Regional Crime Analysis Data-Sharing with ArcIMS: Kansas City Regional Crime Analysis GIS

Steven Hick & Christopher Grubb
University of Denver
National Law Enforcement and Corrections Technology Center
Crime Mapping and Analysis Program
Denver, Colorado 80208
www.nlectc.org/cmap

Twenty-Third Annual
ESRI International User Conference
San Diego, California

ABSTRACT

Regional crime analysis is gaining recognition as the next frontier of crime analysis. Kansas City's mobile society creates a criminal landscape where offenders can commit serial crimes across several jurisdictions. Due to the compartmented nature of local law enforcement agencies, a crime series may go unrecognized or be delayed in detection by a lack of information regarding crime incidents in neighboring jurisdictions. The purpose of this paper is to highlight how ArcIMS is being used to provide a regional crime analysis application to facilitate data sharing for law enforcement personnel in the metropolitan Kansas City area.

About the University of Denver and the National Law Enforcement and Corrections Technology Center

University of Denver

The University of Denver, the oldest independent university in the Rocky Mountain region — founded in 1864 by Colorado Gov. John Evans — enrolls approximately 9,271 students in its undergraduate, graduate and professional programs.

National Law Enforcement and Corrections Technology Center-Rocky Mountain Region (NLECTC-RM)

Located at the University of Denver in Colorado, NLECTC-Rocky Mountain focuses on communications interoperability and the difficulties that often occur when different agencies and jurisdictions try to communicate with one another. This facility works with law enforcement agencies, private industry, and national organizations to implement projects that will identify and field test new technologies to help solve the problem of interoperability. In addition, NLECTC-Rocky Mountain houses the Crime Mapping and Analysis Program, which provides technical assistance and capacity building to State and local agencies in the areas of crime and intelligence analysis and geographic information systems. Sandia National Laboratories has been designated as a satellite of NLECTC-Rocky Mountain and works in partnership with NLECTC to focus on technologies for detecting and neutralizing explosive devices.

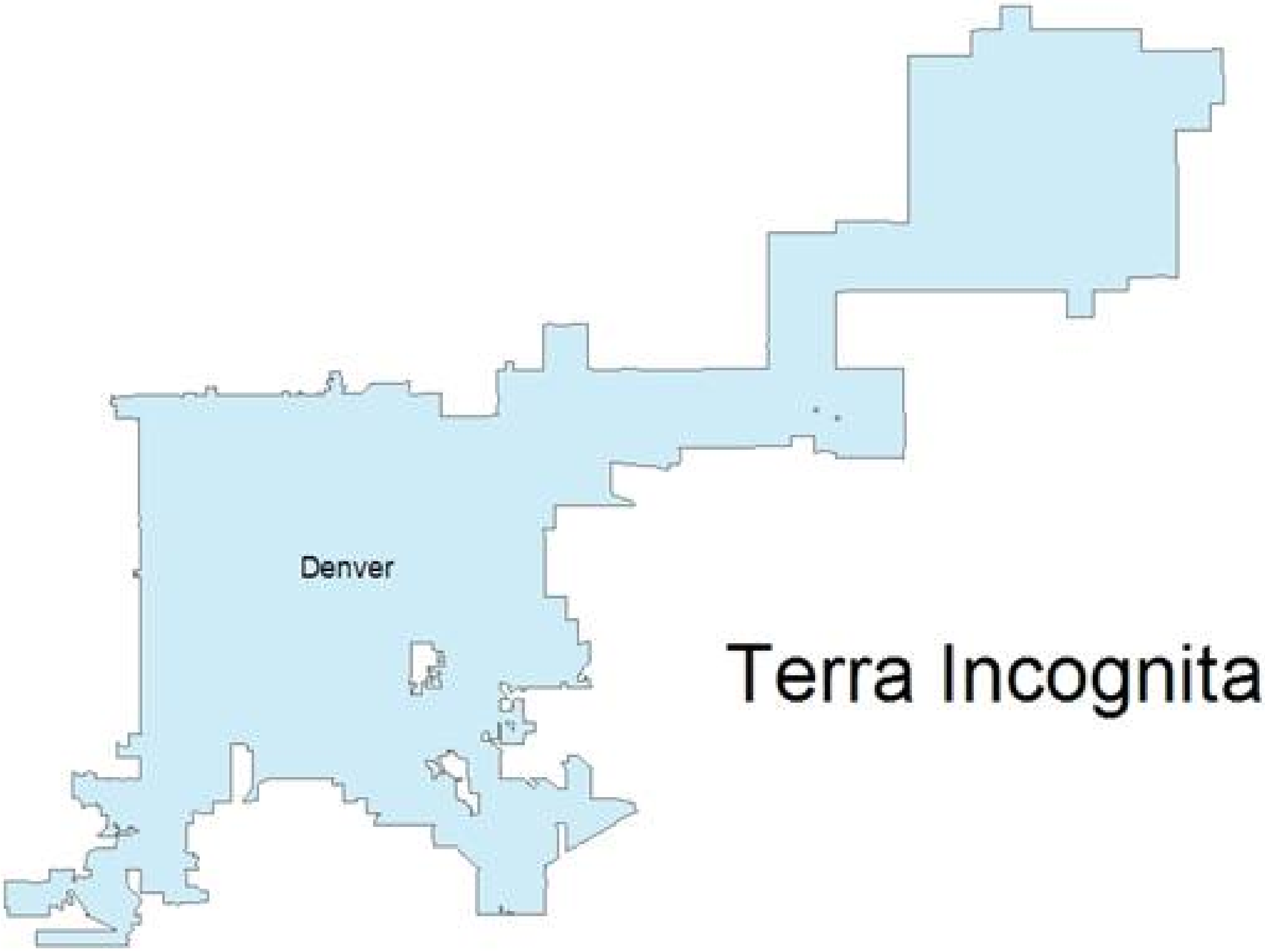
Crime Mapping and Analysis Program (CMAP)

