International Association of Crime Analysts (IACA)

# RMS Technical Requirements for Crime Analysis

Standards, Methods, & Technology (SMT) Committee White Paper 2013-01
August 2013


IACA Standards, Methods, and Technology Committee, Technology Chair:
Matthew Harris, Crime Analyst, Sonoma County Sheriff's Office
IACA Standards, Methods, and Technology Committee Chair:
Jason Elder

IACA SMT Technology Subcommittee Technology Member:
Bryan Hill, Crime Analyst, Glendale Police Department
Travis Taniguchi Ph.D., Senior Research Associate, Police Foundation

Subject Matter Experts:
Stacy Belledin, Crime Analyst II, Lakewood Police Department
Tony Berger, Detective Sergeant, Pierce County Sheriff's Department
Rita De Lucchi, Crime Analyst, Berkeley Police Department
Heather Lane, Crime and Intelligence Analyst, Mansfield Police Department
Abigail Luczon, Sr. Crime Analyst, Glendale Police Department
Jim Mallard, Crime Analyst, Winter Park Police Department

Editors:
Todd Wiles, Crime Analyst Project Coordinator, Cleveland Police Department
Erika Wiles, Crime Analyst II, Akron Police Department



Suggested Citation:
International Association of Crime Analysts. (2013). RMS Technical Requirements of Crime Analysis (White Paper 2013-01). Overland Park, KS: Author.


IACA
INTERNATIONAL ASSOCIATION OF CRIME ANALYSTS

# About the IACA Standards, Methods, and Technology Committee

Through the Standards, Methods, and Technology Committee (SMT), the International Association of Crime Analysts (IACA) is committed to a continuing process of professionalization through standards and knowledge development. In 2011, the IACA chartered the SMT Committee for the purpose of defining "analytical methodologies, technologies, and core concepts relevant to the profession of crime analysis."[1] This document is part of a series of white papers produced by the SMT committee that accomplishes this purpose. The methodology for formulating the positions reflected in the white paper series includes: 1) development of a draft paper through in-depth meetings and discussions of Subject Matter Experts,[2] 2) review and feedback by the IACA Executive Board, 3) review and feedback from an independent editor with knowledge of crime analysis, and 4) review and feedback by IACA members facilitated through the IACA website (www.iaca.net). Questions about this process should be directed to the chair of the SMT Committee at SMT@iaca.net.

# Overview

This document has been prepared by the IACA SMT Committee to address the requirements of records management systems (RMS) as used by law enforcement agencies for crime analysis. A RMS is an agency-wide system that provides for the storage, retrieval, retention, manipulation, archiving, and viewing of records, documents, or files pertaining to law enforcement operations.[3] The primary intended use of this paper is to provide crime analysts, administrators, and other key decision-makers with fundamental RMS requirements for successful crime analysis.

This document explains some of the key applications and functionality needed to support and facilitate sophisticated crime analysis. There are recommendations throughout the paper with regard to workflow within the agency, questions to ask vendors and information technology (IT) personnel, and key things to consider when upgrading or replacing a RMS. At the end of this paper is a list of relevant literature covering these topics in detail; throughout, there are footnotes that point directly to key readings. A checklist that can be used to evaluate RMS packages from a crime analysis perspective can be found in Appendix A. Principles discussed in this document also apply to other systems such as online citizen crime reporting, jail management, traffic citation/collision reporting, and mobile applications for tablets or phones.

**Analyst involvement in procurement and implementation**

Crime analysis units are typically responsible for retrieving data from the RMS for the purpose of reports, analysis, and investigation; therefore, analysts often have valuable input as to how this data should be collected and stored. Analysts use this data in administrative and strategic analysis to lay the quantitative foundation for problem-oriented projects and intelligence-led initiatives. This data is crucial during tactical analysis to identify patterns and trends, and to support investigations by researching people, places, and properties. Poor data entry and lack of proper data flow can hinder investigations and analytical data retrieval. These problems can be mitigated by having analysts involved at the earliest stages of procurement.

---

[1] Drawn from the mission statement of the Standards, Methods, and Technology Strategic Plan (April 2011).
[2] Subject Matter Experts are identified by the Standards, Methods, and Technology Committee based on their special knowledge demonstrated through publications, presentations, and practical experience and their willingness to participate.
[3] See, generally, IACP/COPS Technology Assistance Program (2004).

During the procurement phase, analysts can assist decision-makers by assessing each vendor to ensure that the selected system meets the agency's current and anticipated needs. In addition to the items covered in appendix A, analysts should develop a checklist or consistent and objective set of questions to ask of all software vendors. Examples include:

1) Does the RMS have complex data validation rules that increase data accuracy and quality without complicating the data entry process?

2) When performing data entry, are the steps for selecting existing master records (names, addresses, vehicles, property) simple and straightforward yet fail-proof?

3) Is the system built with open architecture that allows users to view and query against data and documents using third-party search tools?

4) Does the system store geographic coordinate information regarding calls for service, incidents, traffic collisions, contacts with people and vehicles, and property?

5) Are the coordinates stored in an easily useable format? Some vendors store coordinates in a single field that must be parsed, while other vendors use coordinates that have to be divided by 100 before they can be used outside the RMS.

6) Does the system allow for re-geocoding of data if the agency receives updated reference files (centerlines, parcel files, etc.)?

7) Can the RMS provide answers to agency-specific questions in a way that is acceptable? Generally, analysts have a unique insight regarding questions that law enforcement personnel need answered. Does the RMS provide solutions to these problems?

Analysts can also help to ensure the proper setup of the RMS by verifying look-up tables, required fields, mapping infrastructure and automated reports. Taking time to thoroughly review the data configuration in advance will save time in cleaning data, geocoding, and producing reports, leaving more time for investigations and analysis in the future.

