim në kohë reale) dhe i data mart-ëve për IMK dhe IHMK.

Çdo transaksion ndryshimi KOHË dhe DATA do të kapej nga aplikacioni OLTP nga një adaptor, që është mesazh dimensionit i pakonfirmuar te brokeri, i cili pastaj e dërgon në

cilindo sistem që regjistrohet në mesazhet e dimensionit të pakonfirmuar, zakonisht vetëm sistemi menaxher i dimensionit të përshtatshëm.

Sistemi menaxher i dimensionit konform regjistrimit dimensional dhe adaptorit të tij, e dërgon përsëri si mesazh dimensionit të konfirmuar te brokeri, i cili pastaj e përcjellë te të gjitha sistemet që regjistrohen te të dhënat e dimensioneve të konfirmuara, zakonisht sistemet OLTP dhe data mart-ët.

Marrë në shqyrtim këtë shembull: Sistemi ERP update-on rekordet e DATA, pastaj adaptorit ERP zbulon këtë ndryshim, gjeneron një mesazh XML të emërtuar Transaksion DATA i pakonfirmuar nga ERP dhe e dërgon te brokeri.

Më pas brokeri e përcjellë këtë mesazh te menaxheri i dimensionit DATA, zakonisht sistemi i vetëm që regjistrohet në mesazhet DATA të pakonfirmuar nga sistemi ERP.

Menaxheri i dimensionit DATA merr mesazhin dhe e vendosë informacionin e pakonfirmuar të DATA në radhën e punës (ose fazën e stazhimit në mikro-grup) të menaxherit dimension DATA. Menaxheri dimension DATA punon transaksionin dhe nëse rezulton në një ndryshim në një ose më shumë rekorde DATA të konfirmuara, adaptorit menaxher i dimensionit DATA zbulon këto ndryshime që kanë ndodhur, i paketon këto rekorde paciente të rishikuara në mesazhet Transaksione DATA të konfirmuara nga mesazhet Menaxher Dimension DATA dhe i dërgon ato te brokeri.

Duke supozuar se data mart-ët e porosive, data mart-ët e IMK-së dhe sistemet SFA janë regjistruar të gjitha në Transaksionet DATA, të konfirmuara nga mesazhet Dimensionit DATA Menaxher, brokeri i kopjon dhe shpërndan mesazhin te të katër këto sisteme.

Secili prej katër adaptorëve është përgjegjës për aplikimin e ndryshimeve në rekordet DATA në aplikimet e tyre përkatëse.

Pranimi i regjistrimit të konfirmuar të DATA nga sistemet ERP dhe SFA mund të shkaktojë një ndryshim në regjistrimet e tyre përkatëse të DATA, duke lëshuar kështu një set të ri transaksionesh EAI.

Transaksionet e fakteve kapen, gjithashtu, nga adaptorët OLTP, i dërgohen brokerit si mesazh faktesh IMK ose IHMK dhe, pastaj, u dërgohen të gjithë regjistruarve të këtyre llojeve të mesazheve, tipikisht data mart-ëve.

Adaptori data mart kryen të gjitha transformimet e nevojshme dhe e futë transaksionin e ri direkt në data mart-in particion në kohë reale. EAI është një mjet i fuqishëm i sinkronizimit të informacioneve kyçe të AME-se, të dy për rrjedhjen-furnizimin e data mart-ëve dhe për publikimin e shpërndarjen e segmentimeve të nxjerra nga magazina në sistemet OLTP përballë pacientit. Por, mund të jetë komplekse dhe e shtrenjtë për t'u implementuar. EAI është një qasje e shkëlqyer për organizatat, kërkesat e të cilave kërkojnë kohëzgjatje raportimi të ulët, të cilët janë intolerante nga humbja e update-imeve dimensionale ndërditore ose që kërkojnë sinkronizim bidireksional të të dhënave dimensionale midis DWH-së dhe sistemeve operative.

5.11 Ndërtimi i particioneve në kohë reale

Në vitet e kaluara një kërkesë e madhe e re është shtuar në listën e ndërtuesve të DWH-së. Datawarehouse tani duhet ta zgjasë kohën e saj historike ekzistuese deri në rastin aktual. Për më tepër, ne duhet të gjurmojmë statusin për çdo orë të porosisë më të fundit pasi ndryshon gjatë ditës.

Megjithëse boshllëku midis sistemeve operacionale të përpunimit të transaksioneve dhe DWH-së, në shumicën e rasteve është ngushtuar në 24 orë, nevojat sistematike që vijnë nga menaxhimi i fatkeqësive natyrore kërkojnë që DWH-ja ta mbushë këtë boshllëk me të dhëna thuajse në kohë reale.

Shumica e ndërtuesve të DWH-së janë skeptikë se punët ekzistuese ekstrakt-transformim-ngarkesë (ETL), thjesht, mund të përshpejtohen nga një kohë cikli 24-orësh në një kohë cikli 15-minutash.

Edhe nëse hapat e pastrimit të të dhënave lidhen që të ndodhin paralel me ngarkimin final të të dhënave, manipulimet fizike rrotull fakteve më të mëdha dhe tabelave të dimensioneve, thjesht, nuk mund të bëhen çdo 15 minuta. Ndërtuesit e DWH-së po i përgjigjen kësaj nevoje, duke ndërtuar një particion në kohë reale përpara DWH-së së zakonshme statike të të dhënave (Rainardi, 2008).

5.11.1 Roli strategjik i menaxhimit të dimensioneve

Mundësia që lidhë në mënyrë logjike ose fizike fusha subjektsh të veçanta (data mart-ët) së bashku në arkitekturën bus të DWH-së dimensionale, është konformancë e dimensioneve dhe fakteve, e arritur përmes përdorimit të sistemeve menaxhere dimensionit siç përshkruhet më tej. Tradicionalisht, menaxheri i dimensionit është parë si një rol, puna e të cilit është përkufizimi, mirëmbajtja dhe publikimi i një dimensionit të veçantë të konfirmuar i të gjitha data mart-ëve që ndërveprojnë brenda arkitekturës bus të DWH-së.

Së fundi, DWH në kohë reale luan një rol në objektivin më të madh të sigurimit të aksesit të gatshëm për shumicën e të dhënave aktuale dhe informative për të gjithë përdoruesit në sipërmarrje. Për më tepër, për të dhënë shpejt regjistrime faktesh në DWH, mund të gjendet

avantazh tejet konkurrues në sigurimin e sinkronizimit në kohë reale të dimensioneve kyçe, si Dim_kohe ose Dim_data nëpër të gjitha sistemet operationale në organizatë.

Ky funksion informacion-sinkronizim mund të konsiderohet një zgjatim logjik i rolit të menaxherit dimension dhe është një mekanizëm i efektshëm dhe i qëndrueshëm për mbylljen e lakut midis botës operationale dhe asaj të datawarehouse (DWH), duke siguruar një mjet të shpërndarjes së segmentimeve të nxjerra nga DWH-ja dhe të informacioneve të tjera të pasurimit në botën operationale.

Menaxheri i dimensionit Dim_kohe në një DWH strategjike në kohë reale mund jo vetëm të rrjedhë data mart me informacione Dim_kohe të konfirmuar, por gjithashtu mund të bashkëpunojë me ndonjë mekanizëm për sinkronizimin e informacioneve të Dim_kohe nëpër të gjitha sistemet operationale të interesuara (regjistruara). Këto informacione Dim_kohe në kohë reale duhet të përfshijnë inteligjencë Dim_kohe të gjeneruar nga vetë DWH.

Qartësisht, këto objektiva ambicioze, dhe sipas këtij shkrimi, asnjë zgjidhje e paketuar apo set mjetesh nuk mund të thjeshtojë në mënyrë dramatike procesin e ndërtimit të një integrimi aplikacioni ndërmarrjeje bi-direksional (EAI)/zgjidhje DWH në kohë reale. Sidoqoftë, këto sisteme janë krijuar, blloqet bazë të ndërtimit për këto sisteme ekzistojnë dhe po maturohen më shumë. Potenciali për diferencimin e biznesit, i siguruar nga një sistem i tillë është mahnitës, prandaj ka të ngjarë që adoptuesit e hershëm të sotëm do të gëzojnë vend në treg. Avantazhet që nxitin adoptimin më të gjerë të këtyre sistemeve në të ardhmen. Marr parasysh sistemet e ndërtuara sot aspak nuk e pengojnë aftësinë e organizatës për t'u evoluar në një EAI në kohë reale/zgjidhje DWH në të ardhmen.

5.11.2 Vetëm faktet apo edhe ndryshimet në dimension

Arkitektët e DWH-së dimensionale e përshkruajnë këtë fushë në terma faktesh e dimensionesh, por sistemet OLTP nuk i bëjnë këto dallime. Sidoqoftë, duhet të kuptohet se kategorizimi i transaksioneve të të dhënave nga AME-ja OLTP, me interes për përdoruesit finalë, duhet të ndërtohen në mënyrë të përshtatshme.

A janë të përqendruara kërkesat e raportimit në kohë reale të fokusuara vetëm në fakte të freskëta, si, për shembull, kërkesat e bëra nga terreni, komunikimet e fundit, raportimet e fundit të agjencive, e kështu me radhë, apo lidhen, gjithashtu, me transaksionet e freskëta të dimensioneve? Nëse ndryshimet dimensionale në kohë reale janë të nevojshme për raportim, apo ndryshojnë ato ngadalë apo shpejt?

Me fjalë të tjera, a ka nevojë AME-ja për një imazh të saktë të këtyre dimensioneve, ashtu siç ishin në momentin e transaksionit, apo mundet që të gjitha ose disa update-ime dimensionale të shkruhen në formë shkatërruese kur ndodhin përditësime të reja?

A duhet që raportet të jenë të përsëritshme? Tipi 1 - dimensionet që ndryshojnë ngadalë, në të cilat ndryshimet në atributet e një dimensionit mbishkruajnë në formë shkatërruese vlerat e mëparshme, rezultojnë në një DWH që rindërton vazhdimisht historinë, duke raportuar evenimentet jo teksa shikojnë në kohën e transaksionit, por teksa shikojnë në kontekstin e dimensioneve të sotme. Në këtë skenar, raportet e vëna kundrejt të njëjtëve elementë dimensional në pika të ndryshme kohore kanë rezultate paksa të ndryshme ose shumë të ndryshme.

Tipi 1 - ndryshimet janë, gjithashtu, të rrezikshme, sepse mund të zhvlerësojnë përmbledhje historike nëse ndryshimi i Tipit 1 zbatohet në një fushë të përdorur si bazë të një llogaritjeje përmbledhëse.

Tipi 2 dhe 3 - dimensionet që ndryshojnë ngadalë ruajnë një tablo më granulare të imazheve të dimensionit në pika të caktuara në kohë, ndoshta çdo ditë, por përsëri nuk i kapin ndryshimet në dimensione që ndodhin midis ekstraktimeve.

Rifreskimi dimensional në kohë reale mund ta lëvizë këtë granularitet deri në çdo disa minuta ose mund t'i kapë të gjitha ndryshimet e dimensionit. Implikimet arkitekturore nuk janë të lehta.

Duke adoptuar një politikë të kapjes së imazheve të ndryshimit dimensional gjithnjë e më shpesh, DWH largohet nga qëndrimi i saj i mëparshëm si një sistem të matjes periodike (imazheve) dhe drejt një ideali me zero-kohëzgjatje mbështetjes së vendimit.

Ndërsa, DWH dhe teknologjitë e integritetit të aplikacionit fillojnë të përzihen, DWH bëhet, në efekt, një zgjatim i vërtetë logjik i sistemeve operative që vë në punë sipërmarrjen.

Për momentin, si çështje praktike, sistemet ETL me shumë mundësi kanë nevojë të ndërtohen që të japin thujtë zero-kohëzgjatje për faktet e matura sipas mundësisë, por të lejojnë që disa ose të gjitha atributet dimensionale të përditësohen në grupe, ose mikrogrupe, siç është zhvilluar në këtë kapitull.

5.11.3 Integrimi i të dhënave dhe integrimi i aplikacioneve

Duke supozuar që kërkesa e DWH-së në kohë reale detyron, gjithashtu, njëfarë mase të integritit nëpër sistemet operacionale, duhet ta kategorizojme si integrim të dhënash apo integrim aplikacioni.

Në përgjithësi, integrimi që mund të plotësohet, thjesht, duke lëvizur të dhënat midis databazave quhet integrim të dhënash. Shpesh këto zgjidhje janë point-to-point, të ekzekutuara përmes ekstraktimit të fajllave ASCII (për databazat heterogjene), lidhjeve të databazave dhe portave (gateways) ose për shërbimet e replikimit të databazave homogjene ose imazheve të tabelave. Në thelb, të dhënat ndahen nëpër pjesët e pasme të aplikacioneve pjesëmarrëse, duke tejkaluar plotësisht logjikën e aplikacionit. Disa mjete integrimi të dhënash më të sofistikuar sigurojnë mbështetje administrimi të centralizuar për programimin dhe lëvizjen e të dhënave, duke mbështetur pak më shumë kontroll sipërmarrjeje dhe menaxhim për bërthamat e integritit të të dhënave point-to-point.

Integrimi i aplikacionit (ndonjëherë i quajtur integrim funksional) mund të përshkruhet si ndërtim i zgjidhjeve të reja të biznesit, duke ngjitur aplikacionet së bashku përmes përdorimit të disa 'middleware' të përbashkët. Middleware është një kategori softueri që varet nga aplikacioni, duke siguruar një mjet për të kapur, përcjellë dhe ekzekutuar mesazhet transaksionale midis aplikacioneve. Në përgjithësi, përdoren konektorët ose adaptorët për të lidhur aplikacionet pjesëmarrëse në rrjetin e integritit dhe komisionerët përdoren për t'i përcjellë mesazhet sipas rregullave të publikimit dhe regjistrimit.

Point-to-point kundrejt hub-and-spoke

Nëse datawarehouse mbështet, po ashtu, njëfarë shkalle të integritit të aplikacionit (ose funksional), një faktor i rëndësishëm në zgjedhjen e arkitekturës është numri i sistemeve të

publikimit dhe regjistrimit, duke mbështetur të ardhmen e parashikueshme. Ky numër mund t'u ndihmojë të vendosni nëse një zgjidhje point-to-point relativisht e thjeshtë do t'u mjaftojë apo nëse kërkohet një arkitekturë më e fortë hub-and-spoke.

Figura 5.25 Integrimi i aplikacioneve point-to-point (Kimball&Caserta, 2004).

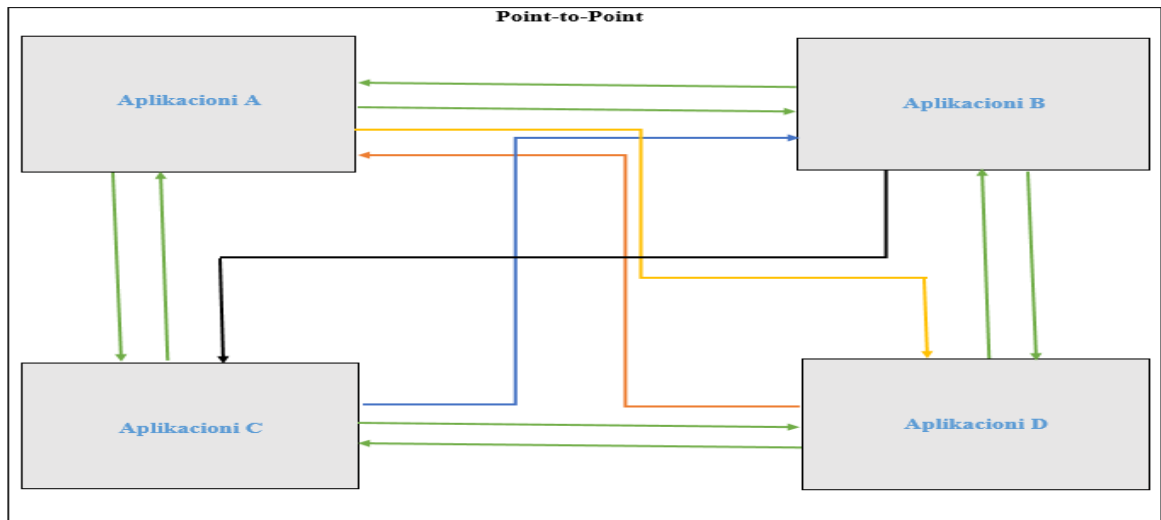
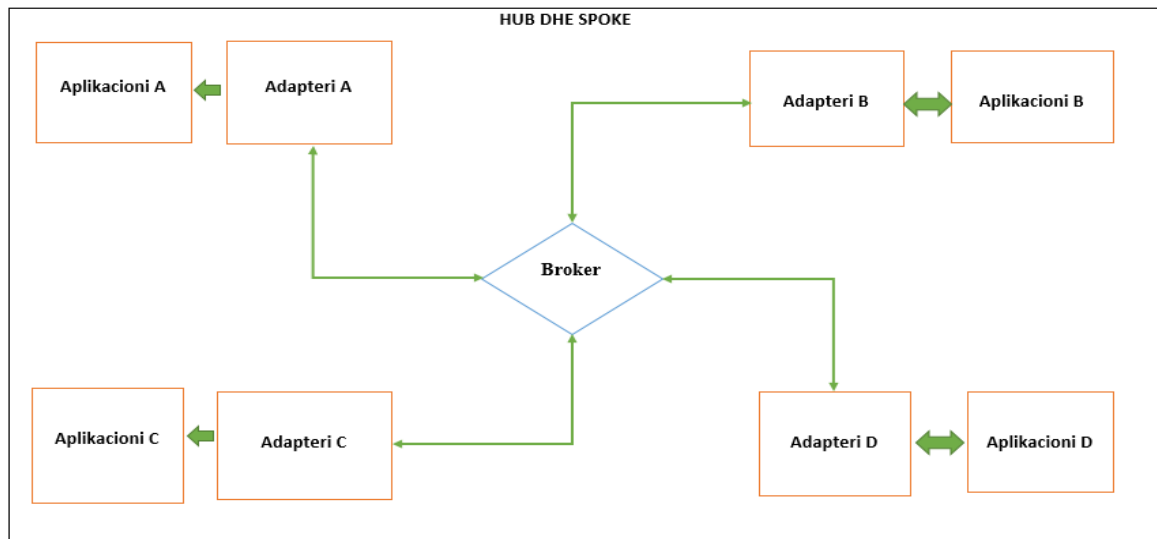


Figura 5.25 tregon se edhe me një numër aplikacionesh relativisht të vogla që shkëmbejnë të dhënat, zgjidhjet point-to-point mund të kërkojnë një numër shumë të madh ndërfaqesh të shkëmbimit të të dhënave, secila prej të cilave kërkon mirëmbajtje kurdoherë që burimi ose aplikacionet e synuara ndryshojnë.

