

DESIGNING DATA WAREHOUSES TO SUPPORT CRIMINAL INVESTIGATION

*Juan C. Rivera-Vázquez, Universidad del Turabo, Gurabo, PR jjn0611@yahoo.com
Lillian V. Ortiz-Fournier, Universidad del Turabo, Gurabo, PR marianavanessa@yahoo.com
Mysore Ramaswamy, Southern University, Baton Rouge, LA mysore@acm.org*

ABSTRACT

The lack of well designed knowledge management systems in the intensive and time critical law enforcement environment poses interesting problems for information technology professionals. The objective of this paper is to study how the different databases at the Police Department of the Commonwealth of Puerto Rico can be integrated. We propose the creation of a data warehouse to store relevant criminal data that will be helpful in crime prevention, locating criminals and finally bring them to justice. The Criminal Information Support System (CISS) project includes the integration of multiple systems having different functionalities. CISS is a cost-efficient web-based system that will enable the law enforcement departments to share information from different data sources. We present a descriptive study of a proposed data warehouse system that aims to improve the criminal data management in different ways such as developing effective plans for crime prevention, finding useful criminal behaviour patterns, and making correlations between similar circumstances to solve difficult cases successfully and to avoid further casualties. We recommend a structure that integrates the different data sources of the Puerto Rico Police Department (PRPD) that can be used to share real-time information through the Internet and the Intranet of law enforcement agencies within the island and the United States. This is the first stage of a larger project that includes online analytical processing (OLAP) tools well suited for advanced extractions and a decision support system.

Keywords: Data Warehouses, Data Quality, Data Marts, Data Integration, Data Mining

INTRODUCTION

The problem of collaboration and information exchange between police units in Puerto Rico have been increasing gradually through the years. The evolution of technology has changed the way criminals break the law challenging more than ever before the law enforcement agencies due to the lack of an efficient collaboration mechanism. According to Hauck [12] the development of new technologies during the last decades has helped to improve the ability of many organizations to understand and disseminate the data collected. Also, the acquisition of better databases adds accessibility, speed of retrieval, and searching flexibility to information systems. In a recent case on April, 2007 [21] in Puerto Rico, police officials arrested a man for drug possession who happened to be a fugitive who had escaped from jail with other seven offenders. At that time this felon was not recognized by the police as a fugitive and was released since the authorities could not find anything to imprison him. Usually the Correction Department takes two hours to report a breakout but the lack of an efficient system for sharing the information with the PRPD brings opportunities for criminals to commit other felonies and leave the country without notice.

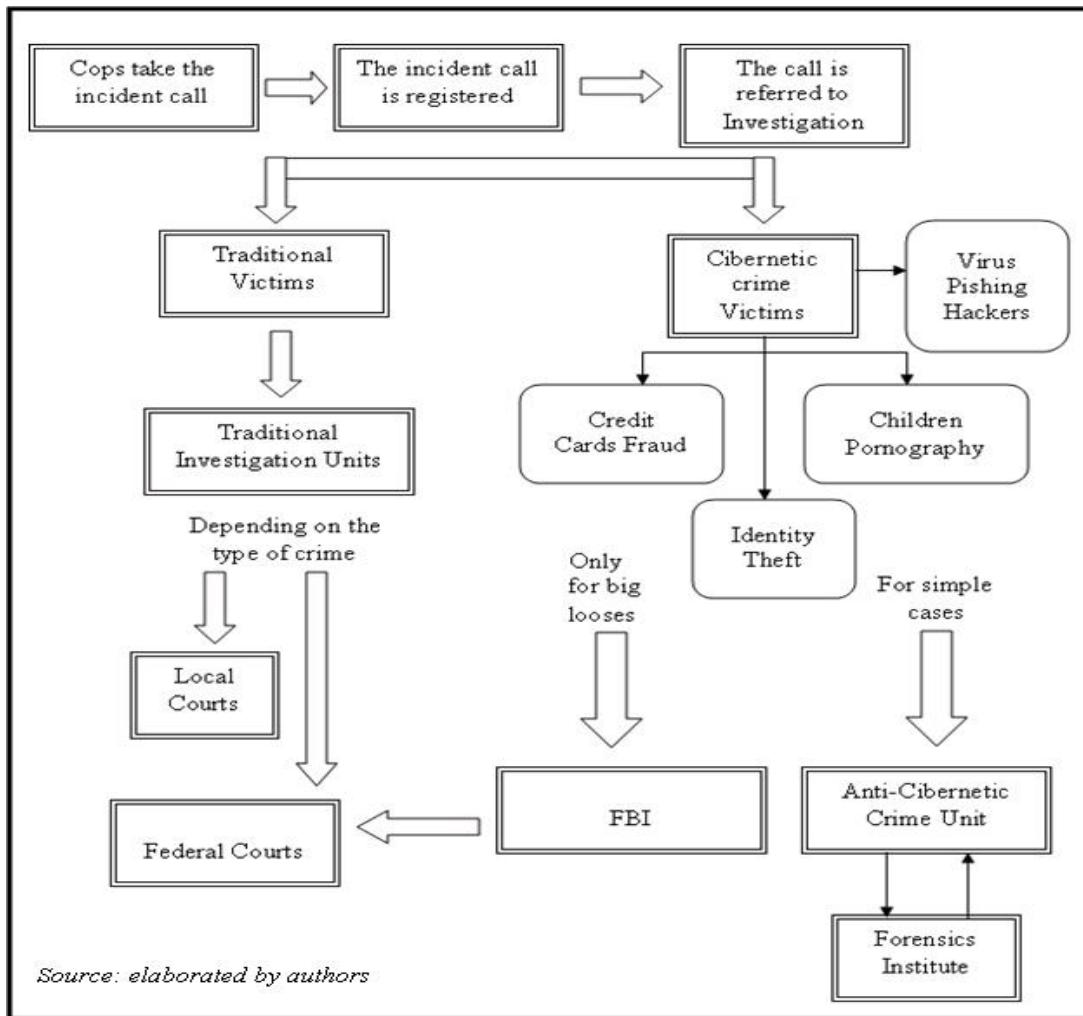
The investigation process in the Puerto Rico Police Department (PRPD) follows several steps. The police officer receives a call and goes to where the incident has occurred. Then, he/she obtains information from the scene, makes arrests if necessary and completes a report. A detective receives the case, reviews it, and does some research and interviews. When appropriate, there are more arrests and bookings and the case is prepared for trial. The PRPD currently do not have integrated information systems but have already started the implementation of a centralized analysis unit. After receiving the call, the police officer notifies the central office of the PRPD which is also in charge of receiving calls from the citizens. At the central office the officers are assigned to cases, the data from criminal incidents is entered into the available information systems and stored in databases, and referrals are made if the situation requires it.

