

System Design Document

Author: Hamad Almarri (SID 200294236)
University Of Regina
Final Year Project (2014-2015)
Project Name: Adala
Supervisor: Abdul Bais

1. Introduction

1.1 Purpose of the system

The purpose of the system is to facilitate law enforcement processes (Police and Investigation) to let them manage and monitor different police cases. It will have also a database of people, cars, addresses, and properties, and it will support work-flow such as open/foreword cases to be special cases for investigation purposes, which will include work-flow actions such as approve and deny processes. Police cases can be incident/accident reports, or arrest reports. In brief, the software will be a tool where police and investigators/detectives cooperate to solve crimes or any investigative cases. The system is web based, and it will have monitoring and reporting functionalities for police cases, and will provide a list of suspect criminals for a new crime based on the data that is available on the system's database. The software also should be able to apply crime prevention for youth at risk by collected and analyzing history data of a young person and the software will be able to classify that person if he/she has high probability of being a criminal in the future.

1.2 Design goals

Adala is consisting of four major parts: the basic Create Read Update Delete (CRUD) operation, case management, and the system intelligence part where it is providing list of suspect people for each case and providing list of young people in risk of being criminal in future by using Machine Learning algorithms. The last part is the software security, which is consists of user login operations, authentication, authorization, and administrator management to manage users' accesses and roles.

1.3 Definitions, acronyms, and abbreviations

ANN	Artificial Neural Network
Machine Learning	Is a scientific discipline that explores the construction and study of algorithms that can learn from data. Such algorithms operate by building a model based on inputs: and using that to make predictions or decisions, rather than

	following only explicitly programmed instructions. (Wikipedia)
CRUD	Are the four basic functions of persistent storage. i.e. create, read, update and delete. (Wikipedia)
Workflow	Consists of an orchestrated and repeatable pattern of business activity enabled by the systematic organization of resources into processes that transform materials, provide services, or process information. It can be depicted as a sequence of operations, declared as work of a person or group, an organization of staff, or one or more simple or complex mechanisms. (Wikipedia)
Law enforcement	Broadly refers to any system by which some members of society act in an organized manner to enforce the law by discovering, deterring, rehabilitating or punishing persons who violate the rules and norms governing that society. Although the term may encompass entities such as courts and prisons, it is most frequently applied to those who directly engage in patrols or surveillance to dissuade and discover criminal activity, and those who investigate crimes and apprehend offenders. (Wikipedia)
Investigative Case	Is the case that is under investigation where officers and investigators collaborate to solve the crime.
Data Mining	Is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. (Wikipedia)
JAAS	Java Authentication and Authorization Service. Java security technology includes a large set of APIs, tools, and

	implementations of commonly used security algorithms, mechanisms, and protocols. The Java security APIs span a wide range of areas, including cryptography, public key infrastructure, secure communication, authentication, and access control. (oracle.com)
JSF	<p>JavaServer Faces is a Java specification for building component-based user interfaces for web applications and exposing them as server side. (Wikipedia)</p> <p>Can be considered as view in MVC.</p>
JPA	<p>Java Persistence API is a Java programming language application programming interface specification that describes the management of relational data in applications using Java Platform, Standard Edition and Java Platform, Enterprise Edition. Persistence in this context covers three areas:</p> <ol style="list-style-type: none"> 1- the API itself, defined in the javax.persistence package 2- the Java Persistence Query Language (JPQL) 3- object/relational metadata. <p>(Wikipedia)</p>
EclipseLink	<p>EclipseLink is the open source Eclipse Persistence Services Project from the Eclipse Foundation. The software provides an extensible framework that allows Java developers to interact with various data services, including databases, web services, Object XML mapping (OXM), and Enterprise Information Systems (EIS). EclipseLink supports a number of persistence standards including:</p> <p>Java Persistence API (JPA) Java Architecture for XML Binding (JAXB) Java Connector Architecture (JCA) Service Data Objects (SDO). (Wikipedia)</p>
GlassFish	GlassFish is the reference

	implementation of Java EE and as such supports Enterprise JavaBeans, JPA, JavaServer Faces, JMS, RMI, JavaServer Pages, servlets, etc. This allows developers to create enterprise applications that are portable and scalable, and that integrate with legacy technologies. (Wikipedia)
MVC	Model-view-controller is a software architectural pattern for implementing user interfaces. It divides a given software application into three interconnected parts, so as to separate internal representations of information from the ways that information is presented to or accepted from the user. (Wikipedia)
Managed Beans	Are Java classes that can be used to inject data into a JSF files. They can be considered as Controllers in MVC.
EJB	Enterprise JavaBeans is a managed, server-side component architecture for modular construction of enterprise applications, and a server-side model that encapsulates the business logic of an application. (Wikipedia)
Entity Bean	Is a type of Enterprise JavaBean, a server-side Java EE component, that represents persistent data maintained in a database. (Wikipedia)
MySQL	Is a relational database management system (RDBMS). (Wikipedia)

1.4 References

- Java (<https://java.com/en/download/>)
- Java EE (<http://www.oracle.com/technetwork/java/javaee/overview/index.html>)
- Glassfish 4 (<https://glassfish.java.net>)
- EclipseLink JPA (<http://eclipse.org/eclipselink/jpa.php>)
- MySQL (<http://www.mysql.com>)
- JSF (<http://www.oracle.com/technetwork/java/javaee/jaserverfaces-139869.html>)

2. Current software architecture

There is no current software that will be replaced. Instead, a survey of current architectures for similar systems will be discussed in this section. This survey includes three existed softwares, which we will illustrate their features.