Shtimi i aplikacioneve në rrjetin e integritetit kërkon, gjithashtu, ndërfaqe të reja të shkëmbimit të të dhënave për të gjitha aplikacionet e publikimit dhe regjistrimit. Sidoqoftë, organizatat që kanë një listë të shkurtër aplikacionesh që kërkojnë integrim dimensionali të konfirmuara dhe që presin që kjo listë të mbetet e qëndrueshme për të ardhmen e parashikueshme, mund të gjejnë se qasja e integritetit point-to-point është mjaft tërheqëse. Shmang kompleksitetin e krijimit të komponentëve middleware, EAI mund të mbështetet

pjesërisht përmes përdorimit të teknologjive të integrit të të dhënave, si Capture, Transform dhe Flow (CTF). Në kontrast me arkitekturat point-to-point, numri i ndërfaqeve klient dhe vartësive ndërsistem mund të minimizohet përmes përdorimit të një qasjeje integrimi hub-and-spoke, më pas kemi edhe figurën 5.26

Figura 5.26 Integrimi i Aplikacioneve HUB dhe SPOKE (Kimball&Caserta, 2004).



Sidoqoftë, barra shtesë e ndërtimit të adaptorëve EAI dhe komponentëve broker nuk është e parëndësishme. Çdo aplikacion që merr pjesë në rrjetin e integrit ka nevojë për një adaptor, i cili do të jetë në gjendje të konvertojë transaksionet e përcaktuara në mesazhe gjenerike dhe të interpretojë e ekzekutojë mesazhet gjenerike në aplikacionin lokal.

Mirëmbajtja e adaptorit kërkohet kurdoherë që një aplikacion host i lidhur ndryshon ose nëse seti i mesazheve gjenerike ndryshon. Rregullat e vështira dhe të shpejta mbi kufirin e vendimit midis arkitekturave point-to-point dhe hub-and-spoke nuk ekzistojnë, por organizatat që presin integrimin nëpër tre ose më shumë aplikacione ose ato që presin një

numër në rritje të pjesëmarrësve në integrim-rrjeti në të ardhmen e parashikueshme, mund të marrin në konsideratë arkitekturat hub-and-spoke EAI (Kimball&Caserta, 2004).

5.12 Konkluzion

Me përfundimin e ndërtimit të modelit data mart, duhet të zhvillohen edhe të gjitha proceset që do të bëjnë të mundur grumbullimin e informacionet nga burimet origjinale, transformimin dhe ngarkimin e tij në destinacion. Të tri këto procese llogariten të jenë proceset më komplekse gjatë ndërtimit të sistemeve ETL. Sipas arkitekturës specifike, destinacioni i të dhënave do të jetë një data mart i ndërtuar për një proces të menaxhimit të emergjencave dhe të fatkeqësive natyrore. Në përgjithësi, synimi i fazës së ekstraktimit është t'i konvertojë të dhënat në një format të vetëm, i cili është i përshtatshëm për përpunim transformimi.

Faza e transformimit aplikon një sërë rregullash ose funksionesh në të dhënat e ekstraktuara nga burimi për t'i nxjerrë të dhënat për ngarkim në vendin final. Për ETL është përdorur mjeti Talend Open Studio for Big Data. Synimi i këtij kapitulli është të dizajnojë dhe të implementojë procese ETL me anë të integritetit të të dhënave të databazave heterogjene që përdorin Talend Open Studio (TOS) dhe, njëkohësisht, të gjejë zgjidhje për sinkronizimin e të dhënave në kohë reale.

Përdorimi i Talend Open Studio (TOS) për dizajnimin dhe implementimin gjysmë-autonom të detyrave ETL në Java mundëson një mënyrë të shpejtë të adaptimit me burime të reja të dhënash. Qëllimi është të tregojë se Talend Open Studio for Big Data (TOS) është në gjendje t'i bëjë të njëjtat detyra si çdo vegël komerciale ETL.

Gjatë procesit të ngarkimit të tabelës Fakt IMK shfrytëzohen referencat e krijuara. Ndërkohë, në Oracle, një tabelë e njohur si FAKT_IMK, krijohet për t'i magazinuar të dhënat e Institutit Mjekësor të Kosovës. Të gjitha tabelat e referencës që nevojiten duhet të ruhen në memorie, që të mundë të kapen shumë lehtë sa herë procesohet një rekord i tabelës Fakt me informacionin e çelësit natyral të dimensionit. Gjatë procesit të ngarkimit mund edhe të bëjmë shtimin edhe të shumë elementeve të informacioneve.

Në ditët e sotme agjencitë e menaxhimit të emergjencave gjenerojnë sasi të mëdha të dhënash, këto fakte tregojnë se është shumë vështirë të përditësohen në mënyrë të efektshme DWH në kohë reale. DWH-ja në kohë reale synon uljen e kohës në komunikimet e rasteve të fatkeqësive natyrore dhe emergjencave. DWH në kohë reale siguron akses në një pamje të saktë, të integruar, të konsoliduar të informacioneve të AME-së dhe ndihmon në dhënien e informacioneve thuajse në kohë reale për përdoruesit e vet. Kjo kërkon teknika ETL të efektshme, që bëjnë të mundur integrimin e vazhdueshëm të të dhënave, po ashtu, edhe inkorporimi i CDC-së, sinkronizimi i të dhënave mund të bëhet i efektshëm në kohë reale.

Një DWH grumbullon të dhënat nga shumë sisteme me burim operacional heterogjen (OLTP - On-Line Transaction Processing) dhe ruan të dhënat e përmbledhura të integruara të AME-së në një depo qendrore të përdorur për aplikimet analitike (OLAP - On-Line Analytical Processing).

Një DWH, disa vite më parë, zakonisht, përditësohej çdo ditë ose çdo javë. Në dy deri tre vjetët e kaluar ka pasur gjithnjë e më shumë kërkesa për të rritur shpeshtësinë e update-imeve. Përdoruesit duan që të dhënat në DWH të update-ohen çdo dy minutë ose edhe në kohë reale.

Për t'u përballur me kërkesat në kohë reale, DWH duhet të jetë në gjendje të mundësojë integrimin e vazhdueshëm të të dhënave, me qëllim që të merret me të dhënat më të fundit të Agjencisë së menaxhimit të emergjencave se problemet e sinkronizimit të të dhënave lindin kur shënimet janë të përbëra nga të dhëna të nxjerra nga disa burime të dhënash që po përditësohen pa dallim. Shumica e ndërtuesve të DWH-së janë skeptikë se punët ekzistuese ekstrakt-transformim-ngarkesë (ETL), thjesht, mund të përshpejtohen nga një kohë cikli 24-orësh në një kohë cikli 15-minutash.

KAPITULLI VI :APLIKIMET E PAJISJEVE MOBILE NË RASTIN E FATKEQËSIVE NATYRORE

6.1 Hyrje

Në vitin 2002 fatkeqësitë natyrore kanë prekur 680 milionë njerëz anembanë botës. Fatkeqësitë natyrore përfshijnë tërmete, përmytje, thatësi, valë të nxehtit, të ftohtë të madh dhe ngjarje që shkatërrojnë vende, biznese dhe pronë private, si dhe mirëqenien e njerëzve. Për shkak të dëmeve dhe humbjeve të mëdha ekonomike që shkaktohen, shumë studime bien dakord për faktin se duhet parapërgatitje me qëllim të minimizimit të dëmeve. Rrjedhimisht, parapërgatitja për këto fatkeqësi natyrore ka qenë objekt i studimeve në rritje. Parapërgatitja për emergjenca dhe reagimi ndaj tyre janë një detyrë komplekse që varet nga strukturat organizative dhe ato ndërorganizative. Mëkëmbja pas një fatkeqësie natyrore lidhet me menaxhimin e saj, që varet nga efikasiteti i parapërgatitjes. Shumë studime propozojnë edhe faktorët që duhen marrë në konsideratë për të minimizuar rreziqet dhe që t'u hapin rrugë aksioneve për rimëkëmbje. Megjithatë, adoptimi i teknologjisë si faktor nuk ekziston ende, madje celularët si mjet parapërgatitjeje janë lënë anash.

Ky përjashtim është befasues, sepse literatura në lidhje me parapërgatitjen na çon drejt rezultatit se pajisjet mobile mund të jenë të dobishme. Sidoqoftë, viteve të fundit janë bërë disa studime në lidhje me përdorimin e pajisjeve mobile në këto situata nga vende në zhvillim dhe vende të zhvilluara. Kjo ka ndodhur ngase një numër në rritje njerëzish përdorin pajisjet mobile çdo ditë. Megjithatë, këto studime nuk i kanë marrë në konsideratë pajisjet mobile për t'i përdorur gjatë situatave të fatkeqësive natyrore. Ky studim është i dedikuar për të mbushur këtë boshllëk.

Synimi i këtij studimi është të ekzaminojë përdorimin e pajisjeve mobile në përpjekjet për përgatitje gjatë sezoneve të fatkeqësive natyrore në një vend në zhvillim. Duke qenë se Kosova është një vend i prekur nga fatkeqësitë natyrore, studimi i tanishëm synon të investigojë adoptimin e pajisjeve mobile për përgatitje për fatkeqësi natyrore. (Assion & Lori, 2012)

Teknologjia mobile ka treguar të ketë një impakt të konsiderueshëm në shoqëri në vitet e fundit. Gradualisht, kjo teknologji është bërë pjesë e jetës së përditshme dhe mënyra më e lehtë për të shpërndarë informacion. Zhvilluesit e aplikacioneve nuk lodhen së eksperimentuari dhe eksploruari aplikacione të reja, kështu që zhvilluesit kanë krijuar aplikacione të reja dhe sfiduese. Pajisje mobile të ndryshme përdorin platforma të ndryshme. (Tanya, 2014)

6.2 Platforma Mobile

Janë katër sisteme më të zakonshme operative në pajisje mobile. Sipas pjesës që zënë në treg gjatë vitit 2013, janë:

- Android nga Google, me dominim të dukshëm 81%,
- iOS, në iPhone e Apple, me 12.9%,
- Windows Phone nga Microsoft me 3.6%
- BlackBerry me 1.7%.

Mes këtyre sistemeve operative ka shumë dallime. Disa janë të dukshme edhe për përdoruesit, si, për shembull, prania e tipareve specifike të ndërfaqes së përdoruesit (UI), e disa janë të padukshëm për përdoruesit, por prekin zhvilluesit e aplikacioneve të pajisjeve mobile. Kur krahasohen platformat më të zakonshme, nga perspektiva e arsimimit, Goodrich and Rogers (2011), thonë se të dy janë të vlefshme për të krijuar kurse, secila me përparësitë

dhe disavantazhet e veta. Një avantazh i Android-it, krahasuar me iOS, është se zhvillimi i aplikacionit ishte i mundur në shumicën e kompjuterëve, kurse zhvillimi i iOS kërkonte përdorimin e kompjuterëve të Apple.

6.3 Kërkesat minimale për funksionim të platformës Android

Asnjë SDK nuk ka kërkesa të veçanta, shumë specifike për harduer. Të gjitha kërkesat mund të përmbushen nga çdo kompjuter i ditëve të sotme, edhe pse ka disa dallime.

Tabela 6.1 Kërkesat minimale për Android dhe Symbian OS

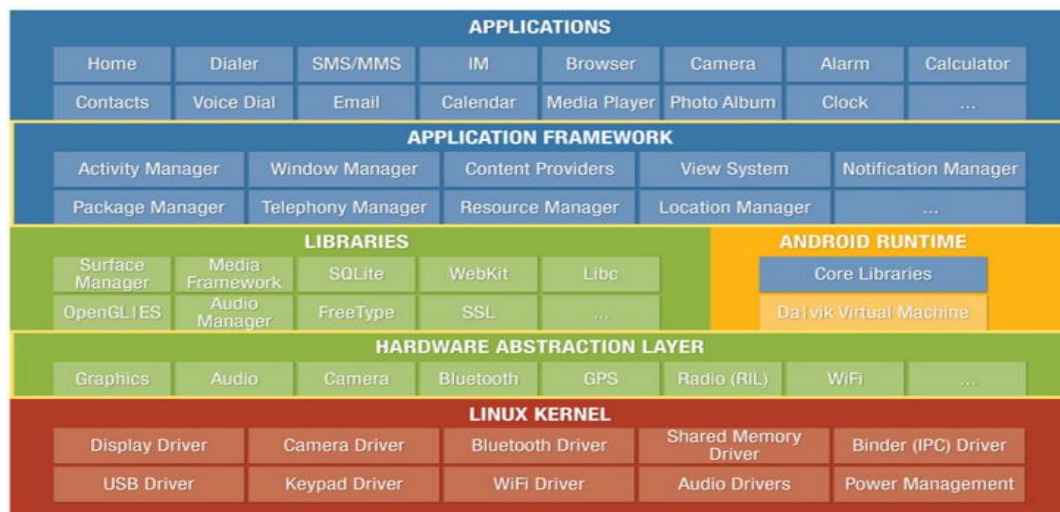
	Android	Symbian OS
Hardware Requirements (minimum)	200 megahertz online processor	1 GHz processor (minimum IDS and OS requirement of processor is 500 MHz)
	32 MB RAM and 32 MB Flash	512 MB RAM (minimum IDE and OS requirement is 128 MB of RAM)
Software Requirements (minimum)	JDK 5 or JDK 6 (JRE alone is not sufficient)	Java Runtime Version 1.5.0 or newer Active Perl Version 5.6.1 or newer
	Windows XP or Vista Mac OSX 10.4.8 or newer (only x86) Linux (tested on Linux Ubuntu Dapper Drake)	Microsoft Windows XP SP2 or Microsoft Windows 2000 SP4.
	Supported Development Environments (IDE): Eclipse 3.2 (Europa) or newer and JDT, WST and ADT plugins for Eclipse. Apache Ant 1.6.5 or newer.	Supported Development Environment (IDE): Eclipse 3.2 or newer and EclipseME version 1.7.7 or newer. NetBeans IDE with Mobility Pack 5.5 or newer.

SDK e Android është në gjendje të ekzekutohet në një sistem me procesor më të vogël dhe me më pak RAM në krahasim me një SDK Symbian OS. Një avantazh i Android SDK është se mund të instalohet në çdo sistem operativ, si: Windows, Mac ose Linux. Në tabelën e mësipërme jepet një krahasim midis kërkesave që ka një Android dhe një Symbian OS. (Aplikime mobile, 2015).

6.4 Arkitektura Android

Duke u nisur nga fundi kemi Linux Kernel. Android është ndërtuar me Linux Kernel. Linux është duke u përdorur gjerësisht prej shumë vitesh. Linux Kernel siguron funksionalitet bazë të sistemit në menaxhimin e proceseve, menaxhimin e memories, menaxhimin e pajisjeve, si: kamera, tastierë, ekranë etj. Gjithashtu, kernel-i menaxhon të gjitha gjërat që Linux shkëlqen në bërjen e tyre, si, për shembull, networking dhe një sërë drajverash për pajisje, duke e transferuar punën te hardueri periferik.