The management process of the criminal incidents at PRPD depends on the type of crime. When a traditional crime is committed the call is registered at the central office and it follows the steps mentioned above until it gets to a trial. Otherwise, if it is a cybernetic crime the case is referred to the Federal Bureau of Investigation (FBI), although,

PRPD is starting to implement an anti-cybercrime unit. The process for collecting criminal information used by officers and detectives lacks an effective integration and collaboration system. There are many databases used for this purpose that are not interconnected making it more difficult to bring criminals to justice as a result of incomplete time-critical information. Figure 1 presents the criminal incidents management process at PRPD.

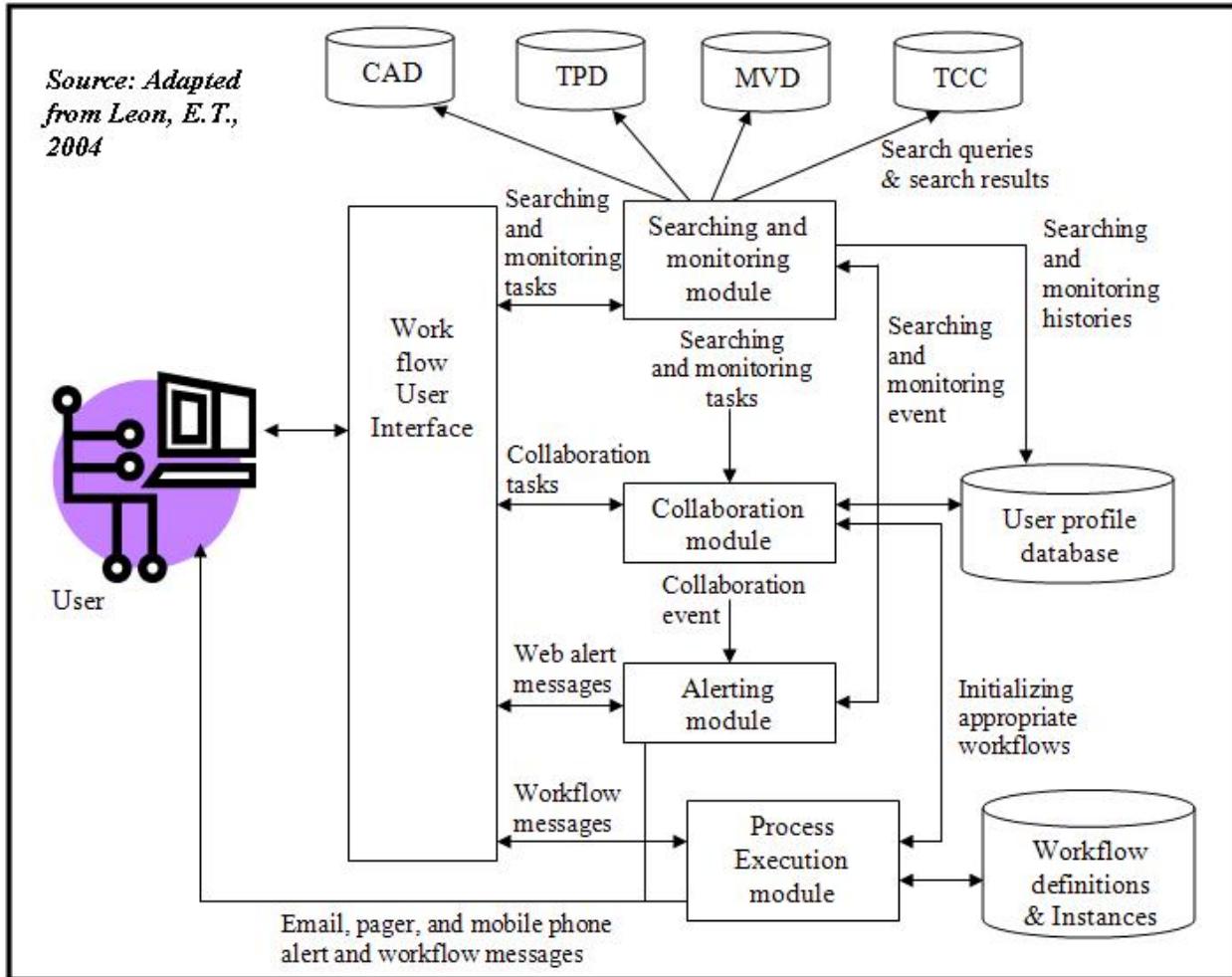
In law enforcement, as in many other applications,