First software is [X-FIRE](#) (pronounced “crossfire”), and it is a product of Agnovi Corporation. It is law enforcement case management software, which provides case management functionality and case relationship visualization feature. It has desktop and web-based versions. This software has been included in this survey due to its simplicity and clarity of representing the data, which an example of the very basic software that meet Adala requirements in terms of CRUD and Case management parts. Agnovi has a link on their web site for requesting a demo, which can be useful to use X-FIRE instead of viewing its feature in bullet lists. Agnovi website is <http://agnovi.com/products/case-management/>.

Second on is [Column Case Investigative](#). This software has what in the previous illustrated software plus that it has approval activity, reports, case prioritization, and data security and privacy. Unfortunately, Column’s demo cannot be found, but they have support, training, and consulting services. Column’s website is columncaseinvestigative.com.

Third software is [i2 Integrated Law Enforcement](#). This software is built by IBM Corporation. In addition on previous shown two softwares, i2 has Build policing strategies, real time maps of cities, sending and receiving messages between users, and person threat assessment. This software doesn’t have a demo, but they post a video of using the i2 for certain scenarios. I2 considered the best among this list based on the features that it has. I2’s website is <http://www-03.ibm.com/software/products/en/integrated-law-enforcement>.

3. Proposed software architecture

3.1 Overview

The software is web-based system that is based on MySQL database hence it is a database driven system. Users will use the system via a web browser and the webserver will handle their http/https requests. Users will login to the system to use it, and each user will have certain roles that facilitate security authorizations for different parts of the system. The basic component on the system is the CRUD where it is web pages that list, add, edit, and view system entities. From CRUD component the administration component will be feasible so the admin can set roles and these roles are to control the access of the users on the CRUD operations. That is a particular role can view and list entities but cannot edit them, where other roles can. The security part will be shown under access control and security section. On top of CRUD layer the case management also there where it uses some system entities that were existed via CRUD. Case management component is the process of recording incident reports and open investigative cases to be worked on and solved by the collaboration between officers and investigators. It includes assigning investigators, adding field interviews, uploading pictures, distributing tasks, adding and evaluating evidences

etc. The last component is the system intelligence. This component depends on the outputs of case management component as inputs to system intelligence modules. The outputs of this component is providing suggestions in terms of criminals characterizations to given crimes, and this will be based on previous data, which is a machine learning operation.

3.2 Subsystem decomposition

As described in Design goals section, there are four subsystems, which are CRUD, case management, system intelligence, and administration.

The CRUD operation is basically data base driven web functionality where the user such as admin can add new officers, investigators, or persons and their properties such as their cars and home addresses. It also includes the uploading files functionality and setting access permissions. The CRUD is the very basic level to the system to be useful since users can make the data persistence by into the database and can view these data, update, and delete them.

The case management part is the bridge between CRUD operations and system intelligent part, and it has the management functionality to operate crime cases from beginning until the case is closed. The roles of this part are officers, investigators, and supervisors. The case is opened by either an officer or a supervisor and after that a group of officers and investigators can be initiated and assigned to the case. A case will pass through workflow points such as requesting investigator to handle the case and the investigator's response (can be accepted or refused). The case input will be filled in at this part of the system, which will facilitate the intelligent part to analyze case input and provide some useful outputs as data mining suggestions.

The system intelligence part is an application of data mining algorithms and methods. The parts of data mining that the system applies are classification and may uses association. The classification method will be applied by utilizing the Artificial Neural Network (ANN). The ANN model will be trained with previous cases, and the method is by using case's inputs such as the tool used to kill or the time of the death of the body and so. These inputs will be used as inputs for ANN model and the outputs will be a classification of the criminal such as the gender, age, education level etc....

For the security part, Java Authentication and Authorization Service (JAAS) will be used to apply login functionality, authentication, and authorization. The web communication must go via secured connection such as SSL AKA https. The system admin will have the ability to assign user roles, and their accessibilities.

Figure 3.2.1 shows a visual representation of the subsystems.

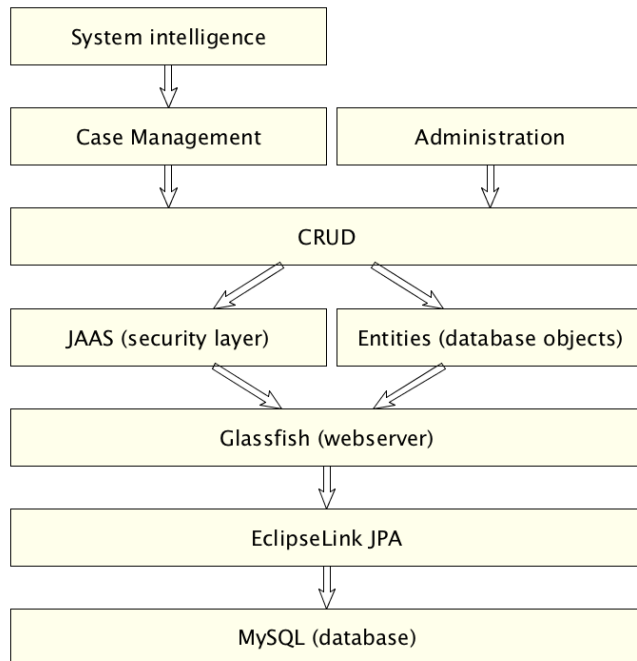


Figure 3.2.1 A visual representation of the Adala's subsystems.