**Recommendation 1: Analyst Involvement in Procurement and Implementation.** The crime analyst should be involved with the procurement and implementation process, especially when connecting to and setting up the RMS mapping infrastructure, building and modifying look-up tables, and determining required fields.

## Data Management

The primary purpose of the RMS is to collect and store crime-related data. The best systems have efficient processes for entering, storing, and retrieving data while at the same time maintaining data quality. Systems that collect and store this data in an efficient and accurate manner drastically improve analytical capabilities. Good systems also store the data in a way that is accessible outside of the system in an easily comprehensible structure. The RMS is one of the main sources that crime analysts use to monitor and analyze crime within their jurisdictions; therefore, they must have a firm grasp on how and where the data is stored within the RMS.

### Open Architecture

Analysts often use third-party tools that utilize information in the RMS through a database connection. Examples of these tools include geographic information systems (GIS), link analysis tools, reporting and statistical software.

The RMS should be designed with an open architecture, meaning that the records are accessible to third-party software packages. It is crucial to avoid systems that store information in proprietary file types or inaccessible locations. Although systems usually provide built-in reporting functionality, these reports frequently fall short of meeting the needs of most agencies. Data should be available via standardized database connectivity protocols such as Open Database Connectivity (ODBC),[4] Object Linking and Embedding Database (OLE DB),[5] Java Database Connectivity (JDBC),[6] etc.

If third-party tools will be used to search or report on RMS data, analysts should question whether the data types in the RMS support their needs. For example, images stored in a database field as file path references are easier to report on than images stored in binary field types. Binary large object (BLOB) field types should be avoided unless absolutely necessary and not used to store images or large unstructured text.[7] Unstructured text elements (such as narratives) are easier to display and filter when stored in text-based fields. Analysts must be able to ask specific questions to determine if the vendor's database structure supports the tools at their agencies' disposal.

**Recommendation 2: Open Architecture.** Systems should allow access for third party tools via standard database connectivity protocols and not use unsearchable field types to store images or large unstructured text.

**Data Dictionary**

The RMS should provide an Entity Relationship Diagram (ERD)[7] showing how database fields are related and joined together. Additionally, the RMS should provide a data dictionary[8] that includes table and data element definitions, and a schema[9] with technical notes for each element, such as type, size, primary and child key(s), and field description. Clear language must be used regarding the data stored in each field, as well as which fields are indexed for faster searching. There should be sections for identifying calculations, look-up tables, and data dependencies. System documentation should include information on entity maintenance, such as dates of creation, last modification, termination and source. The data dictionary and ERD should also clearly delineate fields that are no longer used by the system.

Data management personnel should have the ability to update the data dictionary, and the system should direct them to update documentation when they make changes to data fields or structure. There should also be a comments section for users to communicate special messages, such as, "Data is no longer used," or "Data verified as of October 1, 2012." This feature can be particularly useful for notations regarding legacy data imported during implementation. It is important for departments to make the data dictionary available to crime analysts and other personnel who use the data for analytical purposes. Superior systems will make this collection of RMS metadata accessible as an online or network resource, organized in a logical and easily-searchable format.

**Recommendation 3: Data Dictionary.** Systems should provide an Entity Relationship Diagram (ERD) showing how database fields are related; a schema with technical notes for each element; and field descriptions that explain in clear language what data is stored in each field. The ERD and data dictionary should be available to

[4] More information on ODBC can be found at: http://support.microsoft.com/kb/110093.
[5] More information on OLE DB can be found at: http://office.microsoft.com/en-us/access-help/about-ole-db-and-data-links-HP005260660.aspx.
[6] More information about JDBC can be found at: http://www.oracle.com/technetwork/java/overview-141217.html.
[7] Chen (1976).
[8] A simplified explanation of a data dictionary can be found at: http://en.wikipedia.org/wiki/Data_dictionary.
[9] A simplified explanation of a data schema can be found at: http://en.wikipedia.org/wiki/Database_schema.

crime analysts and other personnel who use the data to create informational or analytical products. Systems should make metadata accessible as an online or network resource in an easily-searchable format.

**Visible Control Source**

ERDs and database schemas can often be cumbersome to navigate and difficult to understand. In order to simplify the process of understanding the database structure, the RMS should provide the ability to see the database table and field names associated with controls in the application. This feature should also provide the ability to identify values stored as codes and clearly identify the look-up tables necessary to decode them. Ideally, the RMS should display this information in a discreet way, such as producing a dialog box only when the user holds the shift key and hovers over the field in question. This open architecture design allows analysts to quickly and easily grasp the data structure, allowing more time to focus on analysis. This feature will also help users communicate with data management personnel and/or vendors if they experience trouble with the system. Figure 1 shows an example of displaying the control source of an address field within a RMS.