Figure1. Criminal incidents management process at the PRPD



collaboration among members of dynamically defined, task-oriented teams plays an important role in everyday operations [4]. In order to have an effective collaboration project, it is necessary to have a system that harvests and processes the data in a standardized way so that any agency can easily retrieve and use it [3, 25]. In 1997 the National Institute of Justice created a project called COPLINK which is a system of local exchange of information and other agencies, in association with the University of Arizona Artificial Intelligence Laboratory and the Tucson Police Department (TPD) at Arizona [23]. Figure 2 presents the architecture of COPLINK. Through this system, stored information can be exchanged between officials and other security agencies almost simultaneously. COPLINK is a solution suite at a tactical, line-level to the problem of inaccessible or irretrievable information resulting from disparate law enforcement information systems that lack of a common language or platform [7]. It is an integrated information and knowledge management environment which represents an effective solution to the interchange and collaboration problem between law enforcement agencies.

Figure 2. Architecture of the Coplink System



The development of COPLINK was mainly motivated by several challenges encountered by law enforcement officials such as (a) need for real-time or near real-time information from dynamic data sources; (b) differences on information sharing needs among law enforcement officers; and (c) a need for a user-friendly interface, user information search history, and user session management, among others [4]. The major goals for the COPLINK project for TPD are (1) to develop an integrated system that allow TPD officers an easy access to information contained in their systems, and (2) to design a prototype system for developing similar systems at other police departments. Knowing that criminals disregard jurisdictional boundaries and, in fact, take advantage of the lack of communication across jurisdictions, COPLINK offers a model for allowing different police departments to share information [5]. The application has two components, COPLINK Connect and COPLINK Detect. The intended users of the Connect component are police officers who are not experienced IT users, but have time and mission critical information needs. The design was mainly guided by user requirements acquired through a multiphase process [6].

The focus of our study is to develop an integrated repository to collect and relevant criminal information at the Police Department of Puerto Rico. We intend to recommend the design of a data warehouse that (1) will retrieve and store criminal data from different sources and platforms; and (2) will improve the integration, collaboration and

information retrieval between law enforcement departments inside the island and with the United States systems. This repository can be connected to OLAP applications that will bring updated information to the officers at the crime scene, enhance the process of criminal analysis, and provide effective methods of criminal activity forecast, among other functionalities that may improve the crime prevention mechanisms in place. It is recognized that the systems based on expert systems methodology have the problem of not having a complete knowledge embedded into them [5]. Taking this into consideration, the proposed system will use the COPLINK project as a model for integrating information faster from different data sources making possible the time-critical update of the data warehouse which will be able to feed decision support systems that will improve investigative tasks.

A CRIMINAL DATA WAREHOUSE

The design of a data warehouse is often a challenging effort for developers. It affects many business areas of the enterprise that manage a workload of thousands of daily transactions. A data warehouse is defined as a repository of integrated information, for available querying and analysis [14]. This means that it is used to store data for consultation that will be useful for a particular company or business. Also, in a data warehouse, time dimension of historical data is used mostly to detect interesting changes which could indicate a particular problem or opportunity. The idea behind data warehousing is to extract, filter and integrate excellent information using consultations outposts [11]. As stated by Nemat et. al. [18] since mid-80s data warehouses have been developed and deployed as an integral part of a modern decision support environment. A data warehouse provides an infrastructure that enables businesses to extract, cleanse, and store vast amounts of corporate data from operational systems for efficient and accurate responses to user queries. It also empowers knowledge workers with information that allows them to make decisions based on a solid foundation of fact.