Figura 6.1 Arkitektura Android (Mattias, 2014)



Duke qenë baza për ambientin e kompjuterizimit mobil, Linux kernel i jep Android-it disa veçori kyçe sigurie, si:

- Një model lejeje i bazuar te përdoruesi
- Izolim të procesit
- Mekanizëm të gjerë për IPC të sigurt
- Aftësinë për të larguar pjesë të panevojshme dhe potencialisht të pasigurta të kernel

Si një sistem operativ multipërdorues, një objektiv thelbësor sigurie i Linux kernel është të izolojë burimet e përdoruesit nga njëra-tjetra. Filozofia e sigurisë e Linux është të mbrojë burimet e përdoruesit nga njëra-tjetra. Pra, Linux:

- Parandalon që përdoruesi A të lexojë fajllët e përdoruesit B.
- Bën të mundur që përdoruesi A të mos harxhojë memorien e përdoruesit B.
- Bën të mundur që përdoruesi A të mos harxhojë burimet e CPU-së së përdoruesit B.
- Bën të mundur që përdoruesi A të mos harxhojë pajisjet e përdoruesit B (për shembull, telefoninë, GPS, Bluetooth).
- **Hardware Abstraction Layer** u jep aplikacioneve akses të drejtpërdrejtë në burimet harduerike.
- Duke kaluar te pjesa e tretë, ndalemi te Libraritë, Android Runtime dhe Dalvik. Libraritë e treguara në foto janë shumë të rëndësishme, sepse pa to aplikacioni nuk mund të funksionojë si duhet. Libraritë për web duhen për navigimin në internet, kurse ato të SQLite duhen për mirëmbajtjen e databazës SQL e të tjera.

- **Dalvik Virtual Machine** është dizajnuar nga Android Open Source Project për të ekzekutuar aplikacionet e shkruara për Android. Çdo aplikacion që funksionon në një pajisje android ka Dalvik Virtual Machine.
- **Android Runtime (ART)** është një alternativë ndaj Dalvik Virtual Machine, e cila është linsuar me Android 4.4 si eksperiment, në Android Lollipop (5.0) ajo do të zëvendësojë tërësisht Dalvik Virtual Machine. Një ndryshim madhor në ART, sepse në Ahead-of-time(AOT) Compilation and Garbage Collection aplikacionet android kompilohen kur përdoruesi i instalon ato në pajisje, ndërsa me Dalvik që përdorte Just-in-time(JIT) bytekodet kompilohen kur përdoruesi përdor aplikacionin.
- **Framework-u i aplikacionit**
Shtresa Application Framework siguron shumë shërbime të nivelit të lartë për aplikacionet në formën e klasave Java. Zhvilluesit e aplikacioneve mund të përdorin këto shërbime në aplikacionet e tyre. (Aditya, 2015).

Pra, platforma Android është një koleksion softueri që përfshin një sistem operativ dhe një numër librarish të niveleve të larta që lehtësojnë detyrën e komunikimit me sistemin operativ. Ajo përfshin, gjithashtu, disa aplikacione që përdoren nga përdoruesit e celularëve, tabletave, e-lexuesve dhe televizorëve, si: aplikacioni për bërjen e telefonatave, e-mail, rrjete sociale, menaxhuesi i kontakteve, Google Maps, Google Search, shfletuesi i uebit, kalendari, lojëra bazë etj. Gjithçka në ambientin e zhvillimit të Android-it, po ashtu, edhe aplikacionet e përfshira, mund të programohet me kombinimin e Java dhe XML falë së ashtuquajturës *runtime*, që gjendet në Android SDK. *Runtime* e përkthen kodin Java dhe XML që ju shkruani në një gjuhë që sistemi operativ dhe pajisja e kuptojnë. Themeli mbi të cilin bazohet Android-i është i ndërtuar mbi Linux 2.6, një sistem operativ që rrallë bllokohet. Linux dhe shërbimet

e veta bazë menaxhojnë telefonin fizik, tabletin, e-book e-lexuesit ose iTV dhe u japin aplikacioneve Android akses total në tiparet e çdo pajisjeje konsumi, përfshirë ekranin me prekje, memorien, të dhënat, sigurinë, marrësit dhe dhënësit e ndryshëm, GPS, Bluetooth, Wi-Fi dhe të tjera. (Wallace, 2012).

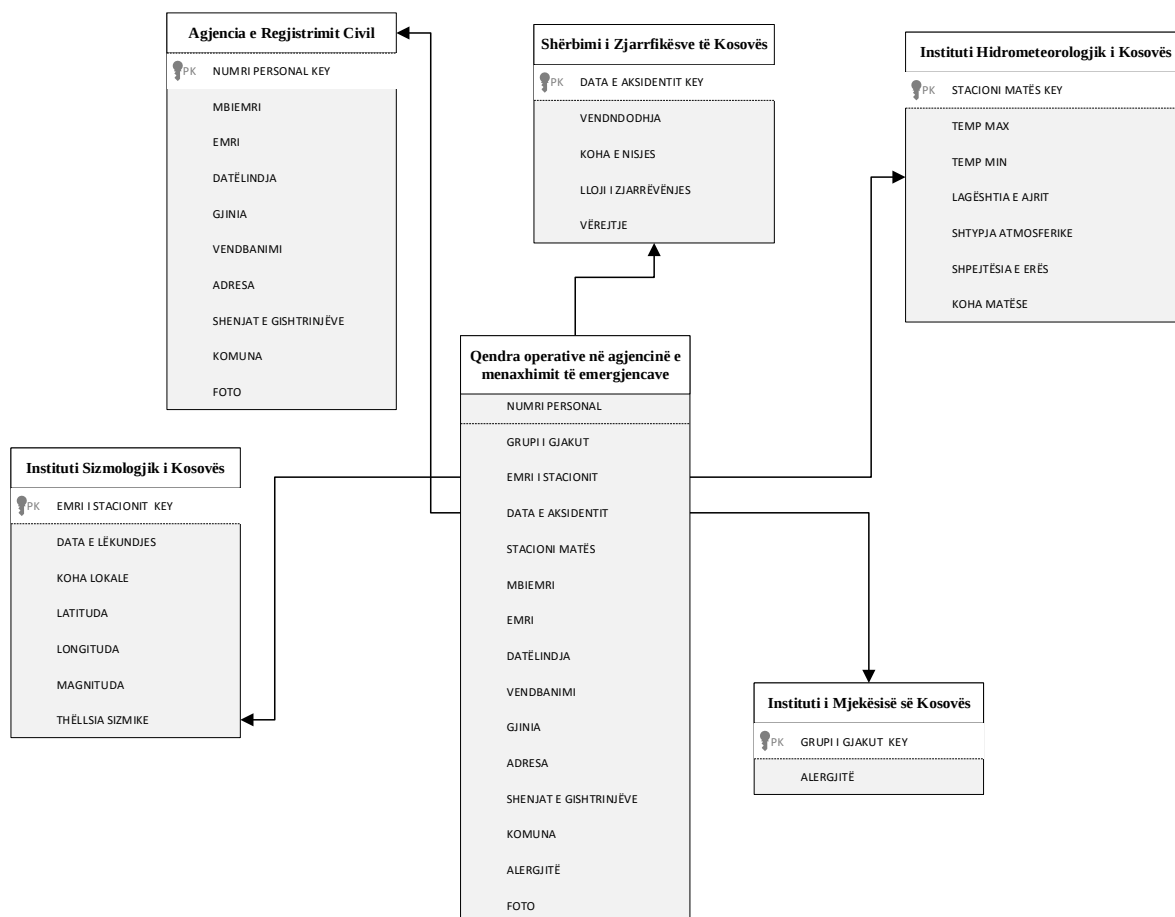
Marrja në konsideratë e Android-it, si platformë target për aplikacionet, vjen pas testeve që kemi bërë në punën tonë, duke përdor teknologjinë Android. Android SDK mundëson zhvillim aplikacionesh pa vështirësi. Ka shumë tipare dhe vegla për këtë në një pajisje Android, e cila mund të integrohet dhe programohet edhe nëpërmjet aplikacionit. (Tanya, 2014).

6.5 Rasti studimor: Integrimi i databazës SQLite_QOAME e implementueshme në pajisje mobile Android

Qendra Operative e Agjencisë së Menaxhimit të Emergjencave kërkon një analizë në baza ditore, javore dhe mujore. Por, mbi të gjitha, rëndësia më e madhe është analizimi i të dhënave në kohë reale. Kështu, rritet nevoja e krijimit, e grumbullimit të të dhënave dhe integrimin në një databazë të vetëm, me qëllim që nga analizat e problemeve në menaxhimin e fatkeqësive natyrore dhe emergjencave të mundemi të përcaktojmë në nivele më të larta të entiteteve nga mund të bëjmë grumbullimin e të dhënave.

Kështu, të dhënat e mëposhtme përbëjnë bazën e modelimit konceptual të databazës QOAME, që do ta ndërtojmë, si në figurën 6.2.

Figura 6.2 Diagrami Entity Relationship i databazës QOAME



Nga përshkrimi entiteteve të databazës QOAME është:

Entiteti Instituti i Mjekësisë së Kosovës përbëhet nga të dhënat: Grupi i gjakut për pacientët dhe nga Alergjitë e mundshme që paraqiten në popullatë. Të dhënat hyrëse nga IMK janë një burim informatash të rëndësishme në nivel kombëtar sa i përket shëndetit të popullsisë.

Entiteti Agjencia e Regjistrimit Civil përfshin fushat që karakterizojnë shënimet bazë mbi të dhënat hyrëse të regjistrimit civil në nivel vendi, duke përfshirë këto karakteristika: numri personal, mbiemri, emri, datëlindja, vendbanimi, gjinia, adresa, shenjat e gishtërinjve, komuna dhe foto. Të gjitha këto attribute të entitetit ARC përfshijnë shënimet e popullatës.

Në studimin tonë si atribut kryesor që do kërkohen të dhënat për një individ do të jenë numri personal dhe shenjat e gishtërinjve.

Entiteti Instituti Sizmologjik i Kosovës përfshin atributet: emri i stacionit nga i cili merren të dhëna në kohë reale mbi lëkundjet sizmike, duke marrë informacione nga instrumentet që bëjnë monitorimin online të lëkundjeve sizmike. Më pas në këtë entitet përfshihet: data e lëkundjes, koha lokale, latituda, longituda, magnituda dhe thellësia sizmike.

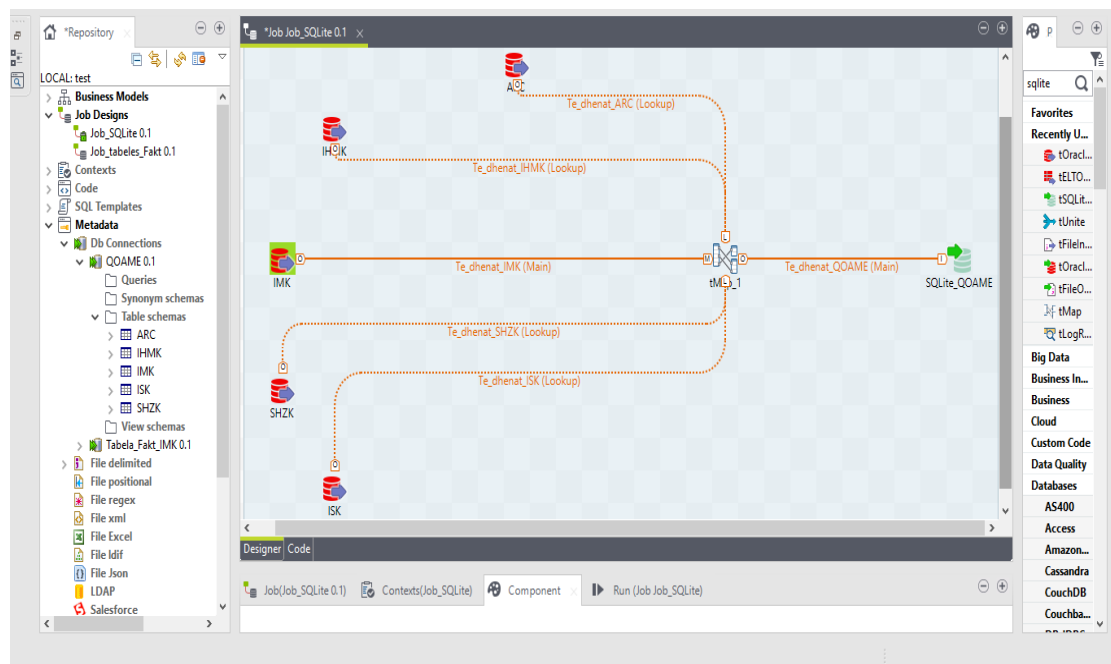
Te entiteti i të dhënave që gjeneron Instituti Hidrometeorologjik i Kosovës përfshin një dinamikë mjaft të theksuar të të dhënave që paraqiten, duke synuar në kohë reale dhe monitorim gjithëpërfshirës në parandalimin e vërshimeve nga reshjet e shiut, të borës, si dhe të thatësisë nga vlerat ekstreme të temperaturave maksimale. Ky entitet përfshin atributet e mësipërme: stacioni matës, temperatura maksimale, temperatura minimale, lagështia e ajrit, shtypja atmosferike, shpejtësia e erës, koha matëse.

Duke pasur nën kontroll monitorimin dhe vëzhgimin e vazhdueshëm të shumë faktorëve që ndikojnë në zjarrvënien, si në rastin e shtëpive, të pyjeve dhe kullosave, atëherë entiteti i Shërbimit të Zjarrfikësve përbëhet nga këto attribute: data e aksionit, vendndodhja, koha e nisjes, lloji i zjarrvënies, vërejtjet. Të gjitha këto japin një mundësi shumë të mirë të marrjes së informacioneve nga ky institucion shumë i rëndësishëm.

6.5.1 Integrimi i të dhënave në SQLite_QOAME përmes Talend Open Studio

Zgjedhja SQLite si databazë kryesore në të cilën integrohen tabelat nga Oracle SQL, si: ARC, IHMK, IMK, ISK dhe SHZK, ku shumica e kompanive përdorin Oracle në aplikacionet e tyre, për rastin e ngarkimit të të dhënave në SQLite_QOAME është ndërtuar job-i i mëposhtëm në Talend Open Studio for Big Data, siç edhe paraqitet në figurën e mëposhtme.

Figura 6.3 Job-i Talend për ngarkimin e të dhënave nga tabelat ORACLE SQL në SQLite_QOAME



Këto burime (IHMK, IMK, SHZK, ISK dhe ARC) heterogjene të dhënash më pas vendosen në databazat SQLite_QOAME, siç paraqitet në skemën TOS.

Hapi i rëndësishëm në zhvillimi e punëve me ETL është të sigurohet lidhje e suksesshme mes TOS dhe databazës SQLite _QOAME.

6.5.2 Projektimi i Databazës SQLite _QOAME

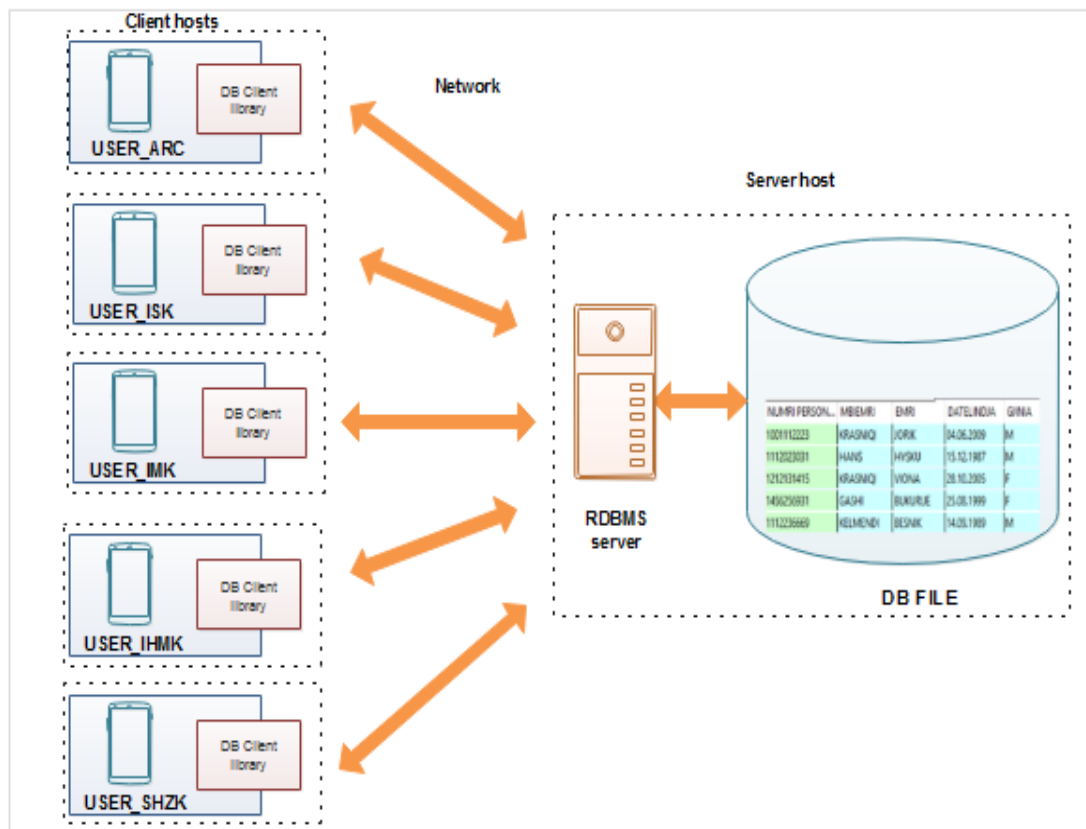
Ndryshe nga produktet e tjera RDBMS, SQLite nuk ka një arkitekturë klient/server. Shumica e sistemeve të mëdha të databazave kanë një paketë të madhe server që përbën veglën e databazës. Serveri i databazës shpesh përbëhet nga procese të shumëfishta që punojnë bashkë për të menaxhuar lidhjet e klientit, fajllat I/O, depot e përkohshme, optimizmi i kërkesave dhe procedimi i tyre.