CMAP's mission is to provide technical assistance and introductory and advanced training to local and state agencies in the areas of crime and intelligence analysis and geographic information systems (GIS). GIS includes crime mapping, global positioning systems, automatic vehicle locator systems, and the use of this technology for the electronic home monitoring of community corrections clientele.

To date, assistance has been provided to over 400 local law enforcement agencies and over 500 crime analysts and law enforcement officers.

Challenges of regional crime analysis

Five years of working with law enforcement agencies in the CMAP reveals little cooperation with regards to data sharing between law enforcement agencies. Phone calls and emails are the predominant mode of exchanging data when data is exchanged. Ironically with GIS technology and a plethora of spatial data, most agencies make a map of their city with nothing appearing outside the municipal boundary.



Background

In 1999, metropolitan Kansas City law enforcement agencies initiated discussions for regional crime analysis. Crime analysts in law enforcement agencies often work in a vacuum with regards to information. Data for their jurisdiction is plentiful. Data about similar crimes in neighboring jurisdictions is difficult to acquire. Neighboring law enforcement agencies often do not have the same data collection and storage systems in the form of dispatch and records management systems. Communication between law enforcement agencies is also often difficult. Busy schedules and competing interests for time makes it difficult even for the most organized crime analysts in a region to meet regularly to share information.

The Crime Mapping and Analysis Program (CMAP) at the University of Denver entered into entered into a memorandum of understanding (MOU) with the Kansas City Regional Crime Analysis Committee in 2001.

The Goal

The goal of the KCRCAGIS project is to create an Internet-based crime mapping system to facilitate the analysis and exchange of crime data across city, county, and state boundaries.

Project Participants

Agencies from the following Kansas and Missouri cities participated in the prototype project:

- Kansas City, KS
- Johnson County, KS
- Overland Park, KS
- Lenexa, KS
- Shawnee, KS
- Kansas City, MO
- North Kansas City, MO
- Independence, MO
- Blue Springs

Project Organization

Three committees were created when the KCRCAGIS project team first convened.

1. Policy

   The policy committee is made up of the leadership of the KCRCAGIS project team. Participants are lieutenants and captains in their respective agencies and generally "have the ear" of the chief of police.
2. Technical

   The technical committee is comprised of interested KCRCAGIS participants that have a vested interest in the data and the technology. Members of this committee are crime analysis and GIS professionals with a working knowledge of GIS.
3. Crime analysts

   The crime analyst committee is obviously comprised of crime analysts. This group provides insight as to how the KCRCAGIS project and Internet-based data sharing help their cause and mission.
4. CMAP staff

   CMAP staff from the University of Denver provides project oversight, communication, programming and data acquisition skills to the project.