The development and implementation of these repositories keep growing in different business areas such as financial, educational, insurance, commodities, among others. This need for data repositories is also increasing remarkably in the security and law enforcement agencies. While each police department has a different database, they might be barely sharing data in the same region, county or state. This results in collaboration difficulties usually increased by the introduction of new criminal behaviours through the Internet. For instance, the methods used by criminals to commit their misdeeds have evolved and crossed over with the sprouting of new technologies. A criminal can be in several places simultaneously committing robberies using an email, and banking frauds using a computer connected to a common network. Take Costa Rica, for example, where a criminal sent an electronic mail from a remote server in China, copied the page of the National Bank of this country and caused damages to 150 clients with a fraudulent request to update their private data [19]. That same criminal could have committed other crimes in different countries or states but because of the lack of integration between agencies it is hardly captured. Furthermore, the problem increases when the information in most cases is stored in different formats and systems using COBOL, Oracle or SQL Server, to mention a few, making it difficult to exchange information between law enforcement departments.

In this paper, we describe the data integration needs of the Puerto Rico Police Department (PRPD). We recommend that this agency migrate the data of all the different databases into a data warehouse to improve data management and sharing between police departments within the island and with the United States. Also, this study intends to model the scheme of the proposed data warehouse that will be connected to a graphical user-friendly interface for any civil personnel or agent to find criminal information. This analysis and design is the first step for further implementation of a decision support system (DSS) that will help the Police Department to prevent crime and manage efficiently criminal investigations in Puerto Rico.

GATHERING DATA WAREHOUSE REQUIREMENTS

Data warehouses are online analytical processing (OLAP) systems well suited to answer questions that involve analytics including aggregation, drilldown, and slicing/dicing of data. They are by definition subject oriented, integrated, non-volatile, and time variant collection of data in support of management's decisions [14]. A large number of data warehousing methodologies and tools are available to support the growing demand of these systems. To improve performance denormalization is usually promoted in a data warehouse environment. Data warehouse implementation activities include data sourcing, data staging (ETL), and development of decision support oriented

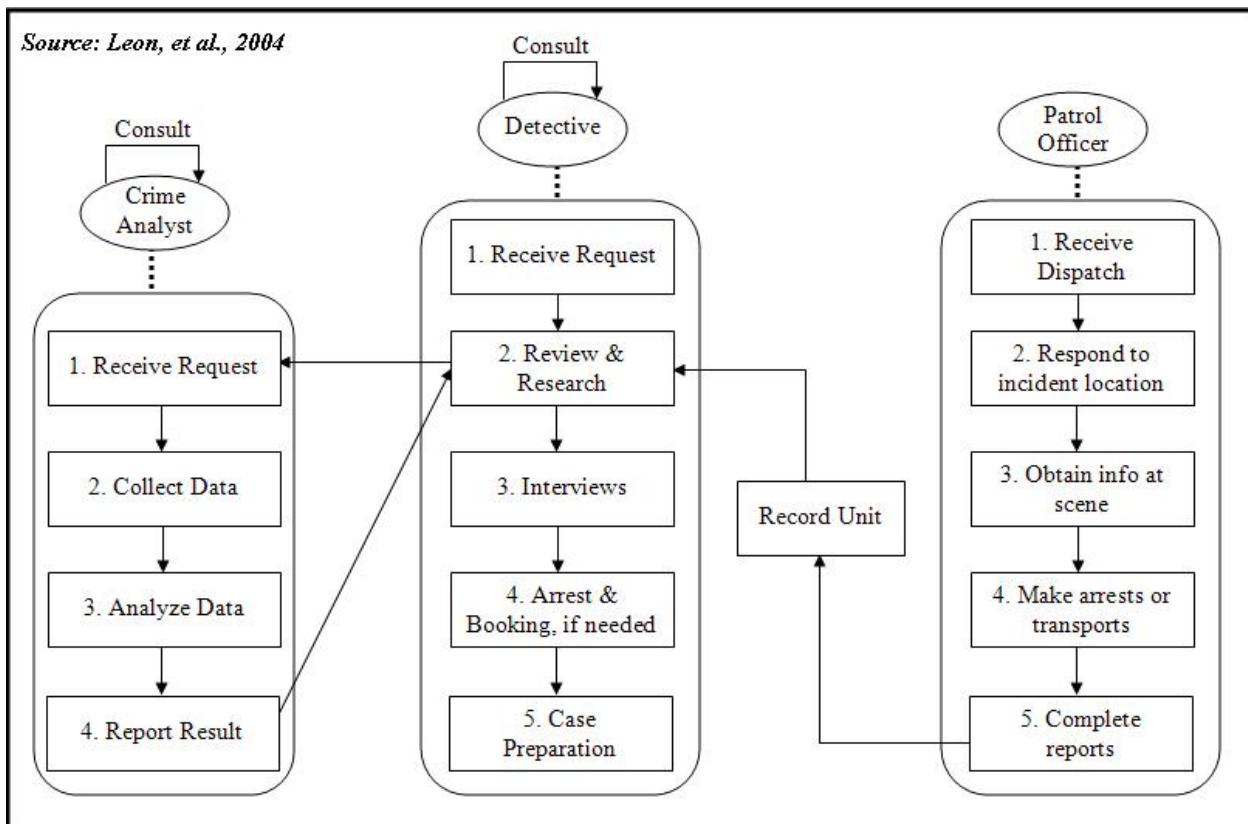
end-user applications [22]. Inmon [14] recommends that instead of starting from requirements, data warehouse development should be driven by data. Data is first gathered, integrated, and then, tested.