Një databazë tipike përbëhet nga një numër i madh fajllash të organizuar në një ose më shumë folder në sistemin e fajllave të serverit.

Për të pasur akses në databazë, të gjithë fajllat duhet të jenë prezent dhe të rregullt. Kjo mund ta bëjë disi të vështirë për të lëvizur ose për të krijuar një kopje rezervë të një databaze. Për të pasur akses në databazë, klienti i databazës siguron edhe libraritë e softuerëve. Këto librari duhet të integrohen në çdo aplikacion klient që kërkon të ketë akses në serverin e databazës. Këto librari klient ofrojnë API për të gjetur dhe hyrë lidhur me serverin e databazës, si dhe rregullojnë dhe ekzekutojnë kërkesa dhe komanda të databazës.

Pra, organizimi arkitekturor klient/server i RDBMS-së është paraqitur në figurën 6.4.

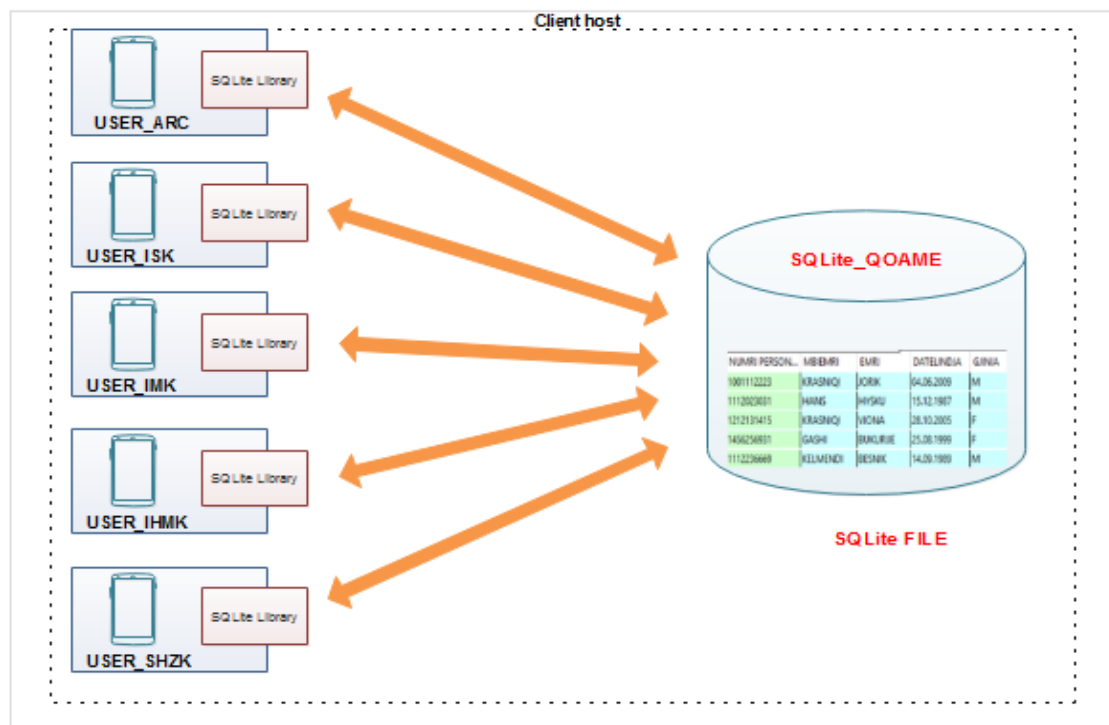
Figura 6.4 Organizimi arkitekturor klient/server për RDBMS (Jay, 2010).



Në kontrast me këtë, SQLite nuk ka server të ndarë. Vegla e databazës është e integruar në çfarëdo lloj aplikacioni që ka nevojë të ketë akses në databazë.

I vetmi burim i ndarë mes aplikacioneve është fajlli i vetëm i databazës që gjendet në disk. Nëse dëshirojmë të lëvizim ose rezervojmë një kopje të databazës, thjesht kopjohen këta fajlla. Më pas do paraqesim organizimin arkitekturor të SQLite si në figurën 6.5.

Figura 6.5 Organizimi arkitekturor i serverit SQLite (Jay, 2010).



Duke eliminuar serverin, arrihet të shmanget një kompleksitet goxha i madh. Kjo thjeshton përbërësit e softuerit dhe thuhet eliminon nevojën për suport të sistemit të operimit. Ndryshe nga një server tradicional, RDBMS kërkon shumë ngarkesë të avancuar dhe komunikim me performancë të lartë mes proceseve.

SQLite kërkon pak më shumë se aftësia për të lexuar dhe shkruar një lloj magazinimi. SQLite është i dizajnuar të integrohet direkt si një i ekzekutueshëm. Kjo eliminon nevojën për librari të jashtme dhe thjeshton shpërndarjen dhe instalimin. Heqja e varësive të jashtme eliminon edhe çështjet e versioneve.

Nëse kodi SQLite është ndërtuar direkt në aplikacion, nuk duhet të shqetësoheni për lidhjet me versionin e duhur të një librerie klient apo për përshtatjen e librarisë klient me serverin e databazës.

Eliminimi i serverit, megjithatë, sjell disa kufizime. SQLite është i dizajnuar të adresojë nevojat e lokalizuara për magazinim, si, për shembull, aksesin në një ueb serveri në një databazë lokale. Kjo do të thotë se nuk është shumë i përshtatshëm për situata ku disa kompjuterë klientë, njëkohësisht, kërkojnë akses në një databazë të centralizuar. Kjo situatë ndodh më shumë me arkitekturën klient/server dhe zgjidhet nga një sistem databazë që përdor të njëjtën arkitekturë (Jay, 2010).

6.6 Zhvillimi hyrës në aplikacionin e QOAME-së

Një aplikacion që shërben për t'u loguar është ekrani që u kërkon kredencialet për t'u loguar në aplikacionin e QOAME-së. Në vijim shpjegohet si të krijohet aplikacioni i tillë dhe si ta menaxhoni sigurinë kur dikush tenton të futet i paautorizuar.

Së pari, duhet të definoni dy TextView që i kërkon përdoruesit të fusë emrin e përdoruesit dhe fjalëkalimin. Fjalëkalimi TextView duhet të ketë inputType për fjalëkalimin. Sintaksa jepet më poshtë:

```

<EditText android:id="@+id/editText2"
android:layout_width="wrap_content"
android:layout_height="wrap_content"
android:inputType="textPassword" />
<EditText
android:id="@+id/editText"
android:layout_width="wrap_content"
android:layout_height="wrap_content"
/>

```

Më pas kemi një buton me tekst logimi dhe vendoseni në onClick Property. Pas kësaj funksionin e përmendur në onClick property në fajllin java.

```

<Button
android:id="@+id/button1"
android:layout_width="wrap_content"
android:layout_height="wrap_content"
android:onClick="login"
android:text="@string/Login"
/>

```

Brenda metodës së onClick merret emri i përdoruesit dhe fjalëkalimi, duke përdorur metodën getText() dhe toString() dhe përputhet me tekstin duke përdorur funksionin equals().

```

EditText username = (EditText)findViewById(R.id.editText);
EditText password = (EditText)findViewById(R.id.editText2);
public void login(View view){
if(username.getText().toString().equals("Jorik Viona ") &&
password.getText().toString().equals("admin")){
//correct password }
else{
// password i gabuar
}
}

```

Gjëja e fundit që duhet vepruar është të sigurohet një mekanizëm sigurie, që të evitohen tentativat e padëshiruara. Për këtë arsye krijohet një ndryshore dhe në çdo tentativë të rreme pakësohet ajo. Kur kjo të arrijë në 0 çaktivizohet butoni për logim.

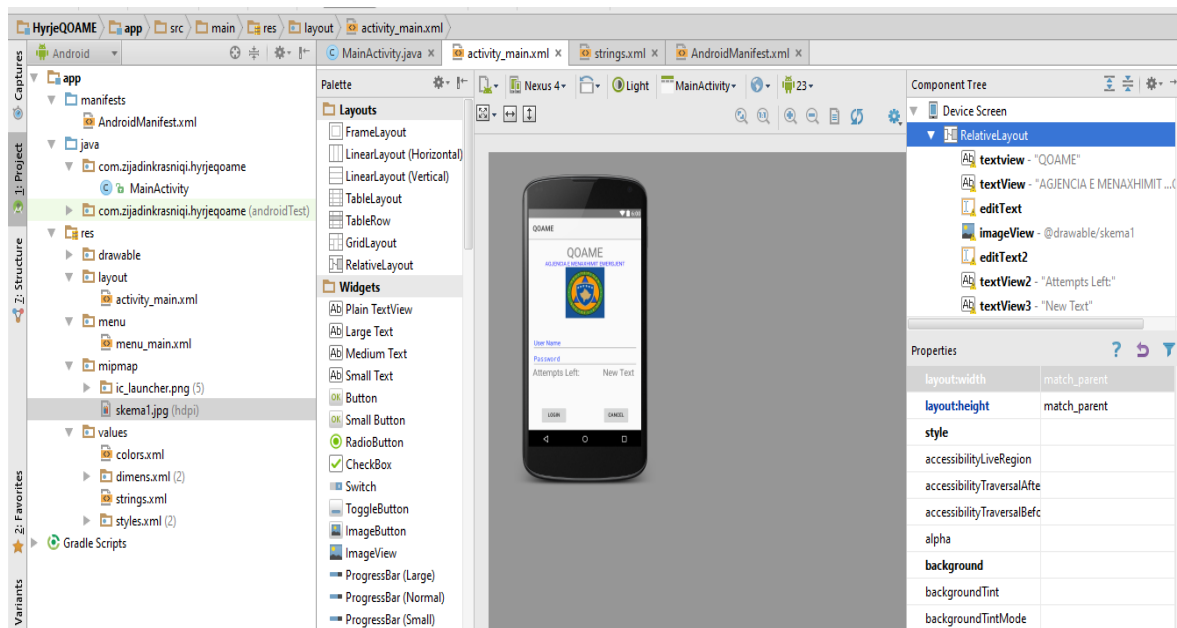
```

int counter = 3;
counter--;
if(counter==0){
// butoni jo aktiv, të mbyllet aplikacioni
}

```

Ai krijon një aplikacion bazë që u ofron vetëm tre herë mundësinë që të kyçeni në aplikacionin QOAME. Në figurën 6.6 është paraqitur forma hyrëse në QOAME .

Figura 6.6 Aplikacioni hyrës për QOAME



Për të nisur aplikacionin nga Android studio hapet aktivitetit i projektit HyrjeQOAME dhe klikohet te ikona Run.

Android studio instalon aplikacionin në AVD dhe e aktivizon. Nëse gjithçka është në rregull me setup-in dhe aplikacionin, do të shfaqet dritarja vijuese Emulator me mundësinë e vendosjes se username dhe password-it në aplikacionin QOAME, siç paraqitet në figurën 6.7

Figura 6.7 Aplikacioni për login në QOAME pas ekzekutimit të emulator-it



6.7 Krijimi i aktiviteteve duke përdorur intents

Një aplikacion mund të ketë zero aktivitete ose më shumë. Tipikisht, aplikacionet kanë një ose më shumë aplikacione dhe qëllimi kryesor i një aktiviteti është të ndërveprojë me përdoruesin.

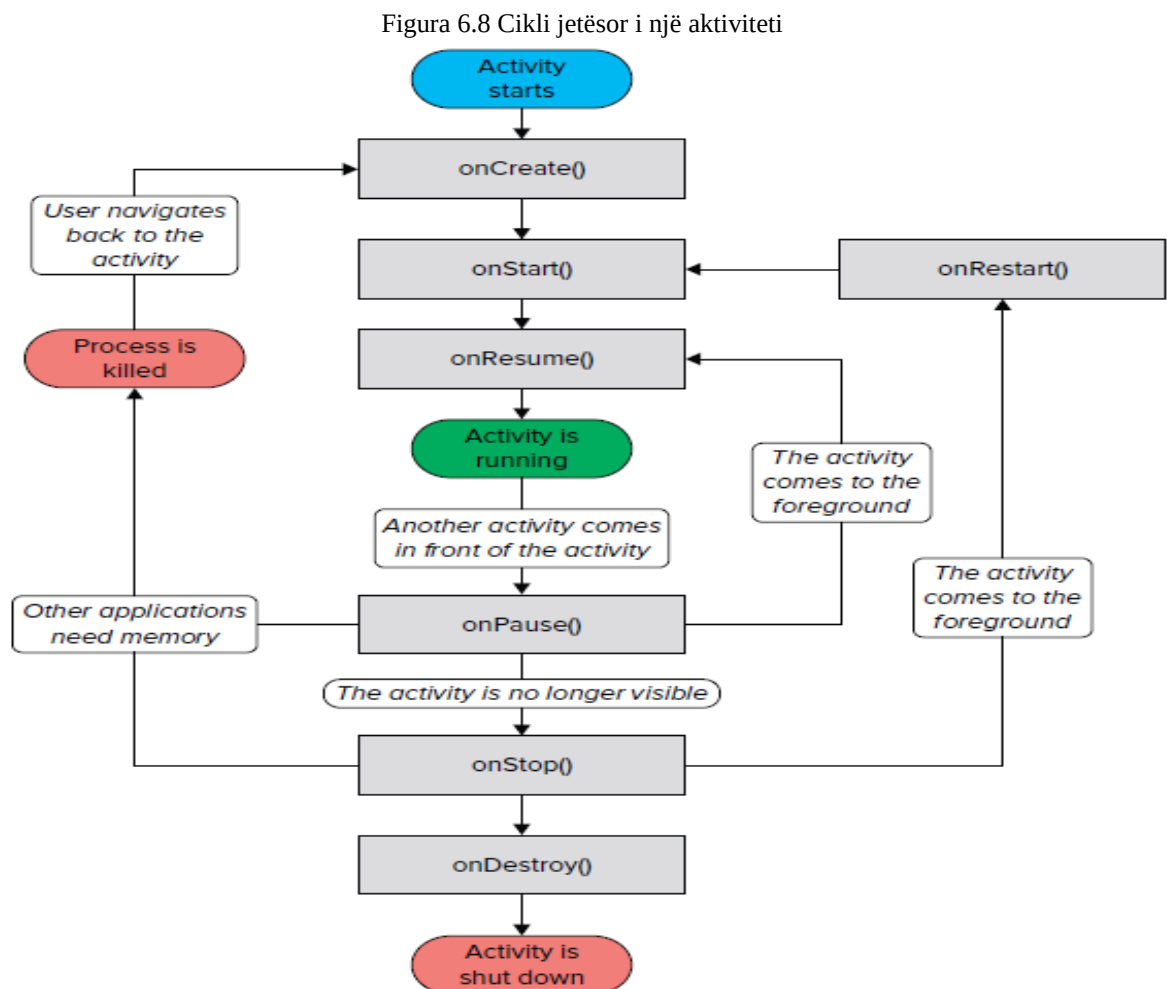
Që nga momenti kur një aktivitet shfaqet në ekran e deri në momentin që nuk duket më, ai kalon nëpër një sërë fazash, të njohura si cikli jetësor i aktivitetit. Të kuptuarit e këtij cikli jetësor është jetik për të qenë të sigurt që aplikacioni punon si duhet.

Krahas aktiviteteve, Android 4.0, gjithashtu, përkrah një veçori që është futur në Android 3.0 fragmente. Sipas kësaj, fragmentet si aktivitete “miniaturë” mund të grupohen për të formuar një aktivitet.

Krahas aktiviteteve, një tjetër veçori unike në Android, është ajo e quajtur intent. Një intent është praktikisht “ngjitësi” që u mundëson aktiviteteve të ndryshme nga aplikacione të tjera të funksionojnë së bashku, duke bërë të mundur që detyrat të performohen si duhet si t’i përkisnin një aplikacioni të vetëm.

Në vijim do paraqesim më shumë për këtë tipar shumë të rëndësishëm dhe si mund të përdoret në rastin e aplikacioneve të tilla, si: Shfletues, Telefon, Harta etj.