3.3 Hardware/software mapping

Adala is a web-based architecture where there is a server machine that contains any operating system that run Java Virtual Machine (JVM). It is recommended to run on Linux operating system that is stable such as CentOS, Debian, or RedHat. The client side is a web browser that can be running on any machine. The connectivity can be either via the Internet or via local network if the server is meant to be a machine connected locally in the organization network.

3.4 Persistent data management

The data is stored on MySQL database. The tables are auto generated by EclipseLink JPA during development phase. EclipseLink JPA generates MySQL tables and their relations based on the system entity beans' configuration. The security part has its own database, which is different database name than the whole system, and this is to enhance the security level. The security database will be shown in Access control and security section. Figure 3.4.1 and Figure 3.4.1 show the UML diagram for the system entities.

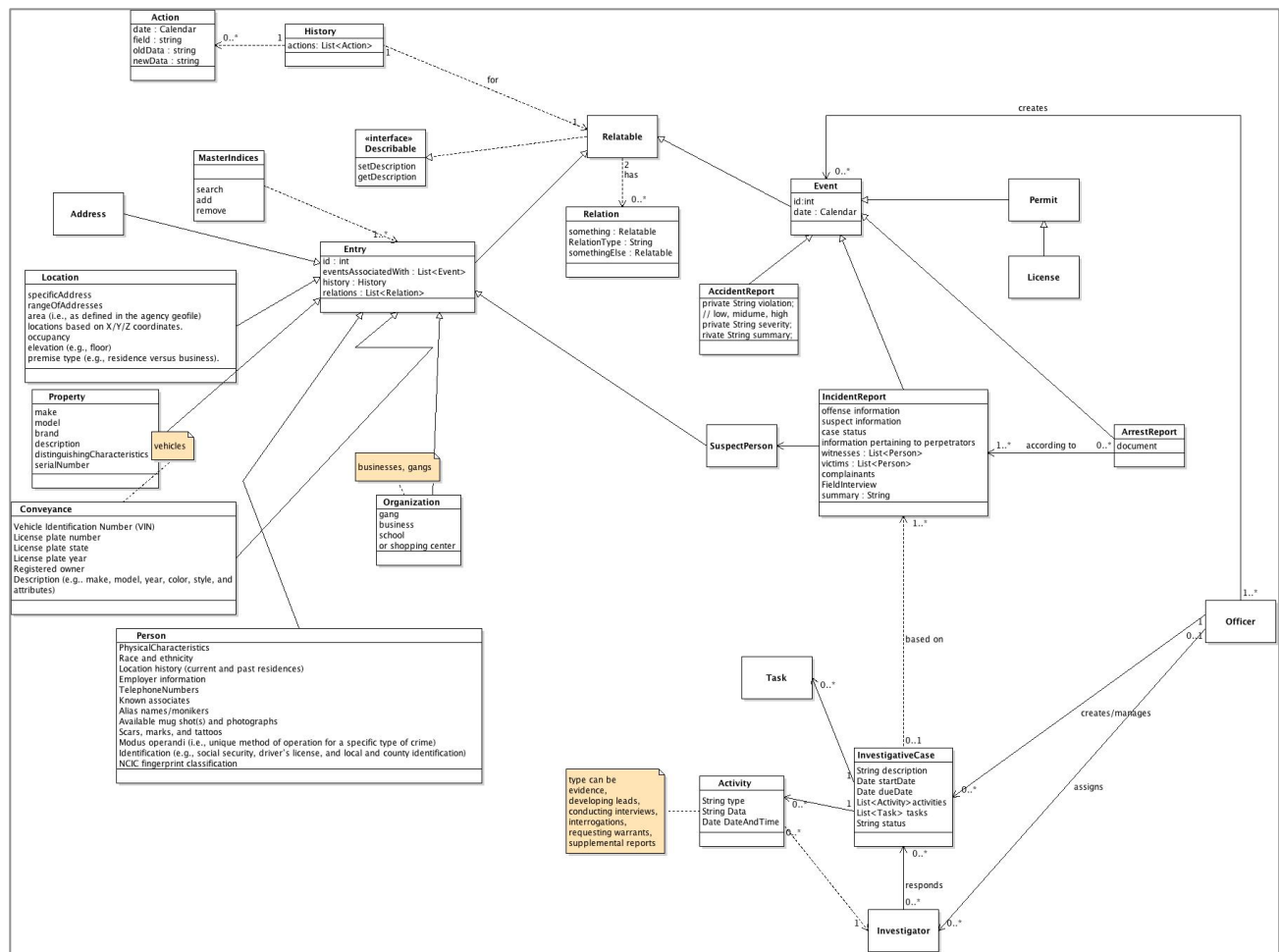


Figure 3.4.1 UML diagram for the system entities. Some fields are without type those are Strings.