There are multiple approaches to build data warehouses. A number of strategies are available for designing the data warehouse architecture, ranging from enterprise wide data warehouse design to data mart design. The organization needs to determine which approach will be the most suitable before adopting a methodology [22]. The data mart design consists of several departmental or local data marts combining into a data warehouse. It is a faster and easier implementation of manageable pieces, inherently incremental and less risk of failure [20]. The data marts are intended for different kind of information gathered from the different databases used by law enforcement agencies.

The data mart or bottom-up approach was selected for developing the proposed data warehouse. For this we will have a functional model that represents how the system would work and a reference architecture that reveals its physical and technical implementation using different software and hardware tools. Using this approach, the data marts of different functional areas are developed first and then integrated into an enterprise-wide data warehouse [2].

The system will be named as Criminal Information Support System (CISS). CISS will be able to collect data from different databases, and consolidate them in a single repository that might be queried to find patterns, correlations, criminal information, and incidents using a single tool [13]. The PRPD recently evaluated the information technology structure and the collaborative processes they follow that are presented in Figure 3.

Figure 3. Collaborative process in a law enforcement agency



These tasks are similar to COPLINK project, except for the Crime Analyst module [17]. From this analysis, the officials at PRPD identified a lack of information sharing, and problems with integration and knowledge

management similar to the ones found by the TPD in Arizona [6]. The PRPD has the following databases which can only be accessed at the main office.

1. CAD (Positron) system for capturing incidents in Microsoft SQL server.
2. Separate system for mug shots (photos taken at the time of tribunal finding causes for arrest).
3. System of Analysis and Statistics (SAEC) to collect all the principal incidents as a diary in MYSQL database.
4. System of Incident Arrest (SADIC) for presenting to the tribunal in MYSQL database.

These systems need to share information through the different units that PRPD has throughout the island. The department requires a good Intranet and multimedia technology to easily access and manage criminal information. It has also several problems which are similar to other law enforcement agencies and are related to the effective development of an integrated system that allows police officers from different units a quick and easy access to relevant information. Other problems that this project intends to solve include the design and development of a user-friendly graphical interface, and the improvement of the analysis applications that support specific law enforcement functions and tasks. As in COPLINK, the capability to easily incorporate and make use of additional information sources is important for law enforcement applications due to the frequent need of cross-jurisdictional collaborations in dealing with crimes that are typically not confined to one geographical location [6].

The following are the main design criteria considered for the CISS project which are adapted from those considered for COPLINK, described by Chen et al. [4].

1. Platform independence: Not every police unit or department uses the same hardware, software, or operating system.
2. Stability and scalability: To be able to grow and expand, we recommend Windows 2003 server 64 bit for the Operative System, Microsoft SQL server for the database, and a Xeon sever with 2GB Ram and 4 hard disks of 300GB in Raid Drive configuration.
3. Easy to use: The graphical user interface should be user-friendly and flexible enough to meet the standards and investigation needs of detectives and officers.

INTEGRATION OF DATA SOURCES

Following the recommendation of Inmon [14], first, we need to analyze the different sources of data used to feed the warehouse. Azhar [2] suggests the use of data profiling to do the examination as it represents a systematic examination of the contents, structure and quality of a data source. The worst case scenario will be the data profiling revealing that the source of data is deeply flawed and cannot support the business objectives, therefore the data warehousing effort may need to be called off [15]. The following steps have to be performed to prepare the data source profiles.

1. Identification of data sources within and outside the organization and examination of data format. The PRPD has eight relevant data sources.
 POSITRON: format=SQL server
 SAEC: format=MYSQL
 SADIC: format=MYSQL
 DRUGS: format=MYSQL (Under construction)
 WEAPON REGISTRY: format=SQL Server
 MUGSHOT: format=SQL Server
 VEHICLE: format=SQL Server
 SEXUAL OFFENDERS: format=SQL Server
2. Locating and staging data of interest: The data in these data sources have similar syntax and are very compatible. This step is necessary to identify similarities between data types, missing values, and values with inconsistent data types. A conceptual data model has to be prepared to verify the facts and dimensions and designing the data mart schema best suited for the data warehouse. Facts represent quantitative (or factual) data about a business entity transaction which is an object or event for which we need to capture and store data. Dimensions contain descriptive data that reflect the dimensions of that entity [2]. Factual data is

Issues in Information Systems

Volume XII, No. 1, pp. 445-454, 2011

more stable than dimensional data because dimensional data change more frequently over a period of time than the factual data [10].

The data from the different data sources are retrieved using the information packages model suggested by Ponniah [20]. This schema serves to understand better the composition of the data marts as presented in the Table 1. The information package shown has the information subject CRIME that was selected from the ones used for COPLINK. These are recommended especially in designing a system that could be deployed across multiple law enforcement agencies to acknowledge that no two agencies would store their incident data in exactly the same way [5]. The selected dimensions are “Person”, “Vehicle”, “Drug”, “Weapons” and “Incident” mainly because this is the data structure used in the PRPD.