Kështu, klasa bazë e Activity definon një sërë ngjarjesh që dominojnë ciklin jetësor të një aktiviteti. Figura 6.8 tregon ciklin jetësor të një aktiviteti dhe fazat e ndryshme nëpër të cilat kalon qëkur fillon e deri kur mbaron aktiviteti.



Kjo klasë Activity definon evenimentet vijuese:

- ▶ **onCreate()** — Thirret kur krijohet së pari aktiviteti.
- ▶ **onStart()** — Thirret kur aktiviteti bëhet i dukshëm ndaj përdoruesit.
- ▶ **onResume()** — Thirret kur aktiviteti nis të ndërveprojë me përdoruesin.
- ▶ **onPause()** — Thirret kur aktiviteti i tanishëm ndalet dhe nis tjetri, i mëparshmi.
- ▶ **onStop()** — Thirret kur aktiviteti nuk është më i dukshëm për përdoruesin.
- ▶ **onDestroy()** — Thirret para se aktiviteti të shkatërrohet nga sistemi (manualisht ose nga sistemi për të ruajtur memorie).
- ▶ **onRestart()** — Thirret kur aktiviteti është ndalur dhe pastaj ristartohet.

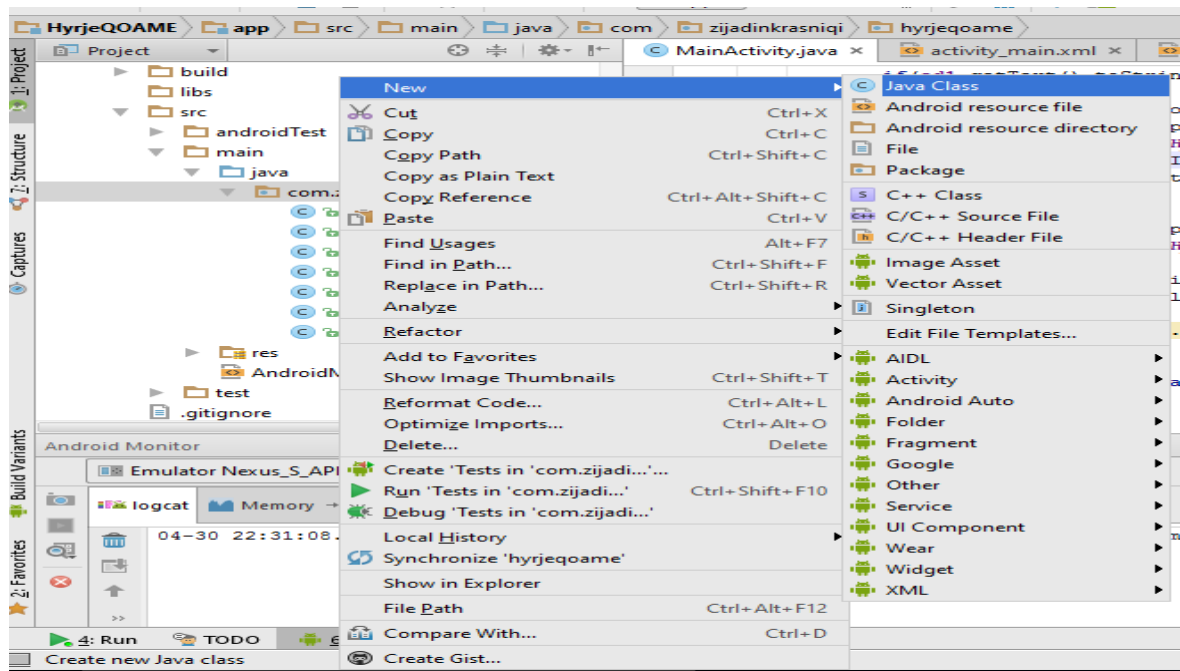
Automatikisht, aktiviteti i krijuar përmban ngjarjen onCreate(). Këtu gjendet edhe kodi që ndihmon të shfaqen elementet UI në ekran.

Mënyra më e mirë për të kuptuar këtë koncept abstrakt shumë të rëndësishëm në Android është se si t'i shtohet një tjetër aktivitet atij ekzistues dhe të navigoni ndërmjet dy aktiviteteve.

Për të krijuar një aktivitet duhet të krijohet në Java klasën që zgjeron klasën Activity:

Më pas klikohet me të djathtë emrin e paketës nën fajllin src dhe selektohet New-Class, emërohet klasa e re Welcome, siç paraqitet në figurën 6.9 në vijim.

Figura 6.9 Klasa Welcome



Në fajllin e krijuar Welcome.java shtohet pjesa e kodit:

```
package com.zijadinkrasniqi.hyrjeqoame;

import android.app.Activity;
import android.os.Bundle;

public class Welcome extends Activity {
    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate(savedInstanceState);
        setContentView(R.layout.welcome);
    }
}
```

Aktiviteti ngarkon përbërësin e vet UI, duke përdorur fajllin XML të përkufizuar në dosjen res/layout. Në këtë shembull do të ngarkojmë UI nga fajlli main.xml:

```
setContentView(R.layout.main);
```

Çdo aktivitet që kemi në aplikacion duhet të deklarohet në fajllin AndroidManifest.xml si vijon:

```
<?xml version="1.0" encoding="utf-8"?>
<manifest xmlns:android="http://schemas.android.com/apk/res/android"
    package="com.zijadinkrasniqi.hyrjeqoame" >

    <uses-permission android:name="android.permission.INTERNET" />

    <application
        android:allowBackup="true"
        android:icon="@mipmap/skema1"
        android:label="@string/app_name"
        android:theme="@style/AppTheme" >

        <activity
            android:name=".MainActivity"
            android:label="@string/app_name" >

            <intent-filter>
                <action android:name="android.intent.action.MAIN" />
                <category android:name="android.intent.category.LAUNCHER"
/>
            </intent-filter>

        </activity>

        <activity

            android:name=".Welcome"
            android:label="@string/title_activity_welcome" >

            <intent-filter>
                <action android:name="com.zijadinkrasniqi.hyrjeqoame.Welcome"
/>
                <category android:name="android.intent.category.DEFAULT" />
            </intent-filter>
        </activity>

    </application>
</manifest>
```

Këtu, me këtë, i kemi shtuar një aktivitet të ri aplikacionit.

```

<activity
    android:name=".Welcome"
    android:label="@string/title_activity_welcome" >

    <intent-filter>
        <action android:name="com.zijadinkrasniqi.hyrjeqoame.Welcome" />

        <category android:name="android.intent.category.DEFAULT" />
    </intent-filter>
</activity>

```

- Emri (klasa) i aktivitetit të shtuar është Welcome.
- Etiketa për aktivitetin e ri quhet Welcome.
- Filteri i emrit të intent-it për aktivitetin e ri është com.zijadinkrasniqi.hyrjeqoame.Welcome

Të tjerë aktivitete që duan ta thërrasin këtë aktivitet do ta thërrasin me emër. Idealisht, ju duhet të përdorni emrin e domeinit të kompanisë suaj së mbrapshti si emër të filtrit të intent-it për të pakësuar shanset që një tjetër aplikacion të ketë filtër intent me të njëjtin emër.

►Kategoria për filtrin e intent-it është android.intent.category.DEFAULT. Më pas duhet shtuar këtë filtër me qëllim që të nisë ky aktivitet nga një tjetër aktivitet, duke përdorur metodën startActivity().

Pra, përdorimi i metodës startActivity() për të shfaqur Welcome, duke krijuar një instancë ose objekt të klasës Intent dhe duke e kaluar atë në emrin e filtrit të intent Welcome (që është com.zijadinkrasniqi.hyrjeqoame.Welcome):

```
Intent intent = new Intent("com.zijadinkrasniqi.hyrjeqoame.Welcome");
```

Aktivitetet në Android mund të invokohen nga secili prej aplikacioneve që përdoret në pajisje. Për shembull, ju mund të krijoni një projekt të ri Android dhe më pas të shfaqni

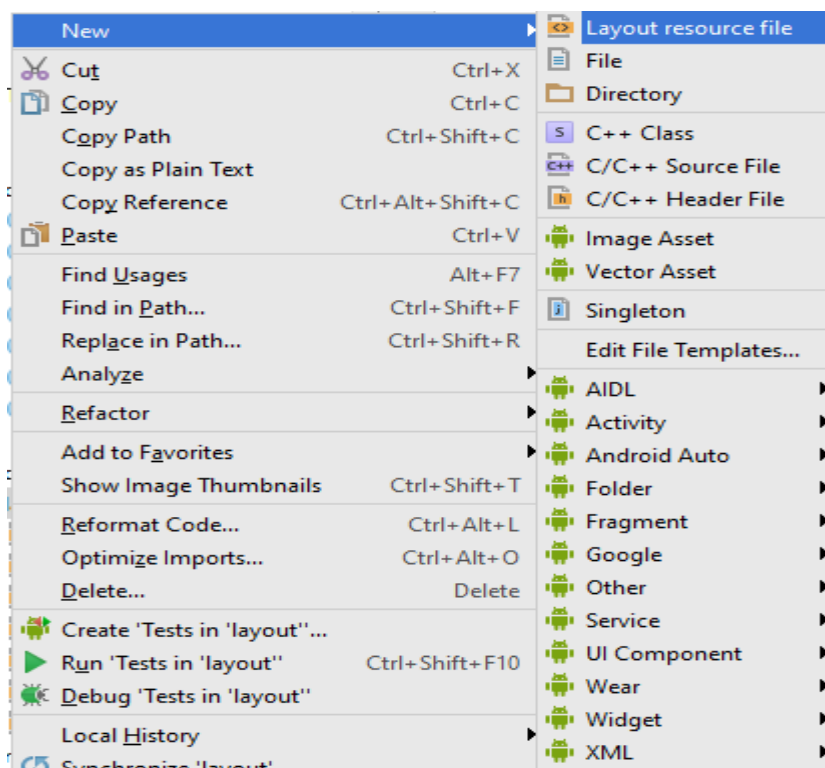
Welcome, duke përdorur filtrin e intent-it të saj `com.zijadinkrasniqi.hyrjeqoame.Welcome`. Ky është ndër konceptet themelore në Android që mundëson të invokohet lehtësisht një tjetër aplikacion lehtësisht. Nëse aktiviteti që dëshirohet të invokoni është i definuar brenda të njëjtit projekt, ju mund të rishkruani stejtmntin e mëparshëm kështu :

```
startActivity(intent);
```

Megjithatë, kjo qasje është e aplikueshme vetëm nëse aktiviteti që doni të shfaqni është brenda të njëjtit projekt si aktiviteti në fjalë.

Pasi të bëhet një kopje e fajllit `main.xml` file (në dosjen `res/layout`), duke klikuar me të djathtën dhe duke selektuar Copy, klikoni me të djathtën në dosjen `res/layout` dhe selektoni Paste. Emërtohet fajlli `secondactivity.xml`. Dosja `res/layout` do të përmbajë fajllin `welcome.xml` (Wei-Meng, 2012), siç paraqitet në figurën 6.10

Figura 6.10 Emërtimi i fajllit `secondactivity.xml`.



Më pas modifikohet fajlli welcome.xml, si vijon:

```
<?xml version="1.0" encoding="utf-8"?>
<RelativeLayout
    xmlns:android="http://schemas.android.com/apk/res/android"
    android:orientation="vertical"
    android:layout_width="match_parent"
    android:layout_height="match_parent">
    tools:context="com.zijadinkrasniqi.hyrjeqoame.welcome">

    <ImageView
        android:layout_width="wrap_content"
        android:layout_height="wrap_content"
        android:id="@+id/imageView"
        android:src="@drawable/skema1"
        android:layout_below="@+id/textView"
        android:layout_centerHorizontal="true" />

    <TextView
        android:layout_width="wrap_content"
        android:layout_height="wrap_content"
        android:textAppearance="?android:attr/textAppearanceLarge"
        android:text="MIRE SE VINI NE QOAME"
        android:id="@+id/textView4"
        android:layout_below="@+id/imageView"
        android:layout_centerHorizontal="true" />

    <Button
        android:layout_width="wrap_content"
        android:layout_height="wrap_content"
        android:text="ARC"
        android:id="@+id/button3"
        android:layout_marginTop="143dp"
        android:layout_below="@+id/textView4"
        android:layout_toStartOf="@+id/button7"
        android:layout_toLeftOf="@+id/button7" />

    <Button
        android:layout_width="wrap_content"
        android:layout_height="wrap_content"
        android:text="IHMK"
        android:id="@+id/button4"
        android:layout_alignTop="@+id/button3"
        android:layout_toRightOf="@+id/imageView"
        android:layout_toEndOf="@+id/imageView" />

    <Button
        android:layout_width="wrap_content"
        android:layout_height="wrap_content"
        android:text="IMK"
        android:id="@+id/button5"
        android:layout_below="@+id/button3"
        android:layout_alignLeft="@+id/button3"
```

```

        android:layout_alignStart="@+id/button3" />

<Button
    android:layout_width="wrap_content"
    android:layout_height="wrap_content"
    android:text="ISK"
    android:id="@+id/button6"
    android:layout_below="@+id/button4"
    android:layout_alignLeft="@+id/button4"
    android:layout_alignStart="@+id/button4" />

<Button
    android:layout_width="wrap_content"
    android:layout_height="wrap_content"
    android:text="SHZK"
    android:id="@+id/button7"
    android:layout_below="@+id/button5"
    android:layout_centerHorizontal="true" />
</RelativeLayout>

```

Nëse gjithçka është në rregull me setup-in dhe aplikacionin, do të shfaqet dritarja vijuese emulator me aplikacionin QOAME, siç paraqitet në figurën 6.11.

Figura 6.11 Aplikacioni bazë i QOAME-së



6.8 Databaza SQLite_QOAME, rasti i aplikacionit android për ARC

Jetëgjatësia e çdo aplikacioni Android është vetë suporti për databazën e saj. Sistemi i databazës nevojitet për të depozituar të dhënat e strukturuar, me përjashtim të rasteve kur aplikacioni merret vetëm me të dhëna të thjeshta.

Android përdor sistemin e databazës SQLite, që është databazë SQL me kod të hapur, e përdorur gjerësisht nga shumë aplikacione popullore.

SQLite është një vegël database i lehtë transaksional që zë pak hapësirë disku dhe memorieje, domethënë është perfekt për krijimin e databazave në shumë sisteme operative për celularë si Android dhe iOS. (Aditya, 2015)

Platforma android ofron mbështetje të plotë për databazat SQLite. Kjo databazë e krijuar do të jetë e aksesueshme me emër nga secila klasë në aplikacion, por jo jashtë aplikacionit. Për rastin tonë të kërkimit të të dhënave nga Agjencia e Regjistrimit Civil në raste të fatkeqësive natyrore, është mundësia e kërkimit të personave gjatë kërkim-shpëtimit, po ashtu, edhe në periudhën e paralajmërimeve nga fatkeqësitë natyrore.

Pra, përmes numrit personal bëhet identifikimi i personave të gjetur në vendin e ndodhur të fatkeqësive natyrore. Nëse marrim tabelën e Agjencisë së Regjistrimit Civil të databazës SQLite_QOAME, si në tabelën vijuese:

Tabela 6.2 Tabela e të dhënave nga Agjencia e Regjistrimit Civil të databazës SQLite_QOAME

NUMRI PERSONAL	MBIEMRI	EMRI	DATLINDJA	GJINIA	KOMUNA	ADRESA
1001112223	KRASNIQI	JORIK	04.06.2009	M	PRISHTINE	"Bajram Bahtiri"
1112023031	HANS	HYSKU	15.12.1987	M	PEJE	"Lidhja e Prizrenit"
1212131415	KRASNIQI	VIONA	28.10.2005	F	PODUJEVE	"Ismail Dumoshi"
1456256931	GASHI	BUKURIJE	25.08.1999	F	MITROVICE	"Bajram Bahtiri"
1112236669	KELMENDI	BESNIK	14.09.1989	M	FERIZAJ	"Lidhja e Prizrenit"
1478523691	ISLAMI	ILIR	15.05.1970	M	PRISHTINE	"Ismail Dumoshi"
2581473691	KADRIU	LIRI	18.02.2000	F	PEJE	"Bajram Bahtiri"
3698521471	GASHI	ISA	23.04.1988	M	PEJE	"Lidhja e Prizrenit"
1234567893	KODRA	GEZIM	01.01.1975	M	PRISHTINE	"Ismail Dumoshi"
1593574562	TELAKU	TRIM	02.02.2012	M	DEÇAN	"Bajram Bahtiri"
1235661114	RUGOVA	JORIK	14.09.2014	M	PRISHTINE	"Lidhja e Prizrenit"
4569871233	HALILI	RRON	14.05.2007	M	PEJE	"Ismail Dumoshi"
1488523691	HALILI	LUM	25.03.2011	M	PRIZREN	"Bajram Bahtiri"

- Kërkimi i të dhënave SQLite_QOAME sipas numrit personal

Nëse marrim për bazë kërkimin e të dhënave sipas numrit personal nga databaza SQLite_QOAME dhe i shfrytëzojmë metodat e rekomanduara për të krijuar një databazë të re SQLite krijohet një subklasë e SQLiteOpenHelper

```
public class DatabaseHelper extends SQLiteOpenHelper {

    private static final int DATABASE_VERSION = 1;

    private static final String DATABASE_NAME = "SQLite_QOAME";

    private static final String TABLE_NAME = "Agjencioni i regjistrimit civil";

    private static final String KEY_NUM_PERSONAL = "Numeri personal";
    private static final String COL_1 = "Mbiemri";
    private static final String COL_2 = "Emri";
    private static final String COL_3 = "Datelindja";
    private static final String COL_4 = "Gjinia";
    private static final String COL_5 = "Komuna";
    private static final String COL_6 = "Adresa";
    public DatabaseHelper(Context context) {
        super(context, DATABASE_NAME, null, DATABASE_VERSION);
    }
}
```

në të cilën komanda SQLite ekzekutohet për të krijuar tabelën e Agjencisë së Regjistrimit Civil në databazën SQLite_QOAME. Mund të shikoni implementimin e SQLiteOpenHelper, duke përdorur konstruktorin që po definohet.