Future Participants

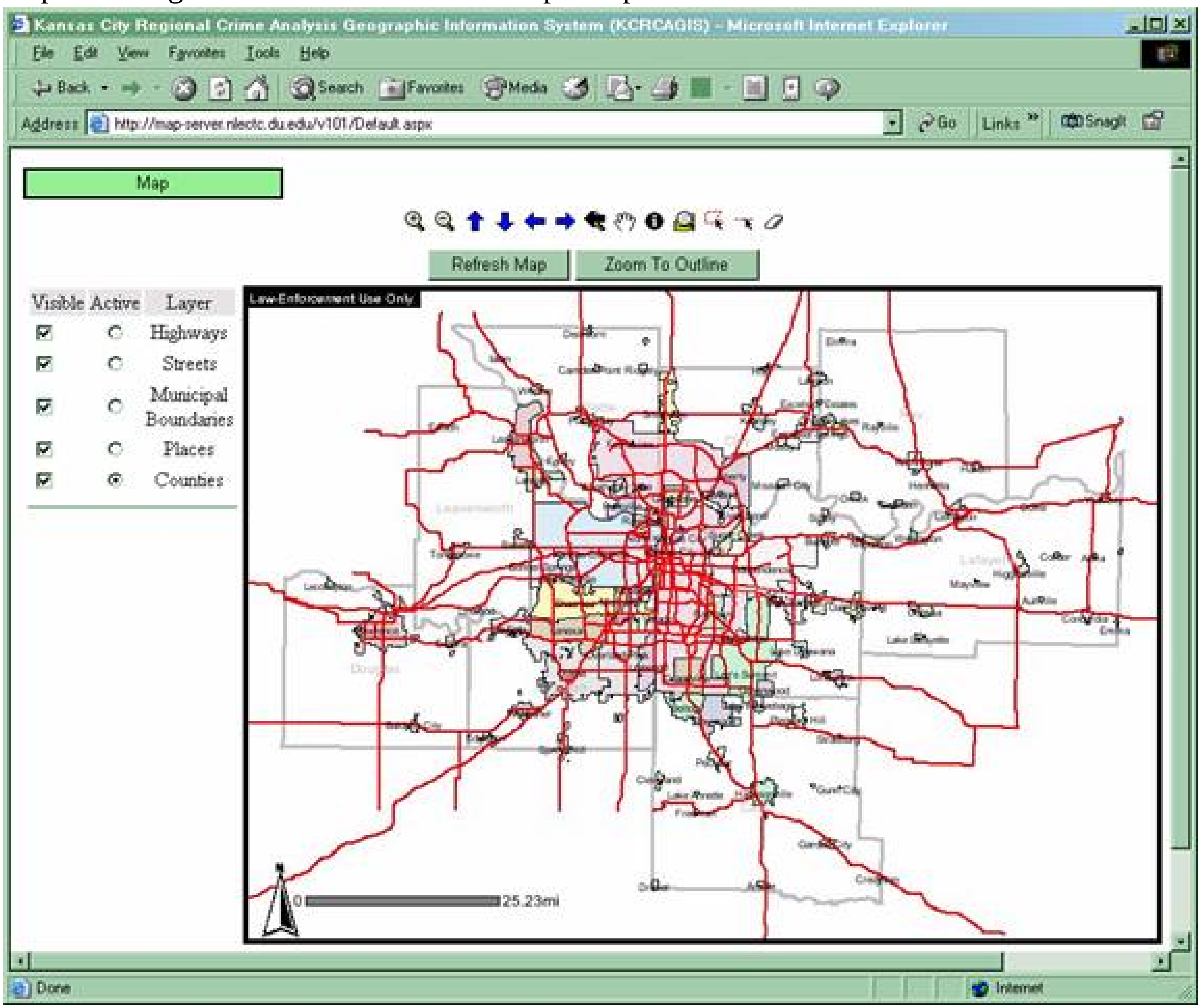
The metropolitan Kansas City area has a population of over two million people. The KCRCAGIS project covers 10 counties in two states representing 85 local law enforcement agencies in the metropolitan area.

Data Sharing

As stated previously, the goal of KCRCAGIS is to provide data sharing capabilities in a secure environment via the Internet. Data is shared across city, county, and state boundaries. Sharing data is not trivial because every agency participating has different reporting standards and different records management systems (RMS). While their location varies, their geographic data requirements are similar.