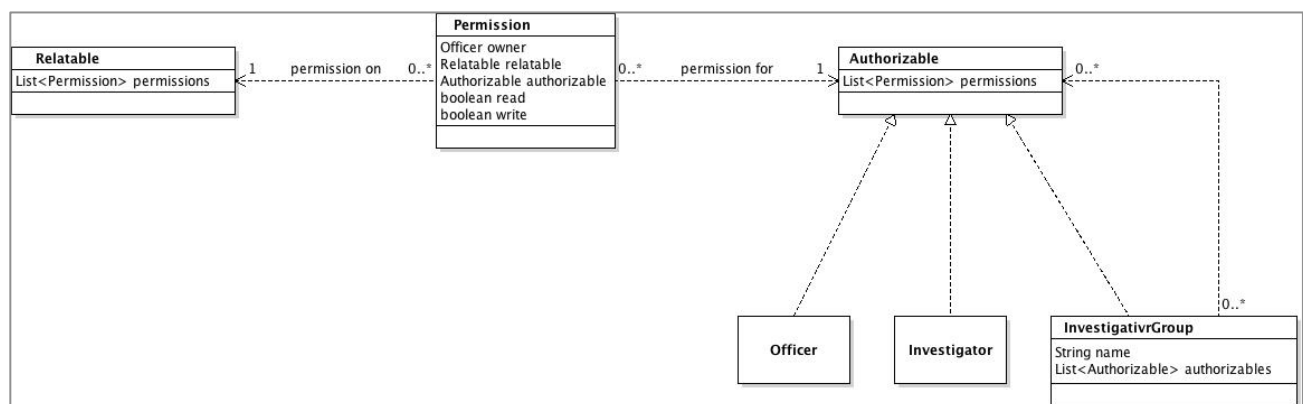


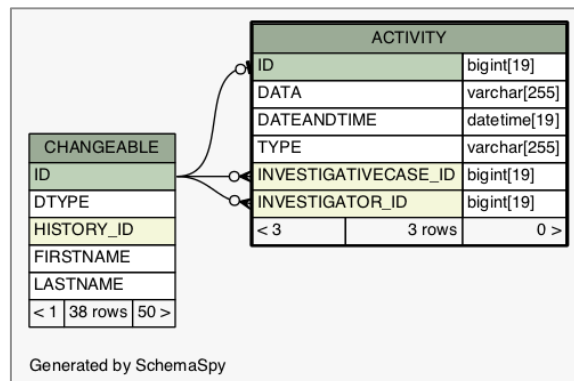
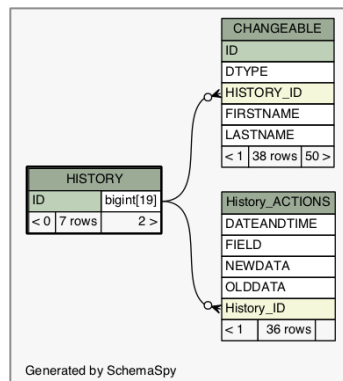
Figure 3.4.2 UML diagram for security permissions. Some classes has no fields those classes refer to the previous UML digram.