The Staging Area model frequently referred to as the model for Extraction Transformation-Load (ETL), is used to extract data from the source systems and transforming different source data standards into a single one [2]. It is considered the core of the data warehousing project because the effectiveness of the data warehouse that is ready to work with a decision support tool depends on how good the data standardization is. Staging is necessary to enforce data quality and consistent standards for data integration in a way that separate sources can be used together and upload the data successfully into data marts [17].

Table 1. Information Package. (Source: Adapted from Ponniah, 2001)

INFORMATION SUBJECT: CRIME					
TIME	VEHICLE	INCIDENT	DRUGS	PERSON	WEAPON
Year	Model	Location	Organization	Name	Model
Quarter	Type	Sector	Member	Alias	Caliber
Month	Location	Agent	DrugsQuantity	Ssn	Description
Date	Owner	Description	OrgLocation	Dob	Owner
DayWeek	SellingYear	Unit	DrugType	Race	Vendor
Holiday		Precinct	ShipmentSource	Gender	
				Weight	
				Height	
				Eyecolor	
				Haircolor	
				GangFlag	
				CautionFlag	
				PhotoFilename	
Facts: CrimeId, CrimeType					

We used the following staging steps as proposed by Azhar [2] (1) extraction and transformation, (2) validation, filtering and corrections, and (3) integration, which targets same data types as well as different data types. The staging will be developed using the Integration Services of Microsoft SQL Server for its cost-effectiveness while Microsoft technology is already available at the PRPD. The transformed data will be stored in temporary data files for assuring data quality and consistency checks. The final version of the data, called targeted data, will be stored in the permanent database files of the data warehouse.

Data validation, filtering and integration are important steps of the staging process. The purpose of validation is to ensure data quality, and correct errors, omissions, or inaccuracies before loading the data loaded into data marts. The following checks are performed to validate data on each record in the data source files.

1. Uniformity check to detect any out of range values or impossible values
2. Conformity check to notice mismatching data units and data types
3. Syntax check to identify erroneous field lengths
4. Check for any missing values

Each faulty record is corrected before moving to the next record. This process needs to follow some additional rules for accurate filtering of criminal data coming from each data source right after data validation, filtering and integration processes are performed. The purpose of data integration is to bring together data from different sources according to the subject area. Two types of integration procedures are performed (a) integration of similar data types from different sources, and (b) integration of different data types per subject areas. While data in different sources could have different aggregation levels, the lowest aggregation level for each data type was adopted to capture the most detailed data [2]. It is important to point out that for this project this process needs to be done in real-time because criminal data is updated every hour.

After the integration process the data needs to be adapted to the selected data mart schema. The most operational data of POSITRON is denormalized for instance the best approach for the schema design is the Star Schema. This schema is a dimensional model composed of a central fact table and set of surrounding dimension tables. A fact table is a specialized relation with a multi-attribute key and contains attributes whose values are generally numeric and additive. A dimension table has a single attribute primary key (usually surrogate) that corresponds exactly to one of the attributes of the multi-attribute key of the fact table [22]. The star schema was selected for our data mart model because access to criminal information needs to be quick and this schema is the simplest database structure that uses denormalized data to provide fast response times [2].

We created a prototype of the data mart using MYSQL to verify the framework of the metadata. Metadata is used as part of the data extraction and load process to map data sources to the common view of their information within the data mart. The metadata contains the rules for data archiving and summarization. Also, it is used as part of the query management process to direct a query to the respective data source [2, 16]. A small view of our data mart's metadata is shown in the Table 2.

The corresponding tasks of the data warehouse manager are to store and manage the data as well as the metadata in the system. There are several commercial and customized software products available for creating all the functionalities of the data warehouse. We recommend the use of Microsoft SQL Server that will also serve as the query manager which performs all operations related with the query process, and as a back-end for the decision-support applications [2]. As a result of the analysis detailed in the above sections Figure 4 presents the suggested architecture for the CISS proposed data warehouse system. It shows the integration of the different data sources of the PRPD that can be used to share real-time information through the Internet and the Intranet of law enforcement agencies within the island and with the United States.

Table 2. Sample Metadata of Data Mart Schema

Name	Primary Key	Aggregation	Load Frequency
Crime	CrimeId + IncidentCode + PersonName	Day, weekly, Month, Quarter	Daily