Për të shkruar dhe lexuar nga databaza thirren në aksion `getWritableDatabase()` dhe `getReadableDatabase()`.

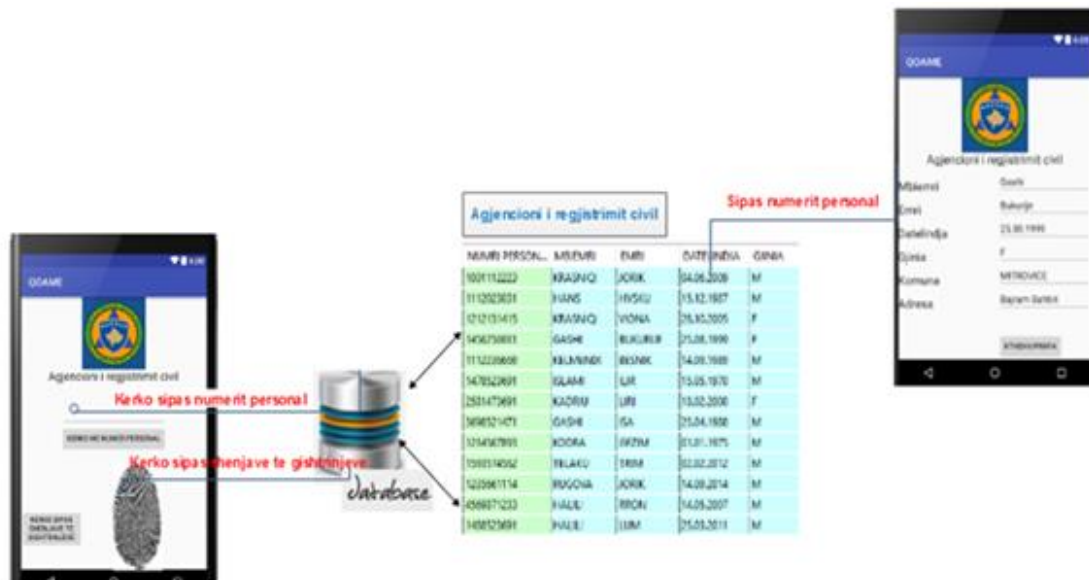
```
void addContact(Contact contact) {  
    SQLiteDatabase db = this.getWritableDatabase();  
  
    ContentValues values = new ContentValues();  
    values.put(KEY_NAME, contact.getName()); // Contact Name  
    values.put(KEY_PH_NUM, contact.getPhoneNumber()); // Contact Phone
```

Të dyja kthejnë një objekt `SQLiteDatabase` që përfaqëson databazën dhe ofron metodat për operacionet SQLite. (Sunguk, 2012).

Kur një databazë është hapur me sukses, SQLite Open Helper do ta aksesojë atë, kështu që ju mund dhe duhet t'i përdorni këto metoda sa herë që bëni një kërkim ose performoni një transaksion në databazë, sesa te aksesori me aplikacionin tuaj.

Tentativa për të përdorur `getWritableDatabase` mund të dështojë për shkak të hapësirës së diskut ose çështjeve që lidhen me dhënien e lejes, prandaj praktika e mirë duhet të përfshijë edhe përdorimin e metodës `getReadableDatabase` për kërkimet në databazë, nëse nevojitet. Në shumicën e rasteve, kjo metodë do të ofrojë të njëjtën instancë të rishkrueshme të databazës, si edhe `getWritableDatabase` (Reto, 2012). Një shembull tregohet në figurën 6.12 .

Figura 6.12 Kërkimi sipas numrit personal në databazën SQLite_QOAME



- **Kërkimi i identifikimit të personave përmes shenjave të gishtërinjve me aplikacionet mobile**

Sipas Pavalloia (2013), aktivitetet më të mira për pajisjet përfshijnë aksesin në informacionet lokale (88%) dhe kërkimin për informacione (82%).

Në këtë punim doktrate synojmë integrimin e sistemit të njohjes nëpërmjet shenjave të gishtërinjve që mund të përdoret në një pajisje mobile, nën platformën Android.

Ne mund të zgjedhim nëpërmjet qasjes kryesore të një sistemi mobil autonom. Për të ndërtuar një sistem autonom ne integrojmë një librari të re në sistemin Android. SourceAFIS është një librari e përdorur për njohjen/përputhjen e shenjave të gishtërinjve. Funksionaliteti esencial i Automated Fingerprint Identification System (AFIS) konsiston në krahasimin e dy shenjave të gishtërinjve dhe marrjen e vendimit nëse i takojnë të njëjtit person. SourceAFIS siguron kërkim të shpejtë në databazën e shenjave të gishtërinjve të regjistruar. Ai ka një API të

thjeshtë (.NET dhe një Java portë eksperimentale) dhe aplikacione e vegla shtesë. Sipas dokumentacionit zyrtar të SourceAfis (2015), tiparet kryesore janë:

- përkrah imazhet e shenjave të gishtërinjve të të gjithë lexuesve të zakonshëm të shenjave të gishtërinjve;
- kërkim i shpejtë (10.000 shenja gishtash në sekondë) në databazë;
- përkrah dhe eksporton shabllonet sipas standardit ISO/IEC 19794-2;
- portabilitet mesatar;
- përfshin vizualizimin e shenjave të gishtërinjve dhe veglën e analizimit.

Ndërfaqja grafike e ndihmon përdoruesin të mbledhë imazhet e shenjave të gishtërinjve dhe t'i krahasojë me shenjat e gishtërinjve nga një databazë e madhe. Procesi është shumë i shpejtë dhe me produktivitet të lartë, siç paraqitet edhe në figurën 6.13.

Figura 6.13 Procesi i krahasimit sipas shenjave të gishtërinjve në databazën SQLite_QOAME



Butoni “Kërko sipas shenjave të gishtërinjve” përdoret për të aktivizuar klasën që implementon logjikën e krahasimit mes shenjave të gishtave. Për eventin onClick përdoret metoda `toFingerprint Recognition ()`. Këtu theksohet krijimi i instancës së dy objekteve `ImageView` te `classFingerprint Activity`. Për t’i krahasuar shenjat e gishtave, është kusht themelor që të konvertohen në formatin `ByteArray` imazhet bitmap në listën e dërgimit. Kjo arrihet përmes metodës së konvertimit `BitmapToByteArray` që jepet më poshtë.

```
public byte[] convertBitmapToByteArray(Bitmap imgJorikKrasniqi) {
    ByteArrayOutputStream stream = new ByteArrayOutputStream();
    ImgTavyIlinca.compress(Bitmap.CompressFormat.PNG, 100, stream);
    byte[] byteArray = stream.toByteArray();
    return byteArray;
}
```

Hapi i radhës është krijimi i një instance të objektit `Person` me qëllim të marrjes së shenjave të gishtërinjve të selektuar. Për këtë qëllim ne kemi definuar një funksion special, `getPerson()`, i cili merr si parametra një vlerë të tipit `int` (që përdoret më pas për të caktuar id) dhe një parametër të tipit `byte[][]`, i cili përfaqëson shabllonin e shenjave të gishtërinjve.

Vlera që marrim është e tipit `Person`. Logjika e kësaj metode është të krijojë një objekt të tipit `Fingerprint`, duke caktuar shabllonin përmes `setTemplate()`.

I tërë funksioni që përshkruan procesin e krijimit të një personi dhe caktimin e shenjave të gishtërinjve jepet më poshtë.

```
private Person getPerson(int id, byte[][] template) {
    Fingerprint Fplist[] = new Fingerprint[template.length];
    for(int i=0; i<template.length; i++){
        Fplist[i] = new Fingerprint();
    }
}
```

```

FPList[i].setTemplate(template[i]);
}

Person ps=new Person(FPList);

ps.setId(id);

return ps;
}

```

Metoda e përdorur për krahasimin e shenjave të gishtërinjve është verify() e ofruar nga libraria e përdorur në projekt. Kjo sjell një vlerë në përqindje, e cila përfaqëson shkallën e ngjashmërisë mes shenjave të gishtërinjve të dy personave. Metoda onCreate() që implementohet për këtë proces jepet në vijim:

```

protected void onCreate(Bundle savedInstanceState) {
    super.onCreate(savedInstanceState);
    setContentView(R.layout.tavy_layout);
    lblMsg = (TextView) findViewById(R.id.txt1);
    Bundle extras = getIntent().getExtras();
    myList = extras.getParcelableArrayList("imagebitmap");
    myFirstImage = (ImageView) findViewById(R.id.image1);
    mySecondImage = (ImageView) findViewById(R.id.image2);
    myFirstImage.setImageBitmap(myList.get(0));
    mySecondImage.setImageBitmap(myList.get(1));
    byte[][] i =new byte[2][];
    i[0] = this.convertBitmapToByteArray(myList.get(0));
    i[1] = this.convertBitmapToByteArray(myList.get(1));
    AfisEngineafis = new AfisEngine();
    afis.setThreshold(12);
    //getting fingerprints
    Person p1 = this.getPerson(-1,new byte[][]{i[0]});
}

```

```

Person p2 = this.getPerson(-2,new byte[][]{i[1]});

float matches = afis.verify(p1, p2);

if(afis.getThreshold() < matches)

{

lblMsg.setTextColor(Color.GREEN);

lblMsg.setText("Fingerprints matched");

}

Else

{

lblMsg.setText("Fingerprints do not match");

lblMsg.setTextColor(Color.RED);

} }

```

Nga pjesa e zhvillimit të aplikacionit kur bëhet kërkimi i personave në situata të emergjencave dhe fatkeqësive natyrore, pra kërkimi përmes shenjave të gishtërinjve, siç paraqitet edhe në figurën 6.14.

Figura 6.14 Kërkimi sipas shenjave të gishtërinjve në databazën SQLite_QOAME



Ky rezultat shfaqet sa herë që përdoruesi nis aktivitetin e përputhjes nga aplikacioni i propozuar prej nesh, duke krahasuar shenjën e gishtave fillestare me atë të gjendur në databazë.

Aplikacioni ynë ka avantazhin e përdorimit të një pajisje mobile të thjeshtë, i cili përfshin një kamerë dhe sistem operativ Android. Shpejtësia e procesimit varet kryesisht nga dy parametra: shpejtësia e procesorit të telefonit dhe madhësia e databazës.

Një përdorim është identifikimi i shpejtë i njerëzve që hasen në zonat e prekura nga fatkeqësitë natyrore, në identifikimin e atyre personave që nuk posedojnë ID Card apo diç tjetër për identifikim, duke përdorur identifikimin përmes shenjave të gishtërinjve.

Gjithashtu, aplikacioni mund të përdoret në sistemet e aksesit në institucionet publike, si: shkolla, biblioteka, spitale etj. (Octavian, 2016)

6.9 Databaza SQLite_QOAME, rasti i aplikacionit android për ISK

- Kërkimi i të dhënave nga databaza SQLite dhe pasqyrimi i tyre në pajisjen mobile android.

Instituti sizmik ka një numër të madh të stacioneve matëse në territorin e Kosovës, por që, njëherësh, këto të dhëna raportohen në këtë institut në baza ditore, javore dhe mujore. Por, me ndërtimin e databazës SQLite_QOAME këto të dhëna raportohen në kohë më të shkurtër ose më mirë të themi në kohë reale nga të gjitha stacionet. Në vijim paraqesim krijimin e tabelës së Institutit Sizmologjik të Kosovës:

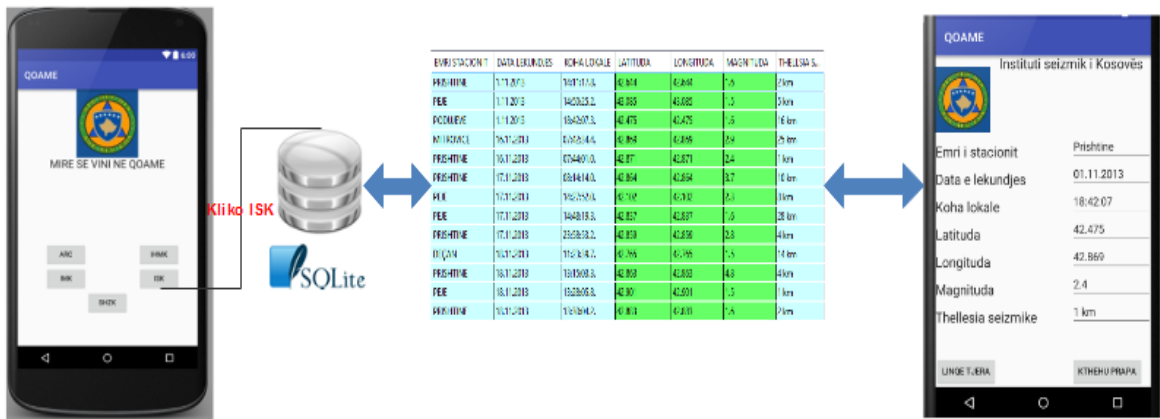
```
public class DatabaseHelper extends SQLiteOpenHelper {  
    private static final int DATABASE_VERSION = 1;  
    private static final String DATABASE_NAME = "SQLite_QOAME";  
    private static final String TABLE_NAME = " Instituti seizmik i Kosoves ";
```

```
private static final String KEY_STAC_MATES = "Emri I stacionit";
private static final String COL_1 = "Data e lekundjes ";
private static final String COL_2 = " Koha lokale";
private static final String COL_3 = "Latituda";
private static final String COL_4 = "Longituda";
private static final String COL_5 = "Mangituda";
private static final String COL_6 = "Thellesia seizmike";

public DatabaseHelper(Context context) {
    super(context, DATABASE_NAME, null, DATABASE_VERSION);
}
}
```

Duke klikuar në butonin ISK në aplikacionin bazik të QOAME-së, kemi qasje në kërkimin e të dhënave më të reja nga ISK. Në vijim paraqitet në figurën 6.15:

Figura 6.15 Qasja në të dhënat e Institutit Sizmologjik të Kosovës



Pasi që është krijuar tabela Instituti Sizmologjik i Kosovës në kuadër të databazës SQLite_QOAME, duke klikuar mbi butonin ISK kërkojmë të dhënat e freskuara për Institutin Sizmologjik të Kosovës, duke përfshirë magnitudën e lëkundjes sizmike, thellësinë sizmike, pozicionin gjeografik latitude-longitude, data dhe koha e lëkundjes sizmike.

Qasja në linjë web.

Fatkeqësitë natyrore nuk janë të përkufizuara në një domain të caktuar gjeografik, por ato përfshijnë një pjesë të madhe gjeografike. Duke u nisur nga ky konstatim, atëherë edhe shteti

i Kosovës, patjetër, duhet të jete pjesë integrale e të dhënave të lëkundjeve sizmike në nivel regional dhe atë kontinental. Në aplikacionin tonë të ndërtuar në android ua kemi bërë të mundur punonjësve të AME-së ta përdorin aplikacionin QOAME, më pas duke klikuar butonin LINQET TJERA, me këtë paraqitet qasja në linkun

```
Uri uri = Uri.parse("http://www.emsc-csem.org/Earthquake");

Intent intent = new Intent(Intent.ACTION_VIEW, uri);

startActivity(intent);

    }

}
```

që lidhet me web faqet më kredibile të monitorimit të të dhënave nga fusha sizmike në nivel botëror, duke u bërë pjesë integrale e informatave. Kjo është ndërtuar dhe paraqitet në figurën 6.16:

Figura 6.16 Aplikacioni QOAME, qasja në linqet për informacione sizmike



6.10 Databaza SQLite_QOAME, rasti i aplikacionit android për IHMK

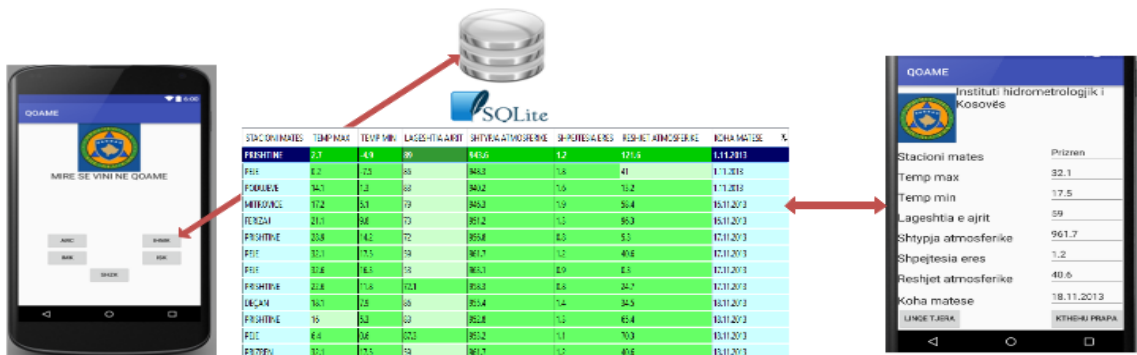
- Kërkimi i të dhënave nga databaza SQLite dhe pasqyrimi i tyre në pajisjen mobile android.