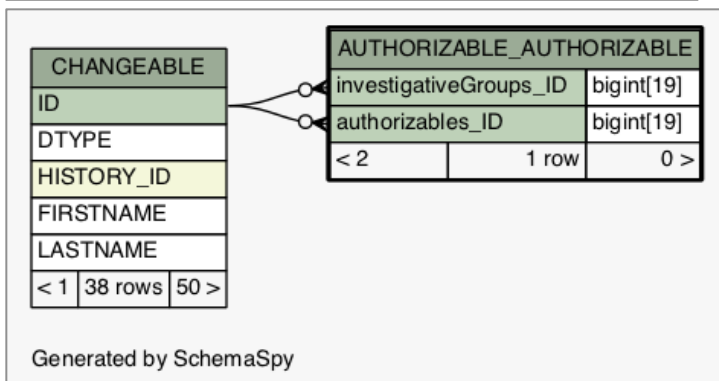
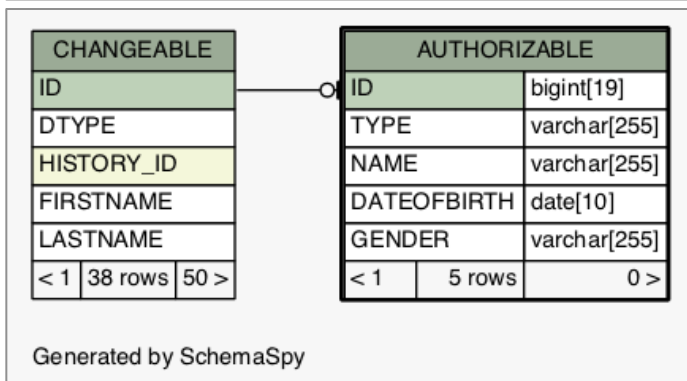
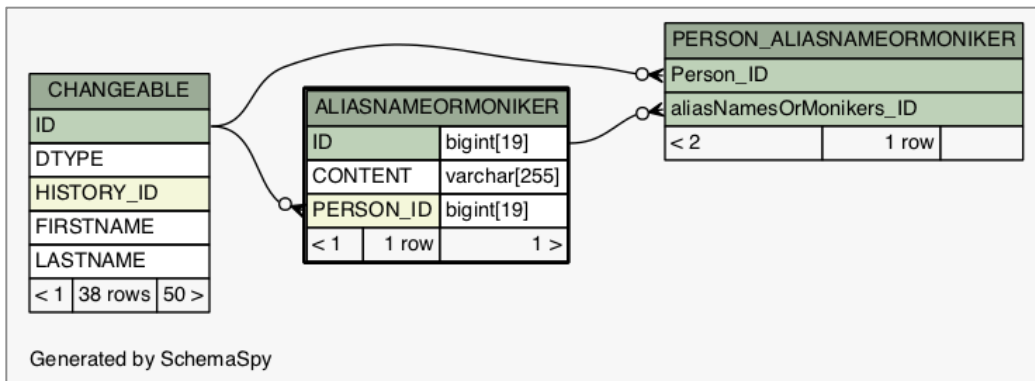
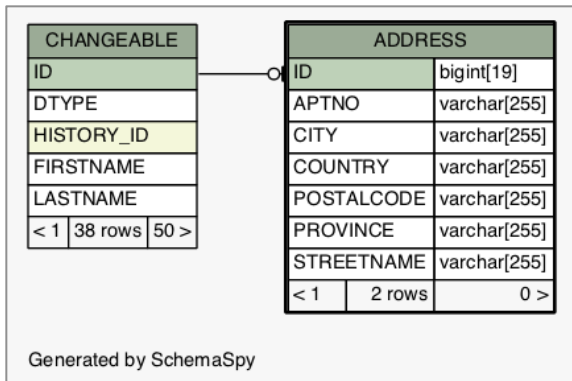
The following is a visualization of MySQL tables and their relationships:-