Base map data

The Mid-America Regional Council (MARC) provides the street centerline base map. In April 2003 the MARC E-911 street centerline file went "live." The E-911 street centerline file was described by one source (remaining anonymous) as a "Frankenstein" map. The map is being sewn together from the parts of several other maps. Each city or county in the region submits their street centerline data to a contractor the merges and project the data into a seamless street map of the region to be redistributed to the participants via MARC.



Crime Data

Crime data sharing is a bigger problem because every KCRCAGIS participant collects, stores, manipulates and analyzes their data differently. This has been without questions the single greatest challenge facing the KCRCAGIS project members.

Part I Crime

The KCRCAGIS members first had to agree upon the types of crimes to collect to share. Initially Part I crimes were the only crimes to be included in the database. After much discussion the following crimes were identified for inclusion in the database. This list is generated via the data submission of each agency.

1. Arson
2. Assault (aggravated assault)
3. Auto burglary
4. Auto theft (motor vehicle theft)
5. Burglary
   Commercial burglary
   Residential burglary
6. Kidnapping/abduction
7. Larceny/Theft
8. Homicide
9. Sex offense

Forcible Rape
Forcible Sodomy
Sexual Assault with an Object
Forcible Fondling

10. Robbery

Data Extraction

An initial decision had to be made about how to extract data from nine different records management systems where the data is being stored in nine different formats and then merge the data into a database where records could be compared across city, county and state boundaries. Every jurisdiction has its own way of codifying data. Across state lines descriptions of crimes vary.

To get the process started we reviewed all crime codes as required by the National Incident Based Reporting System (NIBRS). NIBRS is fairly inclusive with regards to crime, there are some fields missing. Most importantly, address does not exist in NIBRS reporting.

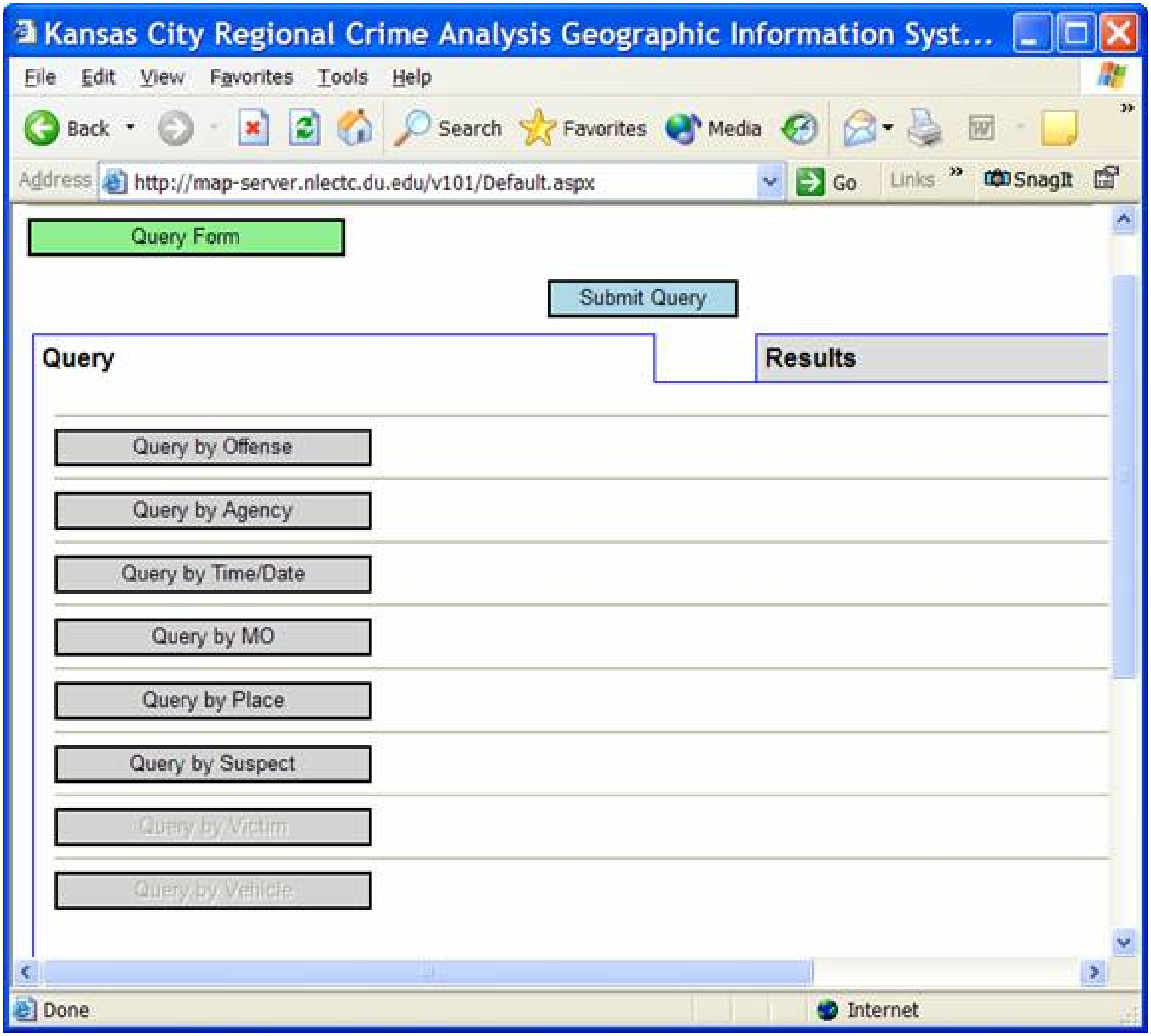
For KCRCAGIS, the express E-NIBRS was coined to express the enhanced (E-) set of attributes being collected for KCRCAGIS that do not appear in NIBRS.
Codes used by each participating agency converted into literal values.
To get around the problem of the differences in codifying crimes, data was converted from the format of the source agency into literal values that everyone could understand. To make this possible, routines were written that varied from agency to agency to extract data from its native system in Microsoft Access tables.