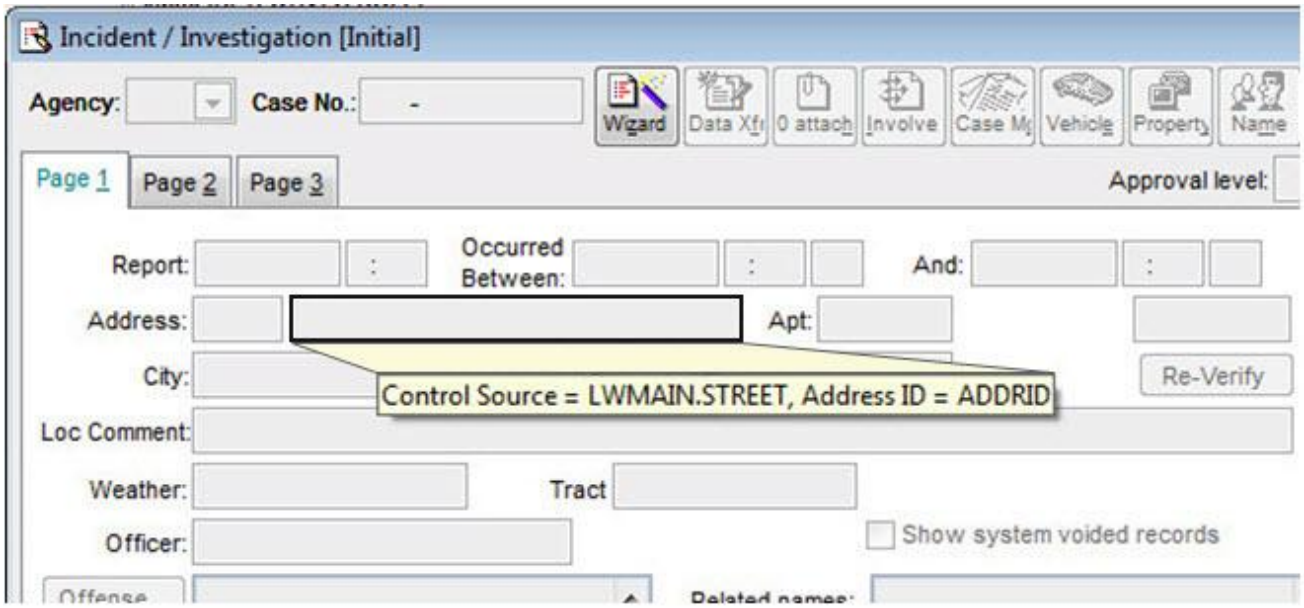


**Figure 1.** Sample RMS "Show Control Source" tool. The pop-up appears as the user hovers over the field in question.

**Recommendation 4: Visible Control Source.** Systems should provide the ability to see the database table and field names associated with controls inside the RMS. This feature should be discreet and provide the ability to identify where look-up tables are stored and how key values are defined.

**Data Types**

The RMS should not be limited to just the criminal event; it should be open to all aspects of law enforcement. It is recommended the RMS be a complete system, able to store a variety of records that are important to an agency. The breadth and scope of the RMS will vary depending on the needs of the agency; for example, some agencies may be fine with simply adding a booking/jail module to the RMS system, but others will need a full jail management system. Some systems are rigid and don't allow for customization while others are flexible, allowing agency-level developers to create their own modules with tables, forms, and reports. Table 1 shows common master indices and data modules that are contained within a typical RMS; for comprehensive documentation,

please refer to *Standard Functional Specifications for Law Enforcement Records Management Systems Version II*.[10] Systems with additional data types would have to be evaluated independently, as they are likely not based upon national standards.

| Typical Master Index Records | Typical Modules | |
|---|---|---|
| Names | CAD events (calls for service) | Pawns |
| Vehicles | Incident reports | Civil process |
| Property | Investigative case management | Protection orders and restraints |
| Location/Address | Property and evidence management | Permits and licenses |
| Organizations | Arrests | Equipment and asset management |
| Employees | Booking/Jail management | Fleet management |
| | Collision reports | Personnel management |
| | Citations | Internal affairs |
| | Field contacts | Known offenders |

**Table 1.** Typical master index records and modules for modern records management systems

# Data Collection

### Importing Legacy Data

During the RMS procurement and implementation process, the agency will need to evaluate the pros and cons of importing their legacy data. Most crime analysts are intimately aware of data quality issues and can be very beneficial during this evaluation. It is particularly important to assess and understand how imported legacy data will impact the new system's master indices. Agencies with data quality personnel may be able to mitigate some of this concern by diligently working to merge duplicate master records. It is possible that after thorough review of legacy data, agencies may find their historical information is not of sufficient quality to bring into the new system. Ultimately, the decision hinges on the possibility of compromising the new system's effectiveness versus retaining seamless access to historical data.

**Recommendation 5: Importing Legacy Data.** Agencies need to evaluate legacy data in advance of the decision to import it into a new system. Crime analysts are intimately aware of data quality issues and should be involved with this evaluation process.

### Importing Computer Aided Dispatch (CAD) Data

Records Management Systems and CAD systems are typically separate, with separate purposes. CAD systems are designed for managing real-time dispatch events, tracking personnel, and keeping a log of activities as they occur in an emergency call center. RMSs are designed to record detailed records of incidents, persons, associations and property. RMS data is typically entered into the system after an incident has concluded, allowing for facts to be collected through observations and interviews.

It is not necessary for the CAD system and RMS to be from the same vendor; however, it is necessary for the two systems to communicate. CAD systems contain details from the initial police response that are valuable to the historical record. Data transferred from CAD to RMS can eliminate the need for duplicate data entry and save valuable time during the report-writing process. Agencies should consider a RMS with the ability to import data

[10] Law Enforcement Information Technology Standards Council (LEITSC), 2003.

from a modern CAD system regardless of the vendor. While there are some benefits to having your RMS and CAD systems from the same vendor, agencies should not necessarily assume this is the best solution. Agencies should take an in-depth look and weigh the costs and benefits of having separate vendors versus a single vendor for these systems.

Adding CAD data to the RMS gives analysts as well as officers the ability to see the full picture beginning with the initial police involvement. CAD data should be transferred or linked to the RMS once the dispatch event is complete and should contain the following data:

- Date, time (including shift), and location (including address, beat, reporting district, and coordinates)
- Initial & final event type
- How received (self-initiated, dispatched, or otherwise)
- Primary officer name and identifier and assisting officers' names and identifiers
- Dispatch/arrival/cleared date/time stamps
- Easily understandable comments, delimited by a date/time stamp and carriage return

CAD systems may also include system-generated comments that indicate actions taken by dispatchers at a given time, such as searching a person or vehicle through federal data systems. These comments are typically not written in plain English but instead are expressed as computer codes. Dispatcher actions can often be insightful when reviewing event records and they should be included in the data transfer from CAD to RMS. If system-generated comments are included in CAD call narratives, they should be presented in the most readable format and clearly delineated into a section separate from user-entered comments.