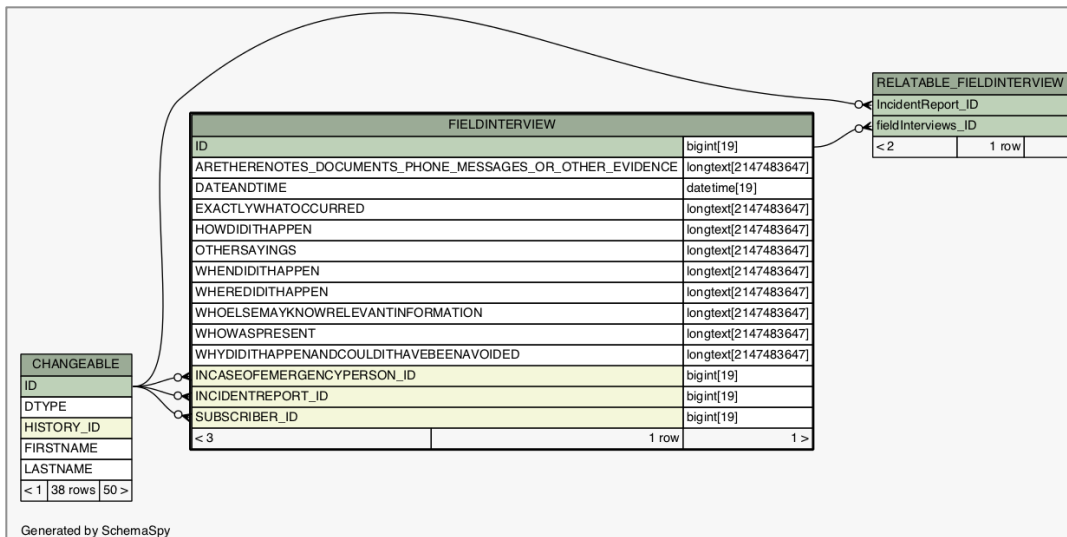
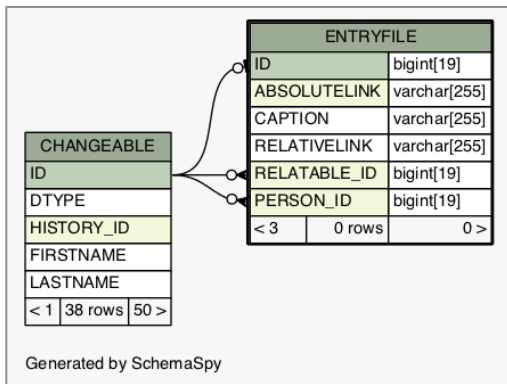
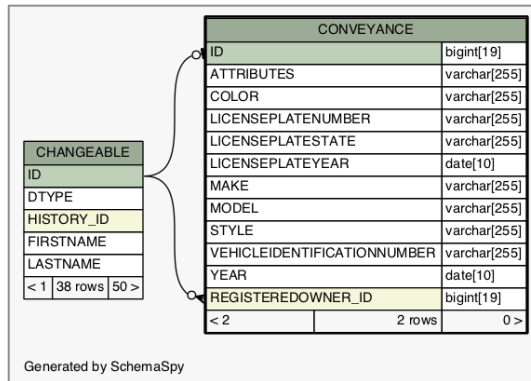
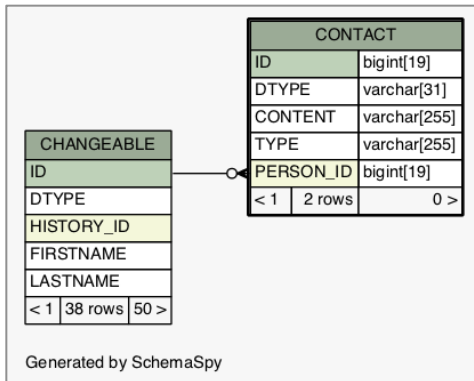
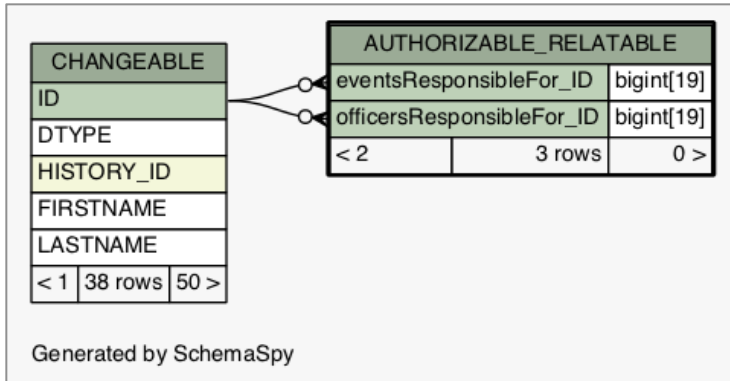
Table	Children	Parents	Columns
ACTIVITY		3	6
ADDRESS		1	7
ALIASNAMEORMONIKER	1	1	3
AUTHORIZABLE		1	5
AUTHORIZABLE_AUTHORIZABLE		2	2
AUTHORIZABLE_RELATABLE		2	2
CHANGEABLE	50	1	5
CONTACT		1	5
CONVEYANCE		2	12
ENTRYFILE		3	6
FIELDINTERVIEW	1	3	14
HISTORY	2		1
History_ACTIONS		1	5
IDENTIFICATION		1	4
LOCATION		1	6
LOCATION_ADDRESS		2	2
MASTERINDICES			1
ORGANIZATION		1	3
PERMISSION		4	6
PERSON		1	15
PERSON_ALIASNAMEORMONIKER		2	2
PERSON_SCARMARKTATTOO		2	2
PROPERTY		1	6
RELATABLE		3	15
RELATABLE_AUTHORIZABLE		2	2
RELATABLE_FIELDINTERVIEW		2	2
RELATABLE_RELATABLE		2	2
RELATABLE_SUSPECTPERSON		2	2
RELATION		3	4
SCARMARKTATTOO	1	1	3
SEQUENCE			2
SUSPECTPERSON		1	8
TASK		1	2

34 Tables

164







CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

IDENTIFICATION			
ID		bigint[19]	
CONTENT		varchar[255]	
NAME		varchar[255]	
PERSON_ID		bigint[19]	
< 1	2 rows	0	>

Generated by SchemaSpy

CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

LOCATION			
ID		bigint[19]	
AREA		varchar[255]	
COORDINATE		longblob[2147483647]	
ELEVATION		varchar[255]	
OCCUPANCY		varchar[255]	
PREMISETYPE		varchar[255]	
< 1	1 row	0	>

Generated by SchemaSpy

CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

LOCATION_ADDRESS			
addresses_ID		bigint[19]	
locations_ID		bigint[19]	
< 2	2 rows	0	>

Generated by SchemaSpy

CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

ORGANIZATION			
ID		bigint[19]	
NAME		varchar[255]	
ORGANIZATIONTYPE		varchar[255]	
< 1	1 row	0	>

Generated by SchemaSpy

CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

PERMISSION			
ID		bigint[19]	
READPERMISSION		varchar[255]	
WRITEPERMISSION		varchar[255]	
AUTHORIZABLE_ID		bigint[19]	
OWNER_ID		bigint[19]	
RELATABLE_ID		bigint[19]	
< 4	1 row	0	>

Generated by SchemaSpy

CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

PERSON			
ID		bigint[19]	
NCIC_FINGERPRINTCLASSIFICATION		varchar[255]	
BIRTHPLACE		varchar[255]	
CITIZENSHIP		varchar[255]	
DATEOFBIRTH		date[10]	
GENDER		varchar[255]	
MODUSOPERANDI		varchar[255]	
FIRSTNAME		varchar[255]	
LASTNAME		varchar[255]	
BUILDCHARACTERISTICS		varchar[255]	
COMPLEXIONCHARACTERISTICS		varchar[255]	
HAIRCHARACTERISTICS		varchar[255]	
HEIGHTCHARACTERISTICS		varchar[255]	
RACE		varchar[255]	
THREATASSESSMENTLEVEL		varchar[255]	
< 1	4 rows	0	>