Software

Data extraction and data queries in KCRCAGIS

The crime analysts in the Kansas City region determined that the data they wanted to query included offense type, originating agency, date and time, modus operandi (MO), and suspect. A query button for place has been included to query by municipality for data from an agency that may have jurisdiction over several places. For example, Johnson County Sheriff has jurisdiction over several smaller municipalities in the county.

Placeholders have been designed in the query screen for victims and vehicles when that data is widely available.



After building a query, the crime analyst can click on the Results tab to get a list of the incidents meeting their criteria.

When satisfied with the results, the analyst can build an Extensible Markup Language (XML) file for downloading into a desktop product, or map the results to appear on the top half of the browser in the map window. The analyst can also download a shapefile consisting of the geocoded crime locations corresponding to the records selected by the query.

Database Design

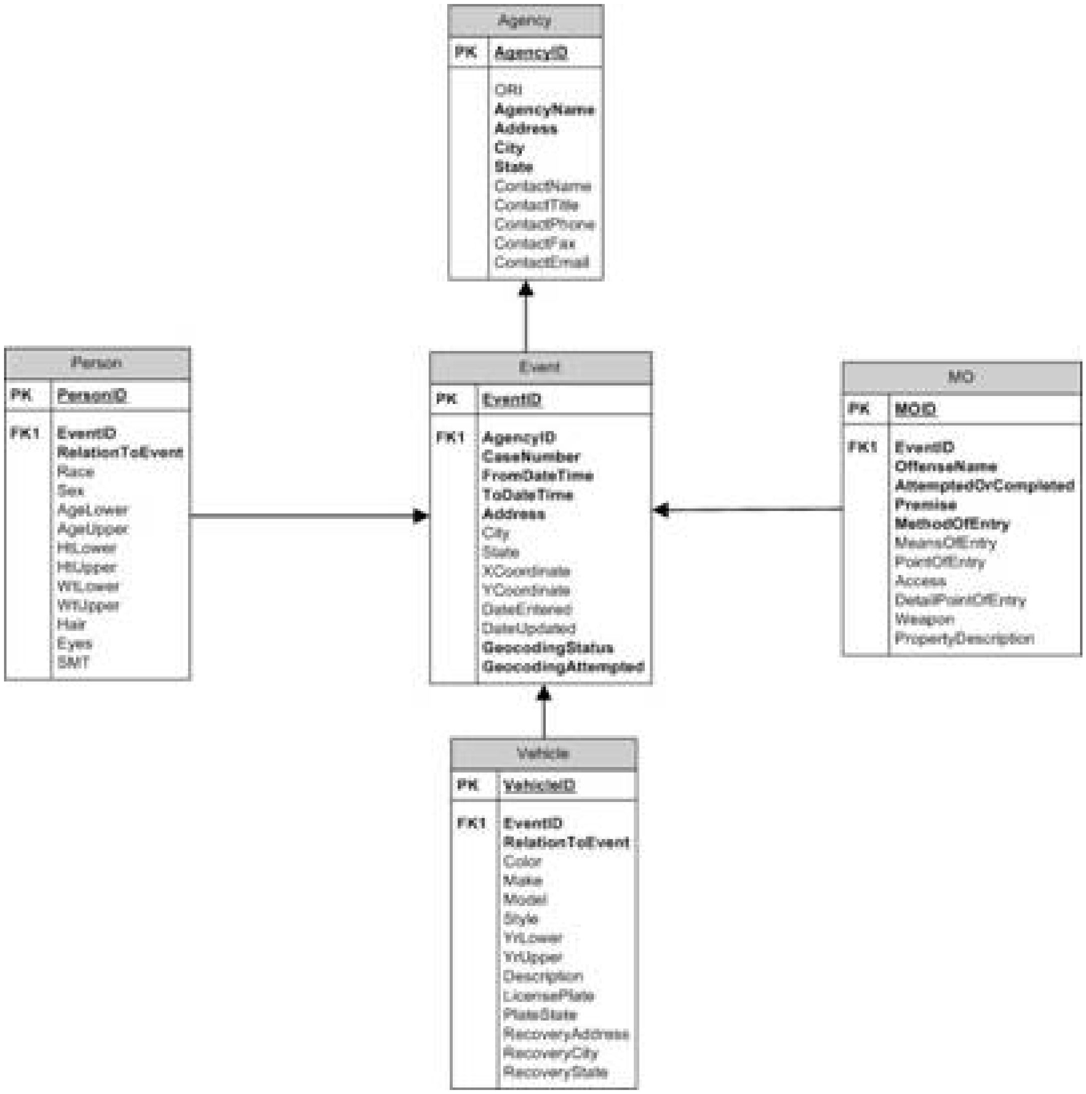
The database design consists of five tables: an Agency table, an Event table, an MO

(*modus operandi*) table, a Person table, and a Vehicle table. The agency table stores information about each agency participating in the system. This information includes contact information, city, state, and the ArcSDE geocoding service to be used in geocoding the agency's data.
The Event table stores a unique identifier, a 32-bit integer that is used as a primary key and serves as a foreign key in the other tables to link the records together. Event data also includes the local case number for the incident, the start and end time/date, the address, city, state, and X- and Y-coordinates. The table also includes a character field signifying whether the address has been matched, and a flag bit to indicate whether geocoding has been attempted for the record.

The MO table stores attributes such as the offense type, whether the offense was attempted or completed, the type of premise, method of entry (forced or non-forced), the means of entry (pry bar, etc), the point of entry (window, door, etc), access (front, rear, side, etc), a detailed point of entry field for more narrative information regarding the entry, weapon used, and a short description of property taken. Each MO record is assigned a unique identifier, and is also linked to the Event table via a foreign key relationship.

The Person table stores information regarding suspects and victims, including race, sex, age range, height range, weight range, hair color, eye color, and scars-marks-tattoos (SMT). The persons are classified by their relationship to the event (i.e. suspect, victim, etc). Each Person record is assigned a unique identifier as well as a foreign key relating the record to a record in the Event table.

The Vehicle table stores identifying information about vehicles used, stolen, or damaged during the incident. Information includes the color, make, model, body style, year of manufacture, a narrative description field for other identifying information, license plate, state of registration, recovery address, recovery city, and recovery state. Again, each vehicle record is given a unique identifier, and a foreign key relating the vehicle record to a specific event.



Data Cleaning/Conversion

Data import and cleaning was accomplished using the Data Transformation Services (DTS) functionality available in Microsoft SQL Server 2000 Standard Edition. When the system is deployed, the data import process will be automated.