Instituti Hidrometeorologjik ka një numër të madh të stacioneve matëse në territorin e Kosovës, por që, njëherësh, këto të dhëna raportohen në këtë institut hidrometeorologjik në baza ditore, javore dhe mujore.

Por, me ndërtimin e databazes SQLite_QOAME këto të dhëna raportohen në kohë më të shkurtë ose në kohë reale nga të gjitha stacionet, duke përfshirë edhe të dhënat nga rrjedhat e lumenjve, të dhënat nga sensorët e vendosur në vendet ku ka reshje të mëdha të borës. Në vijim paraqesim krijimin e tabelës Instituti Hidrometeorologjik i Kosovës:

```
public class DatabaseHelper extends SQLiteOpenHelper {  
  
    private static final int DATABASE_VERSION = 1;  
  
    private static final String DATABASE_NAME = "SQLite_QOAME";  
  
    private static final String TABLE_NAME = "Instituti hidrometrologjik  
i Kosoves ";  
  
    private static final String KEY_STAC_MATES = "Stacioni mates";  
    private static final String COL_1 = "Temp.max";  
    private static final String COL_2 = "Temp.min";  
    private static final String COL_3 = "Lageshtia e ajrit";  
    private static final String COL_4 = "Shtypja atmosferike";  
    private static final String COL_5 = "Shpejtesia e eres";  
    private static final String COL_6 = "Reshjet atmosferike";  
    private static final String COL_7 = "Koha matese";  
    public DatabaseHelper(Context context) {  
        super(context, DATABASE_NAME, null, DATABASE_VERSION);  
    }  
}
```

Figura 6.17 Të dhënat nga databaza SQLite për rastin IHMK



Nga ky aplikacion kemi arritur të marrim informacione të detajuara nga të gjitha stacionet matëse brenda territorit të Kosovës, duke pasur një pasqyrim të paralajmërimeve në rast të vërshimeve nga reshjet atmosferike, mbi shpejtësinë e erës, lagështia e ajrit dhe temperaturave minimale dhe maksimale.

Qasja në linjë web.

Duke vënë në dukje se Qendra operative në kuadër të Agjencisë së Menaxhimit të Emergjencave nuk ka mundësi të operojë edhe me vendet e rajonit për marrjen e informacioneve lidhur me reshjet, temperaturat, prurjet e lumenjve etj., atëherë është e nevojshme që në aplikacionin e QOAME-së të ketë mundësi të marrim edhe informata shtesë, duke përdorur butonin LINQET TJERA, e me këtë paraqitet qasja në linkun

```
Uri uri = Uri.parse("http://www.accuweather.com");
```

```
Intent intent = new Intent(Intent.ACTION_VIEW, uri);
```

```
startActivity(intent);} }
```

aplikacioni tregon në ekran me informacion mbi motin në lokacionin e përzgjedhur, si në figurën 6.18:

Figura 6.18 Aplikacioni QOAME, qasja në linqet tjerë për informim të mëtutjeshëm.



Kjo kërkon vetëm pak sekonda për t'u ngarkuar që nga marrja e vendndodhjes dhe pastaj jep informacione për atë vend. Aplikacioni e jep informacionin në format teksti. Ai tregon temperaturën minimale dhe maksimale, rënia e shiut, koha e lindjes dhe perëndimit të diellit (TANYA, 2014).

6.11 Konkluzion

Teknologjia mobile ka treguar të ketë një impakt të konsiderueshëm në shoqëri viteve të fundit. Gradualisht, kjo teknologji është bërë pjesë e jetës së përditshme dhe mënyra më e lehtë për të shpërndarë informacion. Zhvillimi i aplikacioneve të projektit doktoral për QOAME vjen si rezultat i ndërtimit të një arkitekture në platformën android, më pas zhvillimin e një modeli profesional në Talend open studio, që mundëson integrimin e të dhënave nga shumë burime të ndryshme të të dhënave. Më pas zhvillimin e aplikacioneve në platformën android me gjuhën programuese në objekte JAVA, ku realizuam që të dhënat nga këto burime t'i afrohen limitit të paraqitjes në kohë reale.

Konkluzione dhe rekomandime

Konkluzione

Qëllimi i këtij punimi doktoral është për të dhënë një zgjidhje me përdorimin e teknologjisë së informacionit, për rastin e ndërtimit të një qendre operative në kuadër të AME-së. Për zhvillimin e aplikacionit QOAME në pajisje mobile, që ndihmon në menaxhim të fatkeqësive natyrore.

Andaj, fillimi i punimit doktoral u realizua në ngritjen e hipotezës, si : *“Nevoja e ekzistencës së një informacioni të centralizuar në kohë reale lidhur me situatat e emergjencës dhe të fatkeqësive natyrore në Republikën e Kosovës. Kjo do të realizohej me krijimin e një qendre që do të ndihmonte në menaxhimin e emergjencave dhe të fatkeqësive natyrore, parashikimin dhe parandalimin e mundshëm të tyre. Kjo do bëjë të mundur sinkronizimin e dhënave në aplikacionet e pajisjeve mobile. Rasti studimor: Ndërtimi i qendrës së emergjencave dhe të fatkeqësive natyrore në QOAME”*.

Kështu, për vërtetimin e hipotezës shfrytëzuam dy nga objektivat studimore dhe shkencore, bashkangjitur edhe disa pyetje kërkimore të tjera, të cilat i shqyrtuam, i analizuam dhe në fund zhvilluam hap pas hapi aplikacionin për pajisjet mobile.

Nisur nga shtrimi i pyetjes së parë kërkimore që ka të bëjë me gjendjen e trajtimit të informacionit në situatat e emergjencave dhe fatkeqësive natyrore në Kosovë, është karakteristike të theksojmë se aktualisht Agjencia e Menaxhimit të Emergjencave dhe parandalimi i fatkeqësive natyrore posedon të dhëna nga ISK, që bëhen mbi baza mujore, e që, në fakt, nevoja për informacione është e nevojshme të jenë në kohë reale. Ndërsa, te Instituti Hidrometeorologjik, të dhënat ruhen në baza ditore-javore dhe mujore, ku të gjitha

variablat e motit, të reshjeve atmosferike dhe të vërshimeve duhet të monitorohen vazhdimisht dhe të raportohen nga qendra te QOAME. Andaj, ndër praktikat më të mira sa i përket menaxhimit të përmbajtjeve konfirmohen brenda dy parimeve thelbësore:

- Të dhënat hidrometeorologjike duhet që në mënyrë konstante të monitorohen, në mënyrë që të zbulohet fillimi i një situatë të thatësirës;
- Duhet të ekzistojnë nivele të qarta të alarmimit (varësisht nga statusi i të dhënave hidrometeorologjike) në bazë të të cilave mund të ndërmerren veprime të ndryshme për zbutjen e kësaj situatë.

Por, çka është parashtruar në pyetjen e dytë, që ka të bëjë me ndërtimin e një sistemi komunikimi në mes qendrave operative në Agjencinë e Menaxhimit të Emergjencave (QOAME)? Në punimin tim doktoral kam vënë synim integrimin e sistemit, duke përdor modelin e ndërtimit të datawarehouse për QOAME-në. Më pas, duke argumentuar në nivel literature shkencore dy nga metodat për ndërtimin e DWH-së.

- *Metodika Top-down (Bill Imnon)*
- *Metodika Bottom-Up (Ralph Kimball)*

Jam fokusuar në metodikën Bottom-Up të Kimball-it. Pra, burimet e informacionit që kam arritur nëpërmjet ndërtimit të sistemit DWH t'i integroj, janë:

- Agjencia e Regjistrimit Civil (ARC).
- Instituti Sizmologjik i Kosovës (ISK).
- Instituti i Mjekësisë së Kosovës (IMK).

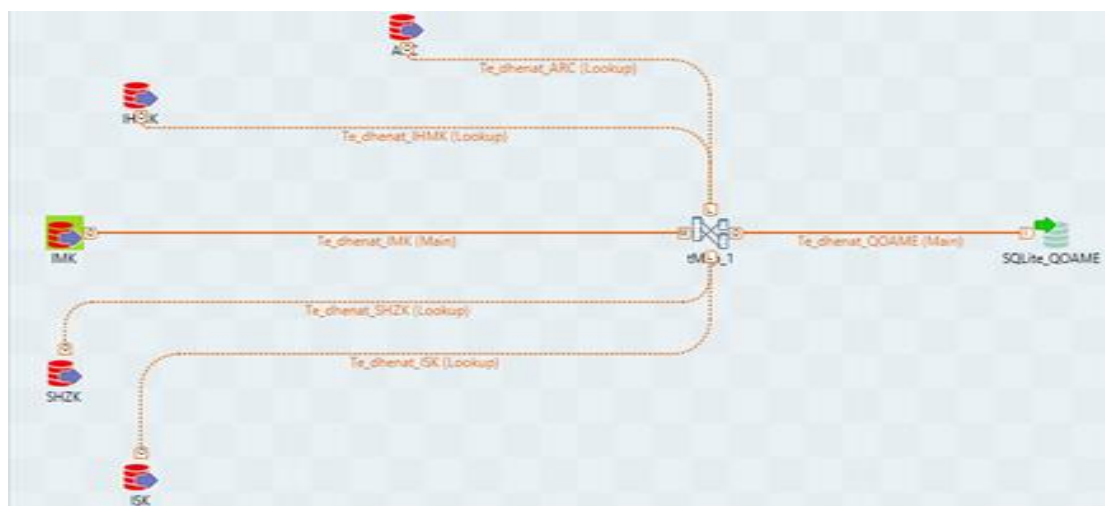
- Instituti Hidrometeorologjik i Kosovës (IHMK).

- Shërbimi i Zjarrfikësve të Kosovës (SHZK).

Në vijim i kam trajtuar edhe versionet e mundshme për të realizuar ndërtimin e një qendre të tillë. Qendra operative e agjencisë për menaxhimin e emergjencave dhe të fatkeqësive natyrore e ka të domosdoshme të grumbullojë dhe analizojë të dhëna masive të fatkeqësive natyrore. Nga versionet që janë shfrytëzuar në këtë punim doktorat është softueri Talend open studio and Big Data, Talend Open Studio përdoret, kryesisht, për integrim mes sistemeve operative, si dhe për ETL (Extract, Transform, Load) për Business Intelligence dhe Datawarehouse, por edhe për migrim, kështu që modeli shumëdimensional u përzgjedh për t'u aplikuar në rastin tonë.

Nga të gjithë data mart-ët e krijuar është arritur që përmes Talend open studio and Big Data, të bëjmë integrimin dhe procesin e ETL-së (Extract, Transform, Load), duke vërtetuar edhe përmes figurës së mëposhtme:

Figura: Integrimi i data mart-ëve ne SQLite_QOAME



Kjo vërteton një segment të hipotezës dhe pyetjeve kërkimore për realizimin e një qendre që integrohen të dhënat dhe pastaj bëhet modelimi i data mart-ëve në arritjen e ekstratimit të të dhënave, transformimin dhe nxjerrjen e të dhënave nga data mart-ët: IHMK, IMK, SHZK, ARC dhe ISK në databazën SQLite_QOAME. Me këtë rast kemi arritur që të grumbullojmë sasi të mëdha të dhënash nga të gjitha këto burime, që më pas t'i integrojmë në një të vetëm.

Në punimin doktoral i është dhënë një rëndësi e madhe edhe sinkronizimit të të dhënave në komunikimin e sistemeve në rastin e emergjencave dhe fatkeqësive natyrore në Kosovë.

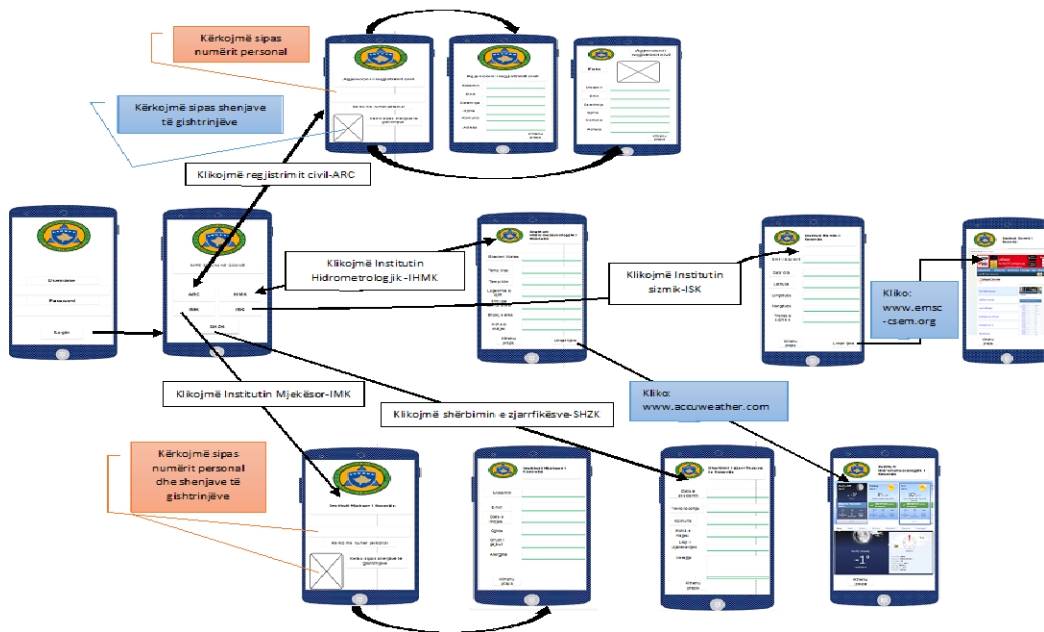
Lidhur me faktin se Agjencia e Menaxhimit të Emergjencave posedon me sasi të mëdha të dhënash, të cilat mund t'i arrijnë me lehtësi shumë megabyte ose gigabyte në ditë, këto fakte tregojnë se është shumë e vështirë të përditësohen në mënyrë të efektshme. Andaj, me këtë rast kemi bërë përpjekje që nëpërmjet shpjegimit teorik, të bazuar në literaturë dhe fakte shkencore, të aplikojmë sinkronizimin e të dhënave në kohë reale.

DWH në kohë reale përbën një numër unik sfidash dhe mundësish për inxhinierin ETL. Në rastin tonë do të përdorim metodat ETL. Proceset ETL-së përbejnë në vetvete zgjidhje software-ike, të cilat janë përgjegjëse për tërheqjen e të dhënave nga burime të ndryshme, për pastrimin e tyre, përshtatjen dhe ngarkimin në databaza.

Njëra nga konkluzionet e fundit që vërtetojnë hipotezën e ngritur dhe japin përgjigje në njërën nga pyetjet kërkimore është roli i ndikimit të aplikimeve mobile në rastin e emergjencave në Kosovë, me aplikimin e përdorimit të platformës Android dhe gjuhës programuese Java në zhvillimin e aplikacionit për Qendrën operative të AME-së dhe fatkeqësive natyrore. Me

zhvillimin e këtij projekti, siç është QOAME në platformën android, bëhet e mundur që të rritet performanca në marrjen e vendimeve për menaxhim të emergjencave dhe fatkeqësive natyrore, sipas figurës së paraqitur në një formë hap pas hapi të zhvillimit të aplikacioneve për këto institucione që janë integruar në një të vetëm, siç është AME-ja.

Figura: Paraqitja hap pas hapi e zhvillimit të aplikacionit QOAME



Aplikacioni QOAME fillimisht është i siguar me username dhe password, më pas është i ndarë në 5 butone, siç janë: ISK, ARC, IHMK, IMK dhe SHZK. Me zhvillimin e këtij aplikacioni që, në fakt, bën të mundur qasjen në informacione në kohë reale, si në rastin e Agjencisë së Regjistrimit Civil, ku janë dy mundësi për identifikimin e viktimave të fatkeqësive natyrore. Duke shfrytëzuar identifikimin nëpërmjet numrit personal, në mungesë të ID, identifikimi mund të bëhet edhe përmes shenjave të gishtërinjve. Pjesë e aplikacionit QOAME është edhe qasja në institutin sizmologjik, institutin mjekësor, institutin

hidrometeorologjik dhe në shërbimin e zjarrfikësve. Në të gjitha këto institucione informacionet janë të përditësuara. Përveç kësaj, në aplikacionet e QOAME-së është zhvilluar edhe mundësia e qasjes në ueb-faqe të rëndësishme për monitorim të kushteve klimatike dhe atyre sizmike në nivel botëror.