CONCLUSIONS AND FUTURE STEPS

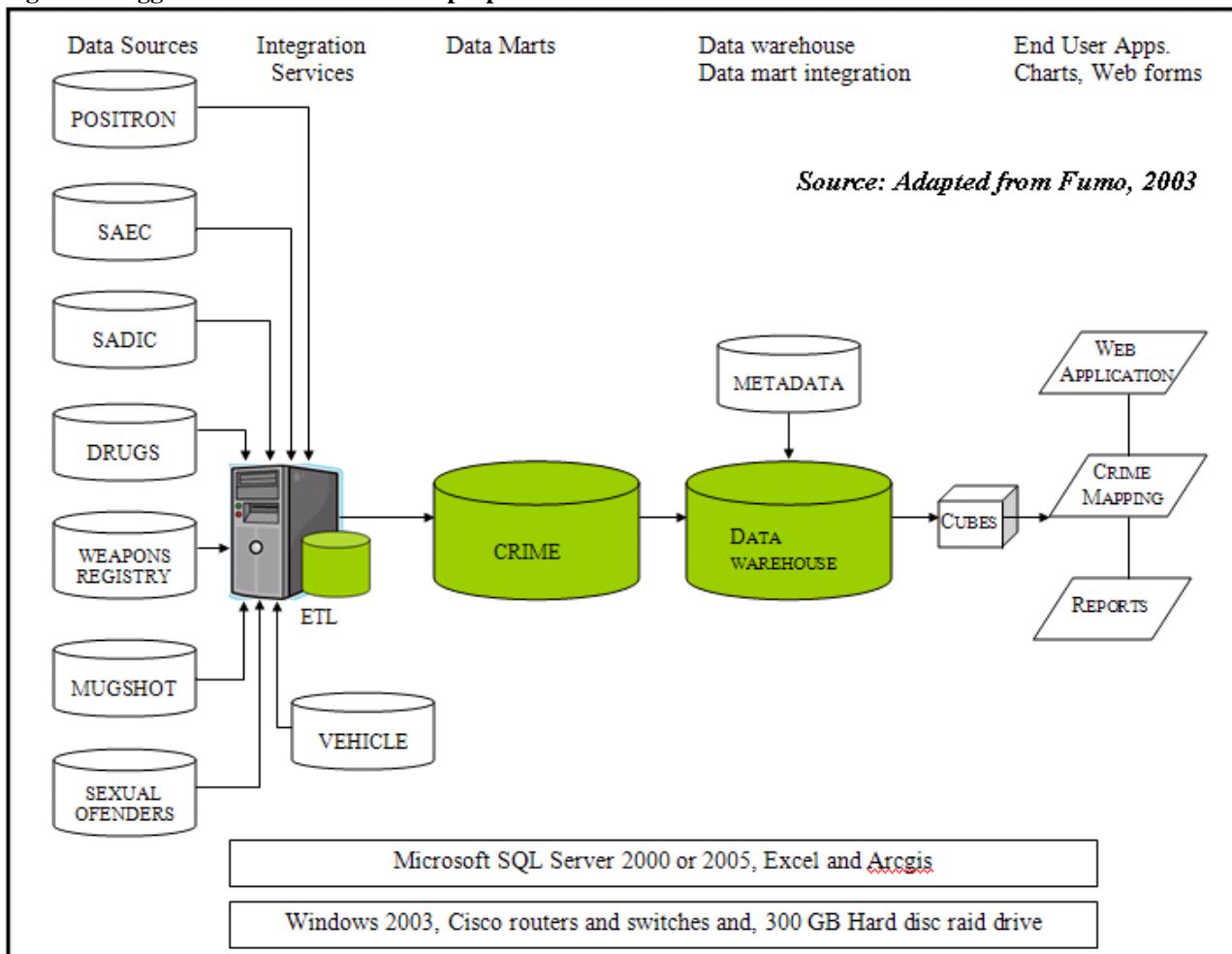
Data warehousing is a valuable alternative to traditional approaches for integrating and accessing data from autonomous and heterogeneous information sources. The warehousing approach is particularly useful when high query performance is desired, or when information sources are expensive or transitory [11]. The purpose of this paper is to describe and recommend the implementation of a data warehouse system in the Police Department of Puerto Rico. The proposed data warehouse system aims to integrate different data source that contains huge amounts of criminal data while improving the collaboration and information sharing between police departments within the island and with the United States. As a result they will have a system that empowers them to promote information sharing between stovepipe information sources and different agencies, and to capture connections between people, places, events and vehicles, based on historical data [5]. The following are future plans derived from our analysis of the proposed system.

1. Connect the data warehouse to a decision support system for tracing patterns, correlations and clustering.

2. Design a Webpage application that is user-friendly for every employee, civilian or officers, to manage efficiently the searching of criminal information.
3. Explore the development of textual mining approaches that support knowledge retrieval from such sources for law enforcement case reports [1].
4. The addition of an integrated multimedia database system to promote information sharing connected to online analytical tools for criminal intelligence analysis.

As Xu and Chen [24] concluded, network analysis is important for understanding the structure and organization of criminal enterprises. Advanced, automated techniques and tools are needed to extract knowledge about criminal networks efficiently and effectively. Our approach constitutes a first step towards this goal since the data warehouse technology includes OLAP tools that are well suited for advanced extractions and analysis. In cybercrime research context, this system can be effectively applied to examine Internet usage patterns, recognition of writing styles in e-mail messages, among others [8]. The remaining challenge is to adapt actual laws as well as prosecution rulings to account for the validity of cybercrime research that may result in more apprehensions and imprisonments of such criminals. Finally, evaluation of these knowledge management and intelligence analysis applications demonstrate their full potential for transforming law-enforcement practices in this age of digital governments [5].