CAD systems often have the ability to place an address on a map through the process of geocoding,[11] which allows responding officers and dispatchers to clearly see the location of an event. Geographic information, including coordinates, should be imported into the RMS during data transfer from CAD. Geography is a vital element to understanding crime; superior systems will capture and geocode all locations including those from the dispatch record. CAD systems may also have the ability to search the RMS for prior police contacts with entities involved in events. Ideally, a RMS with an integrated CAD should automatically identify entities that were searched during a CAD event and associate this activity to the master record. This functionality will thereby store information for every CAD event in which an individuals name has been queried, leading to valuable involvement history.

**Recommendation 6: Importing Computer Aided Dispatch (CAD) Data.** Agencies should evaluate RMS and CAD systems independently, as packaged options may not always be the best choice. CAD data should be transferred or linked to the RMS once the CAD event is complete. The import should include all relevant data elements with easily understandable comments, delimited by carriage returns and date/timestamp. System-generated comments should be separated from user-entered comments. Addresses and associated geographic information should be transferred to the RMS; ideally, systems with an integrated CAD should automatically identify master name records that were queried during a dispatch event.

**Data Entry**

Many data entry errors and data quality issues are caused by processes that are overly complex and confusing. RMS data entry should be straightforward and understandable, easy to navigate, and customizable enough to

[11] Kennedy (2001) defines geocoding as assigning x, y coordinates to tabular data such as street addresses or ZIP Codes so that they can be displayed as points on a map.

reflect differences in local procedures and terminology. RMS software must also be comprehensive enough to capture necessary data elements. Good systems have a data entry tool that is separate and more streamlined than the interface used by those who need to review reports. This data entry tool is often, but not always, made for mobile computers. The mobile interface should be designed with special attention paid to the size and location of buttons that can cause inadvertent data submission.

A potential model for a data entry interface would be a computerized tax filing program that prompts users with a customized list of questions based upon previous responses. These programs allow users to answer questions regarding their specific situations while navigating them through a complex series of steps. They are visually pleasing, easy to follow, and have an automated review process before submission. An ideal RMS will navigate the user through this process in a timely manner without presenting confusing prompts. Additionally, a good system will control data quality with clear input parameters or programming techniques, which limit choices using proven business rules. For example, an agency that desires a robust *modus operandi* (MO) section for theft from motor vehicle incidents could also limit the level of detail for shoplifting incidents. A superior system will be flexible enough to capture the appropriate level of detail required for insightful analysis. This will ease many frustrations with the data entry process, resulting in better data quality.

The RMS should comply with regional and/or national crime reporting standards. For example, systems implemented in the United States should comply with either Uniform Crime Report (UCR)[12] or National Incident-Based Reporting System (NIBRS)[13] submission criteria. These standards are complex and require specialized validation rules to ensure that they are implemented correctly. Both the UCR and NIBRS place special emphasis on the order (i.e., hierarchy) of criminal charges associated with an incident. The order needs to be updated each time a charge is added or removed from a report. Furthermore, the system should allow agencies to easily make changes to statutes or offense tables to reflect legislative changes in the law at the municipal, state and federal level.

**Recommendation 7: Data Entry.** Data entry should be clearly understandable, easy to navigate, and customizable enough to reflect differences in local procedures and terminology. The data entry interface should be separate from the interface used to review crime reports. The interface designed for use on mobile computers should be laid out so that the size and location of buttons prevent inadvertent data submission. Attention should be paid to ensure that this interface collects the data required for regional and national reporting standards, such as the UCR and NIBRS in the U.S.

**Data Validation and Quality Control**

Data validation is essential for quality control; however, excessive validation rules can complicate the data entry process. The ideal RMS should rely upon programming validation rules that increase data accuracy and relieve end users of those tasks. Examples of good validation controls include the ability to prevent physical impossibilities, such as future dates of birth, occurrence chronology (to/from dates) in reverse order, start and end dates/times later than the CAD event date/time, and questionable entries like extreme heights and weights. Some fields will be required based only upon specific circumstances; for example, a robbery should require property fields be

[12] More information about UCR reporting requirements can be found at: http://www2.fbi.gov/ucr/handbook/ucrhandbook04.pdf.
[13] More information about NIBRS reporting requirements can be found at: http://www2.fbi.gov/ucr/nibrs/manuals/v1all.pdf.

completed, but these fields will not be necessary for an assault. Similarly, a burglary will typically have a last known secure date/time and a date/time when the crime was discovered, but shoplifting often requires only one date/time entry. A crucial aspect to this type of "branching" design is the ability of the agency to easily and quickly customize data entry forms based upon their business rules. As society evolves and the complexity of crime increases, policymakers must be able to adapt to changing legal, criminal, and societal issues and capture needed information without delay. Examples range from the decriminalization of historically illegal substances like marijuana, to new laws banning synthetically produced designer drugs, to virtual theft involving digital currencies that were not available a relatively short time ago.

Validation rules need to be balanced against the increased workload they may cause to patrol officers when entering a report. This is a continuum whereby more rules may require more work, and vice-versa. The ideal RMS should allow agencies to adjust these rules to their desired level. For example, during implementation an agency may choose to slowly increase validation requirements to ease officer's expectations and increase comfort level with the new system. In the end, the agency will have to find the right balance regarding data accuracy and simplified data entry.