Rekomandime

Pas procesit të vërtetimit të pyetjeve kërkimore dhe hipotezës së ngritur në këtë punim doktoral, lind nevoja që nëpërmjet këtij studimi shkencor të jepen edhe ca rekomandime. Që, në fakt, janë të nevojshme jo vetëm për rastin tonë studimor në menaxhim të emergjencave dhe fatkeqësive natyrore, por edhe për institucionet e tjera, si: financiare, biznesore, arsimore dhe shkencore.

Rekomandimi 1: Agjencia e Menaxhimit të Emergjencave, si institucion përgjegjës në paralajmërimin dhe parandalimin e fatkeqësive natyrore dhe emergjencave civile, ka nevojë shumë të shpejtë që brenda qendrës operative të bëjë integrimin e të dhënave nga institucionet përgjegjëse, si: instituti sizmologjik, instituti hidrometeorologjik, shërbimi i zjarrfikësve, instituti mjekësor dhe regjistrimi civil i Kosovës. Me këtë rast duhet krijuar një datawarehouse–Big data, që bën të mundur integrimin e të gjitha këtyre burimeve të informacionit.

Rekomandimi 2: Japim rekomandim që AME-ja pas krijimit të datawarehouse-it të bëjë modelimin e këtyre data mart-ëve, sipas modelit Kimball, që do të shërbejë për analiza të hollësishme të institutit mjekësor, të të dhënave nga instituti hidrometeorologjik, si dhe nga institute sizmike. Në grumbullimin e këtyre të dhënave, transformimin dhe ngarkimin e të

dhënave, gjejnë aplikim metodat ETL, që, në fakt, është njëra nga temat me delikate në fushën e komunikimit të sistemeve .

Rekomandimi 3: Pra, procesi ETL ekstrakton të dhënat nga sistemet burim, i transformon të dhënat sipas rregullave të institucionit dhe i ngarkon rezultatet në DWH të synuar. Studimi ka zbuluar se 70% e përpjekjeve për implementimin e softuerëve dhe të mirëmbajtjes së DWH harxhohet në ndërtimin e sistemit ETL. Pra, edhe në punimin doktoral kemi shfrytëzuar softuerin Talend open studio and Big data, që mundëson integrimin e të dhënave nga shumë burime të ndryshme, duke dhënë një shembull pozitiv në integrim dhe modelim të të dhënave.

Rekomandimi 4: Është nevojë më se e domosdoshme që të merret parasysh rekomandimi vijues lidhur me procesin e ETL-së që bën integrimin dhe sinkronizimin e të dhënave. Me këtë rast AME-ja duhet të integrojë përditësimin e të dhënave në kohë, kështu që segmentet e DWH-së përditësohen në intervale të ndryshme në varësi të burimit të informacionit në mjedisin e përpunimit të transaksionit online.

Në punimin doktoral kemi bërë zgjidhje duke shfrytëzuar softuerin Talend open studio and Big data, me këtë kemi integruar data mart-të në databazën SQLite_QOAME, ku kemi arritur që sinkronizimi i të dhënave t'i afrohet gjendjes në kohë reale.

Rekomandimi 5: Andaj, si pjesë e rekomandimit për AME-në është përdorimi dhe zhvillimi i aplikacioneve të projektit doktoral për QOAME, që vjen si rezultat i ndërtimit të një arkitekture në platformën android. Dihet qartë se periudha e fatkeqësive natyrore dhe emergjencave nuk është një proces kur të gjitha informacionet janë në dispozicion, por në këtë periudhë çdo gjë është e rëndësishme e mbi të gjitha informacioni lidhur me faktorin

njeri. Pra, AME-ja duhet t'i këtë parasysh këto lloje situatash dhe t'i paraprijë ngritjes së kapaciteteve në komunikimin e sistemeve, në zhvillimin e aplikacioneve për menaxhim të rasteve emergjente dhe fatkeqësive natyrore.

LISTA E REFERENCAVE/BIBLIOGRAFIA

Libra dhe disertacione

Alejandro.V, Esteban. Z (2014): Data Warehouse Systems, Design and Implementation; Springer-Verlag Berlin.

Andreas M, Thomas L, Thomas R, Thomas K, Holger K.(2002): Design Challenges for an Integrated Disaster Management Communication and Information System. New York City

Armen G,Hachim B,George B (2011): Raport: Vlerësimi i Kapacitetit për Zvogëlim të Rrezikut nga Fatkeqësitë për Kosovë, Prishtinë, Kosovë.

Arwa A J (2011): Transitioning a Clinical Unit to a Datawarehouse, Guildford, England.

Çela.S (2014); Studimi i antenave inteligjente të telefonisë celulare, rrezatimi jo-jonizues në afërsi të tyre, Tiranë, Shqipëri.

Dasho.A (2014): Zhvillimi i aplikimeve të platformës web një risi e re për korporatën elektroenergjetike, Tiranë, Shqipëri.

Golfarelli.M, Rizzi. S (2009): Datawarehouse Design: Modern Principles and Methodologies; McGraw-Hill, Itali.

Hoffer. J. A., Ramesh.V, Topi.H (2011): Modern database management- Tenth Edition; Pearson, New Jersey.

Kimball.R, Ross.M (2002): The Datawarehouse Toolkit - The Complete Guide to Dimensional Modeling, John Wiley and Sons, Inc, New York.

Kimball.R, Ross.M, Thornthwaite.W, Mundy.J, Becker.B (2008): The Datawarehouse Lifecycle Toolkit, Second Edition, John Wiley and Sons, Inc, New York.

Khan, I (2012): Modern methodology & tools for data warehouse development; Athabasca University.

Kimball.R, Caserta.J (2004): The Datawarehouse ETL Toolkit- Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data, Wiley Publishing, USA.

Lito,G (2013): Menaxhimi i riskut të katastrofave-rasti i Shqipërisë, Tiranë.

Juliana L (2014): Studimi i sistemeve datawarehouse dhe ndërtimi i një modeli raportimi të bazuar në teknologjinë business intelligence, Tiranë.

Jay. K (2010): Using SQLite, Published by O'Reilly Media, USA.

Mark, K(2010): Command Center for Disaster Management Communication System Datawarehouse, Purdue University, USA.

Malinowski E, Zimányi E (2008): Advanced Datawarehouse Design-Data-Centric Systems and Applications, Springer-Verlag, Berlin.

Nicholas .O (2012): Smartphones A Platform For Disaster Management, Amsterdam, Holande.

Ponniah. P (2010): Datawarehousing fundamentals for IT professionals, 2nd ed; John Wiley & Sons, Inc., New Jersey.

Rainardi.V (2008): Building a Datawarehouse: With Examples in SQL Server, Apress, United States of America.

Reto. M(2012): Professional Android™ 4 Application Development, John Wiley & Sons, Inc., Kanada.

Santos R, Bernardino. J (2008): Real-Time Datawarehouse Loading Methodology, ACM, Portugal.

Stolba N (2007): Towards a Sustainable Datawarehouse Approach for Evidence-Based Healthcare, Vien, Austria.

Tanya. T (2014): An Android application to support flash flood disaster response management in India, Holande.

Wallace. J (2012): Android Apps for Absolute Beginners Second Edition, Apress.

Wei-Meng. L (2012): Beginning Android™ 4 Application Development, by John Wiley & Sons, Inc., Indianapolis, Indiana.

Revista shkencore dhe burime nga interneti:

Assion. L, Lori. W, Laurence. M, Kinvi. L (2012): “Adoption of cell phones as preparedness efforts for natural disasters” , Vol. 13, f 11-20.

Ali. Sh (2011): “A proposed model for datawarehouse ETL processes”, Vol.23, f 91–104.

Camilovic, D (2009): “A Call Detail Records Data Mart: Data Modelling and OLAP Analysis”, ComSIS, Vol. 6, f 87-110.

Cristina. I, Ferreira. E. J (2006): “Synchronization options for datawarehouse designs”, IEEE Computer, Vol 39, f 53-57.

Hristidis. V, Chen. SH, Li .T, Luis. S, Deng .Y (2010): “Survey of data management and analysis in disaster situations”, Elsevier, fq. 1701-1714.

Kakish. K, Kraft. TH (2012): “ETL Evolution for Real-Time Datawarehousing”, New Orleans Louisiana, Vol 5, fq. 1-12.

Krasniqi. Z (2014): “Application of information technology in emergency and natural disaster management”, Tirana, Albania.

Lenzerini, M (2002): “Data Integration: A Theoretical Perspective”, Universit `a di Roma “La Sapienza.

Lula. J, Kika. A (2012): “Modelimi i një data mart-i për CDR, b u l e t i n i, Vol. 14, fq. 230-242.

Peng.Y, Zhang. Y, Tang. Y, Li. Sh (2010): “An incident information management framework based on data integration, data mining, and multi-criteria decision making”, Elsevier, Vol. 51, fq. 316–327.

Tanasescu. V, Gugliotta, A., Domingue, J., Davies, R., Gutierrez, L., Rowlatt, M., Richardson, M., Stincic, S. (2006): “A Semantic Web Services GIS based Emergency Management Application”, Springer , fq. 959-966.

Agjencia (2011): “Gjendja e mjedisit në Kosovë 2008-2010”, marrë nga www.ammk-rks.net/repository/docs/Gjendja_e_Mjedisit_ne_Kosove_2008-2010.pdf, (parë më 14.2.2015).

Azwa. A, Abdul. W, Nazirah. H: “Integration of heterogeneous databases in academic environment using open source ETL tools”, marrë nga [file:///C:/Users/Administrator/Downloads/dad58d3aa8aab6b692996cdf7775c7d%20\(2\).pdf](http://file:///C:/Users/Administrator/Downloads/dad58d3aa8aab6b692996cdf7775c7d%20(2).pdf) (parë më 4.2.2015).

Aditya. A (2015): “Android Application Security Part 2-Understanding Android Operating System”, marrë nga www.manifestsecurity.com/android-application-security-part-2 (parë më 1.2.2015).

Ambati. K (2012): “DataStage Enterprise Edition”, marrë nga www.scribd.com/doc/79600772/DataStage-OA#scribd, (parë më 15.9.2015).

Aplikime mobile, marrë nga www.mobile.geo.edu.al/wp-content/uploads//Platforma-Android.pdf (parë më 1.2.2015).

CIO.al (2011): “Teknologjia OLAP-revolucioni i parë në analizën e të dhënave” marrë nga www.cio.al/teknologjia-olap-revolucioni-i-pare-ne-analizen-e-te-dhenave/#sthash.o5ëgftdL.dpuf, (parë më 10.10.2014).

ETL(2015):“Extract_transform_load“marrë nga www.en.wikipedia.org/wiki/Extract,_transform,_load. (parë më 12.5.2015).

Efficient and Real-time Data Integration (2009): “Efficient and Real Time Data Integration With Change Data Capture” ,marrë nga www.download.101com.com/tdwi/ww29/Attunity_Efficient_and_Real-Time_DI.pdf , (parë më 15.4.2015).

Daniel. L, Andre. V (2011):”**Extracting, Transforming and Archiving Scientific Data**”, marrë nga <http://arxiv.org/pdf/1108.4041.pdf> (parë më 12.05.2015).

Faulkner. B (2011): “Korniza kosovare për menaxhim të rrezikut nga përmytjet”, marrë nga www.kryeministriks.net/tfu/repository/docs/110421_Kosovo_Flood_Management_Framework_Alb.pdf (parë më 15.12. 2014).

Forumi për siguri (2013): “Reagimet në situata emergjente – përgatitjet dhe mundësitë”, marrë nga www.fiq-fci.org/repository/docs/Reagimet_ne_situata_emergjente.pdf, Prishtinë. (parë më 1.2.2014)

Garg.N, Mithal.S: “Spatial Datawarehouses”, marrë nga www.users.cs.umn.edu/~smithal/ (parë më 04.8.2015).

Instituti Hidrometeorologjik (2014): “Instituti Hidrometeorologjik i Kosovës”, marrë nga www.ammk-rks.net/?page=1,90, (parë më 17.1.2014).

Instituti Sizmik (2012): “Mikrozonimi sizmik i Gjilanit”, marrë nga www.mzhe-ks.net/repository/docs/mikrozonimi_sizmik_i_gjilanit.pdf, (parë më 2.2.2014).

Johnson.E, Jones.J (2015): “Building ETL Processes for Business Intelligence Solutions Built on Microsoft SQL Server”, marrë nga www.erwin.com/content/whitepapers/ca_erwin_building_etl_processes_sql.pdf (parë më 12.3.2015).

Johnson .C (2013):”Surrogate Key in Datawarehouse, What, When and Why” marrë nga <http://www.disoln.org/2013/11/Surrogate-Key-in-Data-Warehouse-What-When-Why-and-Why-Not.html>], (parë më 18.02.2015).

Kristiina, R.(2004): “The Role of Information Technology in Crisis Management”, marrë nga www.einiras.net/conf/conferences/conferences_helsinki2004.cfm, (parë më 14 janar 2015).

MPB (2010): “Plani i Reagimit Kombëtar-Prishtinë”, marrë nga www.mpb-ks.org/repository/docs/Plani%20Reagimit%20Kombetar_14_01_11.pdf, (parë më 18.10.2014).

MSH (2014): “Strategjia për sistemin e informimit shëndetësor në Kosovë 2010-2020”, marrë nga www.kryeministriks.net/repository/docs/strategjia_per_sistemin_e_informimit_shendetesor_ne_kosove_2010_-_2020.pdf , (parë më 1.5.2015).

Muhammad .A, Tripathi .N(2012): A Web-based Disaster Management-Mitigation Framework Using Information and Communication Technologies and Open Source Software; JU Journal of Information Technology (JIT), Vol. 1, June 2012.

Mesime Java:marrë nga www.mesime24.com/v2/index.php/meso-ne/programim/item/122-hyrje-ne-java (parë më 4.3.2015).

Mattias. S (2014): “Development of an Open-Source Mobile Application for Emergency Data Collection”, marrë nga <http://lup.lub.lu.se/luur/download?func=downloadFile&recordId=4252143&fileId=4252157> (parë më 03.03.2014).

Octavian. D, Ilinca. L(2016): “The Recognition of Fingerprints on Mobile Applications – an Android Case Study” , marrë nga <http://www.ibimapublishing.com/journals/JEERBE/2016/813264/813264.html> (parë më 03.03.2016).

Plani i mbrojtjes (2011): “Plani i mbrojtjes nga zjarri për komunën e Prishtinës”, marrë nga www.kk.rks.gov.net/prishtina/getattachment/Municipality/Departments/Sherb-Pub-Mbrojtje-dhe-Shpet-/Sektori-per-mbrojtje-dhe-shpetim/plani-i-mz-prill2011-.pdf.aspx, (parë më 14.7.2014).

Robert. N (2009): “Data Integration with Talend Open Studio”, marrë nga http://www.solver.com/files/_document/Talend.pdf (parë më 13.7.2014).

Sam. B, Lance. M (2009): “Geospatial Web Services and Cross-Boundary Information Sharing During Disasters”, marrë nga www.ogcnetwork.net/node/573 (parë më 24.5.2014).

SIME (2010): “Sistemi i Integruar i Menaxhimit të Emergjencave”, marrë nga www.mpb-ks.org/repository/docs/Sistemi%20i%20integruar%20per%20menaxhimin%20e%20emergjencave%20ALB.pdf (parë më 1.2.2014).

Sriganesh, V., Iyer, P.: “Information Resources and Systems for Disaster Management”, marrë nga www.qmed.org.in/wp-content/uploads/2013/02/pi_info_sys_disastermgnt.pdf (parë më 15 mars 2015).

Sunguk .L (2012): “Creating and Using Databases for Android Applications”, International Journal of Database Theory and Application Vol.5, No.2.marrë nga www.sersc.org/journals/IJDTA/vol5_no2/8.pdf (parë më 1.1.2015).

Talend (2015): ”Datawarehouse” marrë nga www.datawarehouse4u.info/Talend-Open-Studio.html (parë më 2.2.2015).

Zyra e Kryeministrit (2014): “Grupi i punës për menaxhimin e thatësirës. Mësimet e nxjerra dhe rekomandimet”, marrë nga www.kryeministriks.net/tfu/repository/docs/2014_Thatesiat_mesimet_e_nxjerra_rekomandimet.pdf, (parë më 2/9/2015).

WFP (2014):”Using mobile phone activity for disaster management during floods” , marrë nga http://unglobalpulse.org/sites/default/files/Mobile_flooding_WFP_Final.pdf (parë më 2/10/2015).