Geocoding

A custom geocoding application was developed using ESRI ArcObjects 8.2 and Microsoft Visual Basic 6.0. This application queries the database for records with addresses that have not yet been matched, and geocodes them using the appropriate ArcSDE geocoding service. Each agency is assigned a geocoding service containing street centerlines corresponding to its geographic jurisdiction. A point is inserted into the ArcSDE crime locations layer and assigned an event identifier corresponding to the Event table primary key, linking that record to the crime attributes database tables. X- and Y-coordinates are inserted into the crime attributes database in the record corresponding to the geocoded point. This application is intended to run automatically on a scheduled basis.

ArcIMS

A map service was designed and deployed using ArcIMS 4.0 to display the base map layers and geocoded incident locations. The base map layers include street centerlines, county boundaries, municipal boundaries, and major highways for the full ten-county study area

SQL Server/ArcSDE

The database component consists of two databases: a geodatabase developed using ArcInfo 8.2 and ArcSDE 8.2, and a separate database used to store the crime attribute data (described above). The geodatabase stores the base maps including street centerlines, municipal boundaries, county boundaries, and highways. The geographic data coverage includes Johnson, Leavenworth, Wyandotte, Douglas, Jackson, Clay, Platte, Cass, Ray, and Lafayette counties. The geodatabase also contains separate street centerline files for each city and county involved in the prototype project to be used in geocoding the data supplied by each agency. Separate centerline files were used to minimize the chance of duplicate street names and numbering systems across the ten-county area. Geocoded incident locations are also stored in this geodatabase.

.Net

A web browser-based map display and query interface was designed and implemented using Microsoft Active Server Pages (ASP.Net) and the ESRI ArcIMS ActiveX connector, and deployed using Microsoft Internet Information Services 5.0 as the web server.
All code for the application was written using Microsoft Visual Basic.Net. The mapping interface enables the user to pan, zoom, identify features, and select crime incidents. The query interface allows the user to construct queries based on offense type, reporting agency, location, suspect/victim physical characteristics, vehicle description, or a combination of all or some of these factors. The user can also download a subset of data in XML or ESRI shapefile format.

Hardware

For development purposes, the ArcIMS and ASP.Net application was hosted on a 1.5ghz Intel Pentium IV workstation with 512Mb of RAM. The SQL Server/ArcSDE portion was hosted on a server with one 2.0ghz Intel Xeon processor, a 100Gb RAID 5 array, and 2Gb of RAM.

Client access

The Midwest HIDTA provides a region-wide virtual private network (VPN) with Smartcard access for user authentication. The ASP.Net application will be accessible only through the VPN, with no access from the general Internet. The application can be accessed using any recent version of a standard Internet browser.

Conclusions/Lessons Learned

The most troublesome aspect of the project from an implementation perspective has been the extraction and transmission of data from the participating agency's records management systems. Each agency uses a different RMS vendor, and each agency has unique policies dictating the frequency of data entry and extraction. In some agencies, the primary responsibility for data extraction rests with the IT staff, while in others the task is performed by crime analysts or sworn officers. The diversity of data extraction techniques initially led to a diverse range of file formats, including text files, dBase IV tables, Microsoft Access databases, and Microsoft Excel spreadsheets. To streamline the automation of the data import process, a common file format with standard field names and structure was needed. For the prototype project, Microsoft Access was chosen for data transmission. This was primarily driven by the widespread availability of the Microsoft Office suite in the participating agencies, and the fact that the file format enforces a predictable structure and is self-contained.

Geocoding has presented another challenge. To date, all geocoding has been performed centrally after the data was submitted. Significant problems with this approach included a low 'hit' rate (approximately 80% overall) and the difficulty in re-matching several thousand unmatched addresses by persons unfamiliar with the local jurisdictions. In addition, low administrative overhead was one of the original system design goals, meaning that there would not be a full-time administrator available to interactively re-match unmatched addresses. However, there are also complications involved in using locally geocoded data, such as a diversity of projections and coordinate systems, and differences in the accuracy of local street-centerline data. Local geocoding also potentially increases the workload of the IT or crime analysis staff, presenting a barrier to participation. Another potential problem with local geocoding stems from the fact that not all agencies currently have access to geocoding software.

Steven Hick
GIS Director
Department of Geography
University of Denver
shick@du.edu

Christopher Grubb
Graduate Research Assistant
Department of Geography
University of Denver
cgrubb@du.edu