Data should be entered from look-up tables and master records rather than with free text formats, whenever possible. Free text fields lead to spelling errors and variations that need to be accounted for when searching and analyzing data. However, free text fields are sometimes necessary and should not be completely avoided; one example includes the use of free text as a note field on a master record. Agencies should put a considerable amount of forethought into the development of look-up lists. A balance must be struck between having enough choices to be useful, but not so many that the list is unwieldy. In addition, the wording and word order should be carefully considered. For example, instead of having choices worded as "Nikon Camera" and "Canon Camera," it would be better to word them as "Camera – Canon" and "Camera – Nikon" so that they are adjacent in the list.

Comprehensive look-up tables will cover a variety of potential values without similar choices that can confuse the user. The agency should refer to national standard codes (such as those used in the NCIC[14]) and set up a process to update the tables to reflect agency changes as well as national standard changes. Look-up tables should have the flexibility to allow multiple items to point back to the same national standard code.

**Recommendation 8: Data Validation.** Systems should have complex validation rules that increase data accuracy without complicating data entry. Validation controls should prevent physical impossibilities such as future dates of birth, to and from dates in reverse order, and so forth. Data should be entered from look-up tables rather than free text formats when appropriate. Look-up tables should be based upon national standard codes with the flexibility to have multiple items pointing to the same code. Agencies should balance validation rules against increased data entry workload. Systems should have conditionally required fields in an effort to balance validation needs with streamlined data entry.

Good systems make it simple to choose a master record that already exists within the database; they also have the ability to generate automatic notifications if similar matching records are entered. The ideal system would make suggestions while the user is typing, similar to features used by many internet search engines.[15]

No matter how well a RMS is built, errors will still occur; however, the system must have clean data in order to be most effective. A good RMS will have data quality utilities such as a "name merge" function, which allows users to

[14] More information about NCIC can be found at: http://www.fbi.gov/about-us/cjis/ncic.
[15] See, for example, Google Instant Search: http://www.google.com/insidesearch/features/instant/about.html.

reduce duplicate records by combining multiple instances of the same name into a single record after a mistake is discovered. It is highly recommended that agencies employ data quality personnel with the responsibility of ensuring that data is accurately entered and verified. This goes beyond ensuring that records meet the agency's protocol policies. These personnel should establish processes to verify the data upon entry and to clean it of misspellings, duplicates, incorrect entries and other quality issues that impact the overall health of the RMS. Costs of this investment are offset by actual cost savings in many areas. Clean data increases the users' potential for locating desired information, allowing them to spend more time analyzing the data and less time conducting the initial research. Furthermore, clean data reduces liabilities that come from producing products with questionable data. A proven track record of accurate products based on clean data creates a sense of trust and reliability in the records system. Ultimately, clean data increases the consistency of statistical reports that help demonstrate true needs of the department and community.

**Recommendation 9: Data Quality Control.** Systems should make it simple to choose a master record that already exists within the database. Systems should make notifications of duplicate master record entries. Agencies should consider employing a data quality control unit/person with the responsibility of ensuring that data is accurately entered and verified.

**Historical Records**

The RMS should store historical data including changes made to records, such as previous addresses, phone numbers, and physical descriptions. This is especially true of all data associated with names, vehicles and addresses. Agencies often consider this data important for auditing or tracking changes to fields, but analytical products and investigations can benefit from the historical knowledge stored within the system as well. Consider a subject's previous address history; historical residences can prove useful when identifying a person's associates or approximating a subject's activity/awareness space.

**Recommendation 10: Historical Records.** Systems should store historical records of updated fields. This is especially true of data associated with name, vehicle and address records.

**Mapping Data**

An ideal RMS must work hand-in-hand with the organization's mapping system. Effective mapping systems for crime analysis require standard map layers, such as street networks, beat boundaries, and population estimates. When working together data in the records system becomes a part of these map layers allowing analysts to evaluate all possible geographic variables. Mapping layers used for data processing and need to be continuously updated.[16]

The RMS should store geographic coordinates and spatial attributes for each location and address. Coordinates should be derived from a composite geocoding process that uses multiple data sources, including online resources for outside jurisdictions. A typical two-pass geocoding process would look for matches in a parcel file first; if no match was found, it would attempt to interpolate a position from a street centerline reference file. A bypass function should allow users to manually set points on the map for locations without a mappable address. Addresses entered using the bypass function will then need to be reviewed by the data quality personnel and evaluated as possible new master locations.

[16] Boba (2009).

The geocoding process needs to be used for every location associated with an event. These addresses can be associated with a victim, offender, witness, crime scene, arrest, pursuit, vehicle recovery or anything else in the event record. Once the geographic location has been identified, the record should be assigned the spatial attributes for the given area, such as patrol beat/zone, reporting district, tow zone, school district, and residential district. Ideally, systems should allow users to update these spatial attributes as time goes by and boundaries change, while at the same time allowing users to maintain the previous attributes in a historical record.

Systems should allow for incidents to have multiple address points, such as with pursuits or other transitory crime. Transitory location data should be stored as either points or routes, so that officers can clearly indicate where incidents have taken place and which geographic areas have been impacted.

**Recommendation 11: Mapping Data.** Systems should store geographic coordinates for every address entered. Systems should have a composite geocoding process that uses multiple data layers, including online resources for outside jurisdiction mapping. When an event does not have a valid geographic location, systems should have a bypass function with which end users can manually add a point to the map. Systems should allow for incidents to have multiple address points or routes, such as with pursuits or other transitory crime.

**Data Sharing, Importing, and Linking to Third-Party Data**

With the growth of information sharing systems it is essential for RMS developers to build their systems using the National Information Exchange Model (NIEM)[17] standards, with the ability to export directly into the framework of the Global Justice XML Data Model (Global JXDM).[18] Systems built with these standards have the flexibility to import or link to data from third party online reporting tools, and push data to regional/shared justice information databases.