Generated by SchemaSpy

CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

SCARMARKTATTOO			
ID		bigint[19]	
CONTENT		varchar[255]	
PERSON_ID		bigint[19]	
< 1	1 row	1	>

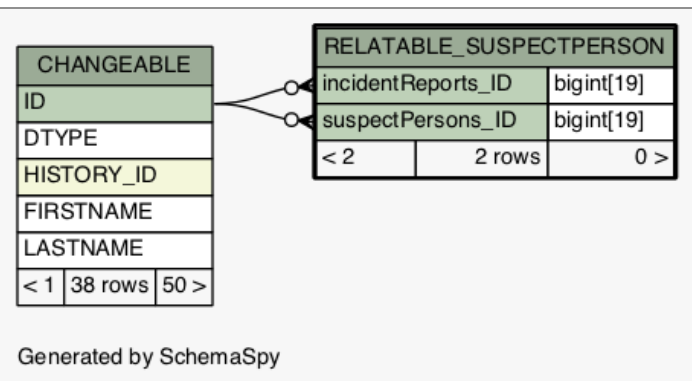
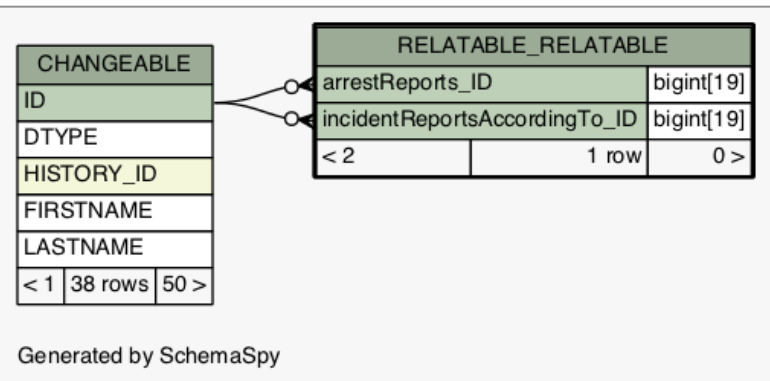
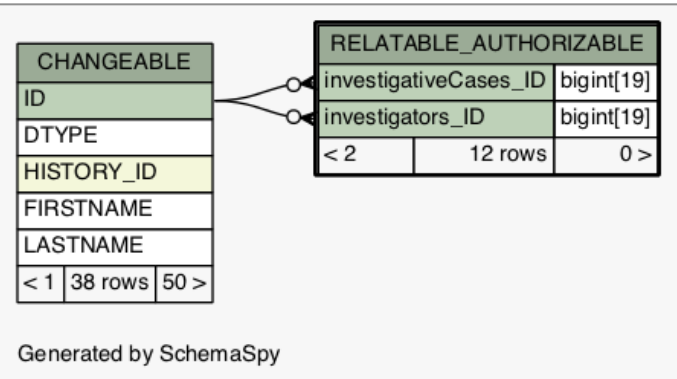
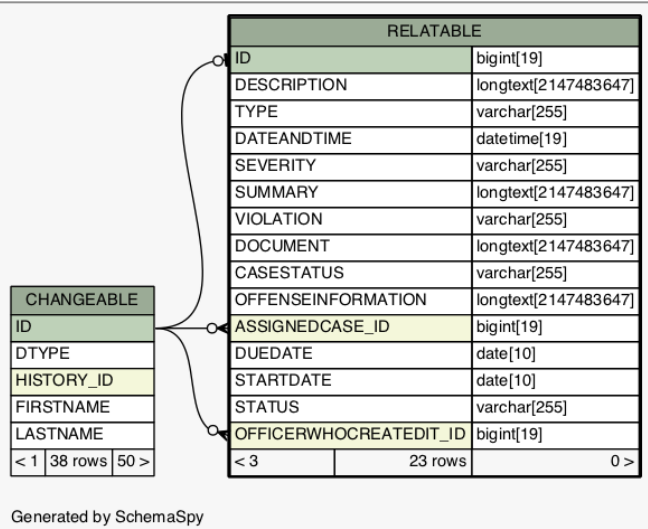
PERSON_SCARMARKTATTOO			
Person_ID			
scars_marks_tattoos_ID			
< 2	1 row		>