Figure 4. Suggested architecture for the proposed data warehouse for CISS



REFERENCES

1. Atabaksh, H., Schroeder, J., Chau, M., Xu, J., Zhang, J. and Bi, H. (2001), Coplink: Knowledge Management for Law enforcement: Text Analysis, Visualization and Collaboration, *Proceedings of the National Conference on Digital Government Research on Digital Government Research Centre*, Los Angeles, California.
2. Azhar, S. (2005), Information Systems to Support Decision-Making in Construction Owner Organizations: A Data Warehousing Approach, *Dissertation Abstracts International*, 235, UMI No. 3206015.
3. Bruckner, R., List, B. and Schiefer, J. (2001), Developing Requirements for Data Warehouse Systems with Use Cases, *Proceedings 7th Americas Conference on Information Systems (AMCIS)*, 329-335.
4. Chen, H., Atabakhsh, H., Zeng, D., Schroeder, J., Petersen, T., Casey, D., Chen, M., Xiang, Y., Daspit, Y., Nandiraju, S. and Fu, S. (2002a), COPLINK: Visualization and Collaboration for Law Enforcement, *Proceedings of the 2002 annual national conference on Digital government research*, Los Angeles, California, 1 – 7.
5. Chen, H., Schroeder, J., Hauck, R. V., Ridgeway, L., Atabakhsh, H., Gupta, H., et al. (2002b), Coplink Connect: Information and Knowledge Management for Law Enforcement, *Decision Support Systems*, 34(1), 271–285.
6. Chen, H., Zeng, D., Atabakhsh, H., Wyzga, W. and Schroeder, J. (2003), COPLINK: Managing Law Enforcement Data and Knowledge, *Communications of the ACM*, 46(1), 28-34.
7. COPLINK. (2003), Knowledge Computing Corporation Secures \$2 Million Series B Investment From Diamond Ventures To Accelerate National Deployment of COPLINK, [Online], Available: http://www.coplinkconnect.com/pr03_fund1.htm
8. De Vel, O. (2000). Mining e-mail authorship, *Proceedings Workshop on Text Mining on ACM International Conference on Knowledge Discovery and Data Mining (KDD'2000)*.
9. Fumo, T. G. (2003), Health Information Systems Integration: A Data Warehouse Architecture Model for the Ministry of Health in Mozambique, *Doctoral dissertation* [online], Available: <http://urn.nb.no/URN:NBN:no-11501>.
10. Gray, P. G. and Watson, H. J. (1998), *Decision Support in the Data Warehouse*. Prentice Hall.
11. Hammer, J., Garcia-Molina, H., Widom, J., Labio, W. and Zhuge, Y. (1995). The Stanford Data Warehousing Project. *IEEE Data Engineering Bulletin, Special Issue on Materialized Views and Data Warehousing*, 18(2), 41-48.
12. Hauck, R. V. (1999), COPLINK: Exploring usability of a multimedia database application for law enforcement [online], Available: <http://ai.eller.arizona.edu/coplink/publications/nij.pdf>.
13. Hauck, R.V., Chau, M. and Chen, H. (2002), COPLINK: Arming Law Enforcement with New Knowledge Management Technologies, In W. McIver and A. K. Elmagarmid (eds.) *Advances in Digital Government: Technology, Human Factors, and Policy*, Boston.
14. Inmon, W.H. (2005), *Building the Data Warehouse*, 4th Edition, Wiley Publishing, Inc.
15. Kimball, R. and Casetera, J. (2004). *The Data Warehouse ETL Toolkit* (2), NY, Wiley.
16. Krishna, S. J. (2004), *Data Warehousing: Design and Development Perspectives*, India, ICFAI Books.
17. Leon, J., Henry, H., Chen, H., Zeng, D. D., Lin, C. and Chau, M. (2004), Process-Driven Collaboration Support for intra-agency crime analysis, *Decision Support Systems*, 41(1), 616-633.
18. Nemati, H., Steiger, D., Iyer, L. and Herschel, R. (2002), Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing, *Decision Support Systems*, 33(1), 143-61.
19. Primera Hora (2007), “Fraude bancario vía Internet”, press release on March 10, 2007.
20. Ponniah, P. (2001), *Data Warehousing Fundamentals*, John Wiley and Sons.
21. Rodríguez, F. (2007), “Falla la comunicación con Corrección”, *El Nuevo Día*, press release on April 16, 2007.
22. Sen, A. and Sinha, A.P. (2005), A Comparison of Data Warehousing Methodologies, *Communications of the ACM*, 48(3), 80-84.
23. Schroeder, J. (2001), COPLINK: Database Integration and Access for a Law Enforcement Intranet, *National Criminal Justice Reference Services*, NCJ 190988.
24. Xu, J. and Chen, H. (2005), CrimeNet Explorer: A Framework for Criminal Network Knowledge Discovery, *ACM Transactions on Information Systems*, 23(2), 201-226.
25. Zhuge, Y., Garcia-Molina, H., Hammer, J., and Widom, J. (1995), View Maintenance in a Warehousing Environment, *Proceedings of the ACM SIGMOD on International Conference on Management of Data*, San Jose, California, 316-327.