The ability to import or link to third-party data allows analysts to become more efficient and leads to enhanced analytical work, such as the ability to rule out potential suspects based on incarceration dates that conflict with offense date. Ideally, imported records should be automatically matched or related to the appropriate master index records and questionable matches forwarded to data quality personnel.

Agencies are typically required to send standardized crime data to the state and federal government. The ability to electronically submit reports directly to state and federal agencies reduces staff hours required to process these submissions. Although this responsibility does not often fall upon the crime analysis unit, eliminating this time-consuming process allows agency personnel more time to focus on data quality.

**Recommendation 12: Data Sharing, Importing, and Linking to Third-Party Data.** Systems should be built using the National Information Exchange Model and have the ability to export directly into the framework of the Global Justice XML Data Model, allowing the flexibility to integrate with outside systems. External records should be automatically matched or related to existing master index records. Systems should electronically submit reports to regional, state, and federal agencies.

**Attachments**

Systems should allow for files such as PDFs (portable document formats), pictures, spreadsheets, and text documents to be attached to database records. The process of attaching documents should be simple and open

[17] More information on NIEM can be found at: https://www.niem.gov/Pages/default.aspx.
[18] Information and documentation on the Global JXDM can be found at: https://it.ojp.gov/jxdm.

to users writing or updating reports. Users should also have the capability to review all documents attached to a record. Figure 2 depicts a RMS tool used to attach and view files associated with an incident record.

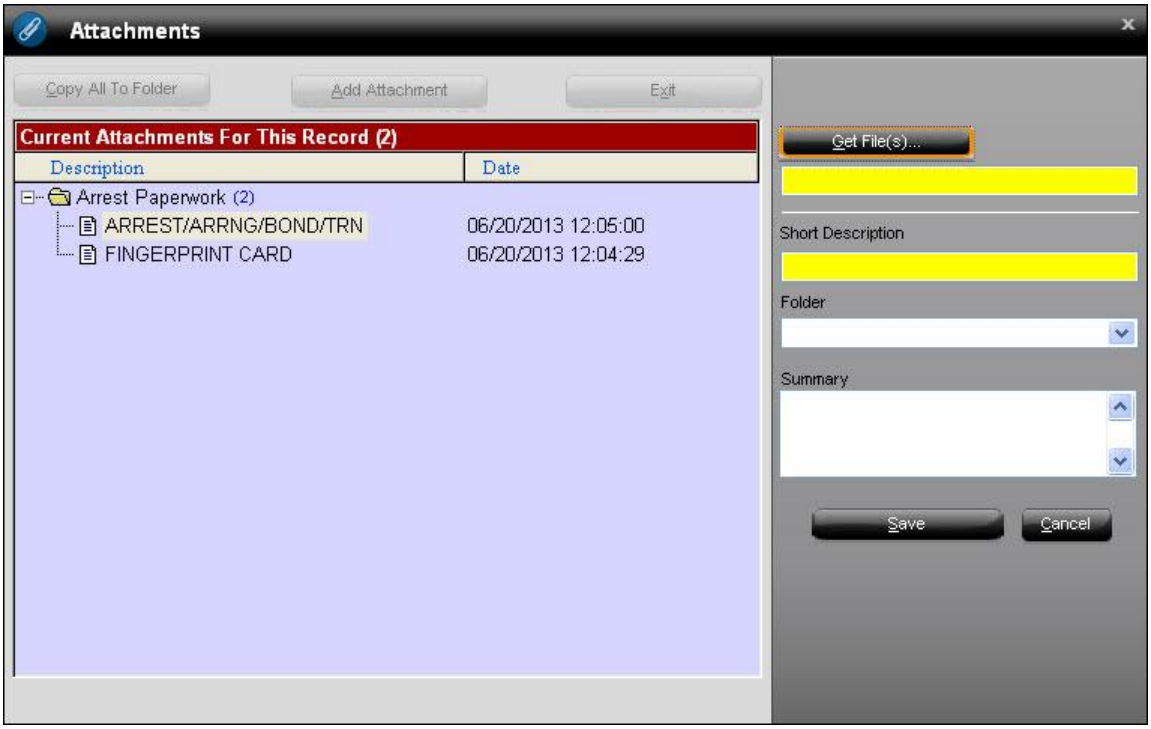


**Figure 2.** Sample attachment screen from a RMS vendor

The import process should include automatic optical character recognition (OCR)[19] to convert the document into a searchable format. The OCR process should allow users the choice to review and validate the results to ensure that the text is captured correctly.

It is important for all data to be accessible for browsing and use by third party tools, and this also holds true for attachments. This is important to note because attached files may be part of an official record, and RMS vendors may restrict access in an effort to protect file integrity for court proceedings. One approach to making files accessible while maintaining integrity is to store them in read-only format; another approach is to store a replicated copy in a secondary directory.

**Recommendation 13: Attachments.** The process of attaching documents should be simple and open to the appropriate users. Attachments should be automatically processed using OCR technology and the resulting text should be searchable. Systems should store attached documents and images in non-proprietary file types and allow folder browsing of these files.

[19] See, generally: http://en.wikipedia.org/wiki/Optical_character_recognition.

# Data Analysis

### Data Relationships

The relationships of people, property, and locations to calls, crime reports, traffic collisions, arrest records and other data are very important for analysis. The RMS should be able to capture various relationship types and make them available for research, analysis and investigations. Modern RMS systems should be built using a relational database management system (RDBMS)[20] or equivalent system allowing for master records to be linked to each other and to file records. Common relationships within crime and traffic data pertain to people (victims, offenders, witnesses, and property owners), property (stolen, recovered, and damaged), and locations (reported, started, ended, and occurred). Users should be able to view both multiple directions of relationships; for example, one can identify all vehicles associated with a specific person as well as all people associated with a specific vehicle. In certain instances users should be able to identify all associates of a person via secondary links, such as crime reports, traffic collisions, vehicles, and locations. When secondary links are used to display associations between records, they should be ranked by level of closeness.[21] An ideal system will allow users to build and export a tabular relationship specification file detailing association strength, direction, and chronology. This file should have the ability to be saved in an easily accessible, non-proprietary file type, such as comma separated values (CSV), so that the user can incorporate it into other analytical tools at their disposal.