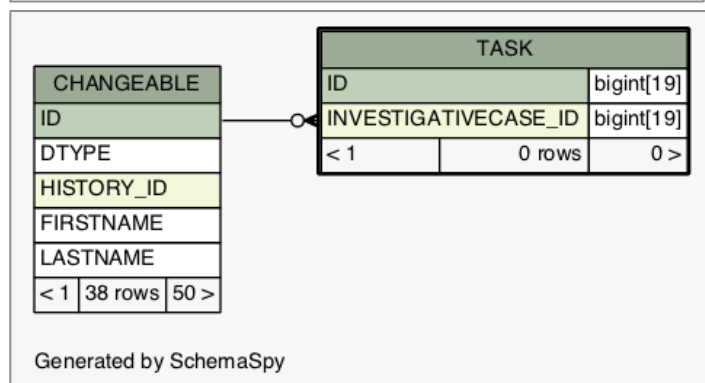
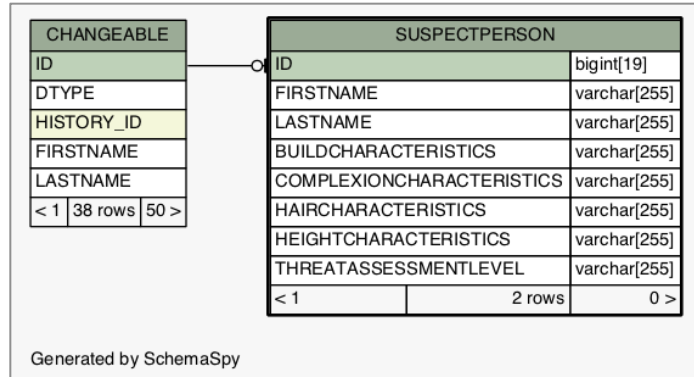
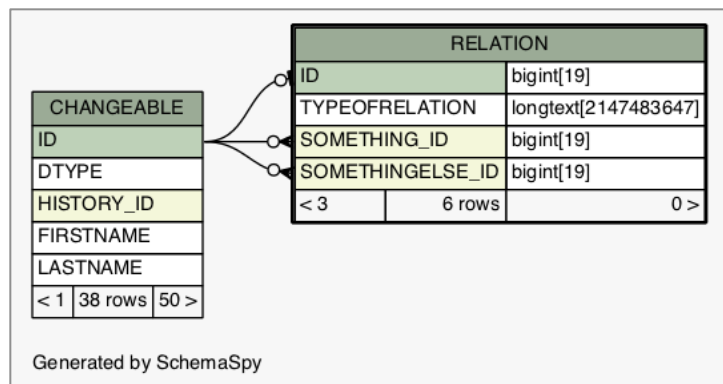
Generated by SchemaSpy

CHANGEABLE			
ID			
DTYPE			
HISTORY_ID			
FIRSTNAME			
LASTNAME			
< 1	38 rows	50	>

PROPERTY			
ID		bigint[19]	
BRAND		varchar[255]	
DISTINGUISHINGCHARACTERISTICS		varchar[255]	
MAKE		varchar[255]	
MODEL		varchar[255]	
SERIALNUMBER		varchar[255]	
< 1	1 row	0	>

Generated by SchemaSpy



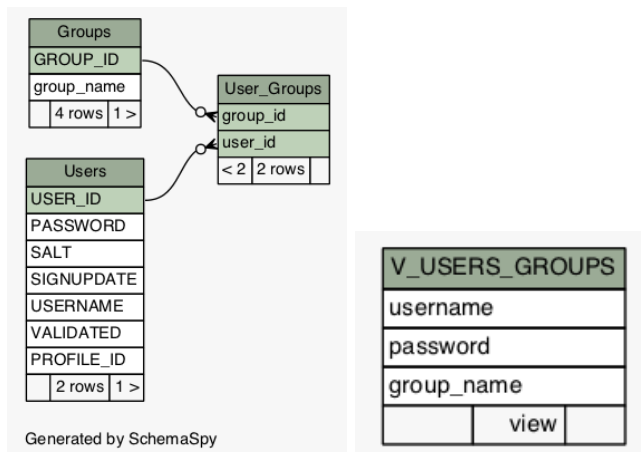


3.5 Access control and security

The software uses JAAS to control the access and security. Users will be under certain roles and each role has a specific authorizations. Also, there is a set of permissions that will be applied on information accessibility, means that a supervisor can create a set of permission on other users or group of users that on any Entity object in the system. The permission record contains an id of entity object, an id of authorizable entity (i.e. officers, investigators, supervisors, or an id of an investigative group object entity that it is an authorizable object itself and contains a list of other authorizables, see Figure 3.4.2), and read/write booleans. There are four main roles: admin, supervisor, officer, and investigator. The admin role is able to use all system functionality with no limit including setting the roles to other users and activate/deactivate users. Supervisor role will be able to use edit operations on system entities and set read/write permissions. Officers will be able to view data and create new data such as new investigative cases. Investigator role is limited to investigative process only.

The following figures show the database tables of the security part.

Table / View	Children	Parents	Columns
Groups	1		2
User_Groups		2	2
Users	1		7
V_USERS_GROUPS			3
3 Tables			11
1 View			3



3.6 Global software control

The software is web-based system that relies on Glassfish webserver. Glassfish is able to handle http, https, and other type of requests. It also has session management and it uses a pool of threads methods to manage new connections in threads effectively. However, Glassfish is a Java application, which needs to be care of in terms of JVM configurations such as garbage collector and maximum heap size and other stuff. Also, Glassfish itself has a configuration page to tune the size of initial threads number, the number of new thread if the old ones were consumed, and threads timeout and http sessions timeout as well. All theses configurations must be tuned after the system is deployed to ensure the stability and reliability of the system.

3.7 Boundary conditions

Memory consumption is a main concern to Adala system since it is built in Java where it consumes memory resources relatively more compared with other technologies such as LAMP stack. This may leads for low performance if the number of users requests is higher than what server RAM size can hold. It is recommended to have 4G RAM size or more for medium size group (1000 user approx.).