When evaluating a system's ability to capture relationships, it is important to pay close attention to linking points and the quality of master index records. It is common for systems to adequately match name or business records to crime and traffic data; however, outlying or secondary links are often overlooked or omitted. A RMS may have the ability to record a person's employer or school information but not to create a relationship between the person and the master business record. In this example, if the relationship was properly established, users would gain the ability to see all other persons associated to the business, giving the analyst valuable details about potential coworkers. In other scenarios, systems may have robust linking ability but do a poor job of maintaining clean index records, voiding the value of sophisticated multilevel linking. For example, an RMS with the functionality to identify if property was stolen and then pawned could be under-utilized because of dirty data diminishing the ability to find a valid match. Attention to detail when developing these linking routines can help to highlight valuable information often missed in the noise of large datasets. As previously stated in the section on data management and quality control, this level of detail will have to be balanced against the increased workload required to maintain accuracy.

**Recommendation 14: Data Relationships.** Systems should be built using relational database management system technology. Systems should be able to capture, show, and export all common relationships including known associates of individuals based on prior contacts, crime reports, and traffic incidents. Systems should also capture outlying links such as emergency contact, employer, and school. Ideal systems will allow the visual display of relationships with link diagrams, and the export of tabular relationship specification files.

### Data Review

Due to the fast-paced nature of police work, crime analysts need to have access to RMS data as soon as it is entered. This is especially important when identifying emerging trends, tracking suspect activity, and providing

[20] See, generally: http://en.wikipedia.org/wiki/Relational_database_management_system.
[21] Sabidussi (1966).

valid case leads. Analysts should not be prevented from seeing data because it is not yet finalized or still being completed. This is especially important to consider when utilizing a RMS that is split between a report writing system and an administrative records management system. Split systems should allow analysts to see queued data immediately, as this information is critical for time-sensitive investigative and intelligence work. However, analysts will need to be cognizant of the fact that information in unapproved records is still tentative and may never become finalized or official.

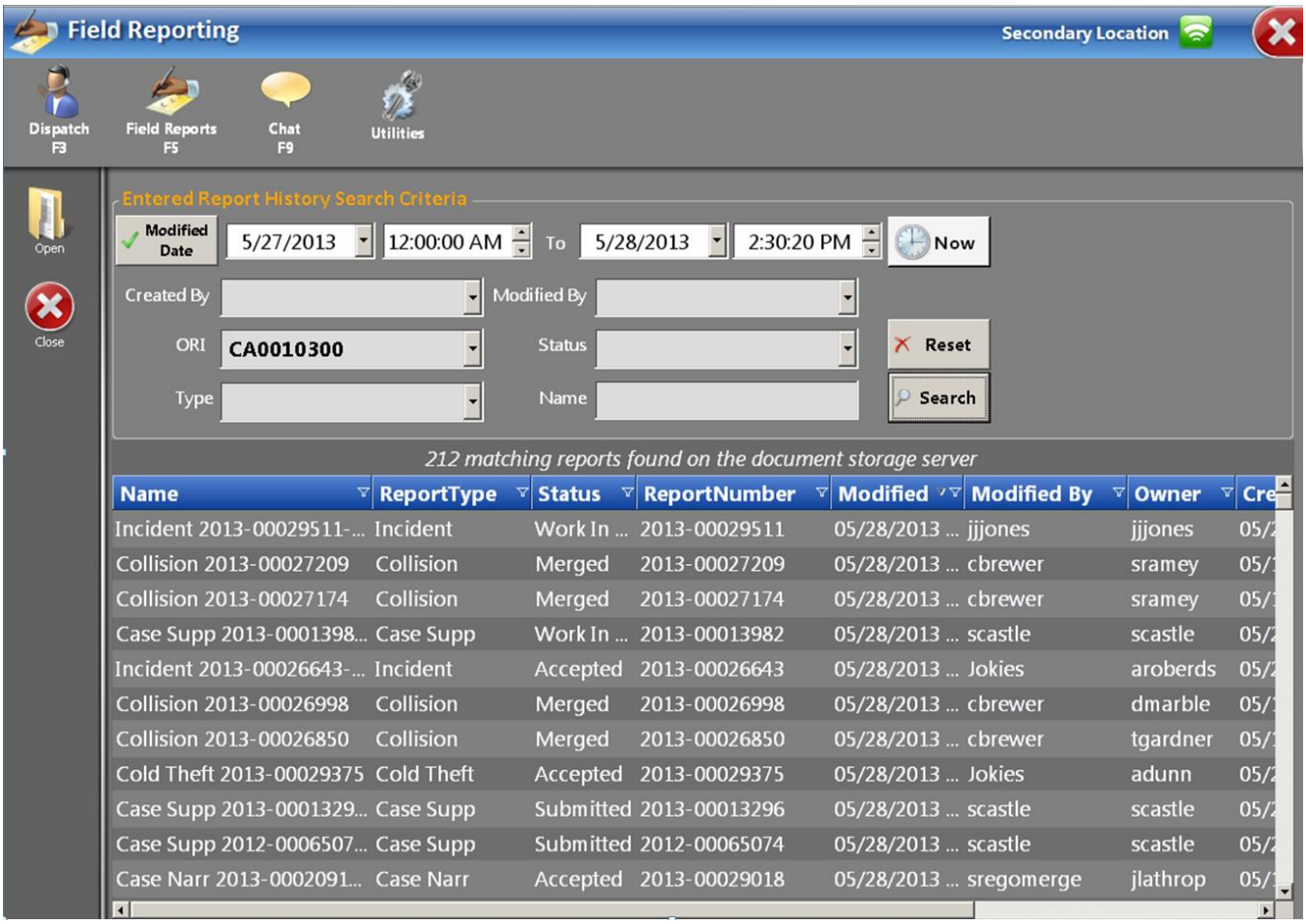


**Figure 3.** Sample of a RMS mobile module report queue that allows access to documents created in the mobile report writing system

**Recommendation 15: Data Review.** To assist in identifying emerging trends, crime analysts should have access to data that is not yet finalized. Systems that are split between mobile data entry and administrative records management should allow for analysts to review data in the queue.

### Search Functions

The search functionality of the RMS is crucial, as users must be able to quickly locate desired information. The RMS should include the ability to search against any field, including narratives and attached files. Standard query types should include wildcards, single character wildcards, word proximity, ranges, and phonetic matching.

In addition to standard field search tools, a good RMS should include a single point search similar to an internet search engine. A single point search tool should be able to search against all fields at once, including narratives

and attached files, and should accept common logical operators (e.g., and, or, not). Single point search results should be categorized by record type, such as name, incident, collision, location, or field contact.

Ideal systems should include a feature to perform pattern-matching searches for telephone numbers, email addresses, and social security numbers, using regular expressions.[22] Additionally, a superior RMS should include the ability to perform federated searches of outside data sources based upon record data stored within the RMS. For example, users should be able to search a person against state driver's license data, supervision status (probation or parole), booking photos, and custody records (jail or prison). The same should be true for searching vehicles against state registration data, and real property against parcel datasets.

**Recommendation 16: Search Functions.** Systems should allow users to search against individual fields, including narratives. Systems should include a single point search tool that searches against all fields at once, including narratives and attached files. Search tools should be able to utilize wildcards, single character wildcards, ranges, phonetic matching, word proximity, and pattern searches for telephone numbers, email addresses, SSNs, etc., using regular expressions. Systems should include the ability to perform federated searches of external databases based upon local record data.

**Reporting Functions**

RMSs are designed for collecting, managing, and reviewing data, but not for complex analytical reporting. For advanced reporting and analysis, users should instead look to third party tools with better capabilities. RMS vendors that provide standardized reports within their systems should also provide easily accessible metadata for these reports. Metadata should include source information and data functions (calculations, filters, and restrictions), explaining what is/is not represented. Systems that allow for building customized reports should also allow users to create the metadata as described above. Ideal systems allow for customized reports that can be created ad-hoc, stored for later retrieval, and queried from multiple data tables.

**Recommendation 17: Reporting Functions.** RMS vendors that provide standardized reports within their systems should provide easily accessible metadata, including data definitions explaining what is/is not represented in these reports. RMS vendors should provide a reporting tool that allows users to query all data fields across multiple data tables. This tool should allow users to quickly store and retrieve recurring reports.

**Vendor Partnerships**

It is recommended that complex statistical calculations and reports be left to third party tools or analytical tools provided in partnership with your RMS provider. Agencies looking for crime analysis products within their RMS should consider vendors that have established partnerships with companies specializing in analytical reporting and mapping functions. The IACA maintains a list of analytical software on its website (www.iaca.net), via a forum in which members can review and post comments. When purchasing a new RMS, it is especially important for the new system to be compatible with existing software. This is typically accomplished through open architecture as discussed above.

**Recommendation 18: Vendor Partnerships.** Agencies looking for crime analysis products within their RMS should consider vendors that have established partnerships with companies specializing in analytical reporting and mapping functions.

[22] Regular expressions are special text strings used to describe search patterns. More information is available at: http://www.regular-expressions.info/.

**Data Warehouse**

Ideally, in lieu of a crime analysis module, vendors should provide replicated data in a data warehouse format with denormalized (i.e., flat file) tables. This will allow the crime analyst to search data without having to identify look-up values, account for voided records, or write complex data joins. The data warehouse is also a place to store data from various sources that can be joined together, such as criminal records, parole and probation records, telephone toll data, documents and images. Data replication and warehousing also prevents system performance issues that occur when analysts run complex queries against large datasets.

**Recommendation 19: Data Warehouse.** Ideally, in lieu of a crime analysis module, vendors should provide replicated data in a data warehouse format with standardized tables that allow the crime analyst to search de-normalized/flat file tables without using look-up values or complex data joins.

## Summary

The optimal RMS should be intuitive for users with a wide range of skill levels, and allow an agency to conduct automated processing, enter data with ease, and record and find any piece of information. Data should be easily imported into the RMS as well as exported from it. The system should be designed to prevent unclean data, but must also have a way to easily make data corrections. The workflow process of an agency should be part of all phases of the RMS implementation; workflow should dictate the RMS, not vice-versa. The RMS needs to be a flexible, living piece of software that grows and evolves with the agency. Off-the-shelf systems trying to apply the same functionality to all agencies will not work. This is a major investment by an agency and the decision will affect users at all levels of the organization for years after the purchase; agencies should not settle for off-the-shelf systems without customizations to suit their local needs.

## Relevant Literature

Berger, Tony. (2006). *Review of standard functional specifications for law enforcement records management systems (RMS) v.1* – draft by the international association of crime analysts, International Association of Crime Analysts.

Chen, Peter Pin-Shan. (1976). *The Entity-Relationship Model: Towards a Unified View of Data*. ACM Transactions on Databases, 1, 9-36.

Groff, Elizabeth, McEwen, Tom. (2008). *Identifying and Measuring the Effects of Information Technologies on Law Enforcement Agencies: A Guide for Law Enforcement, The Making Officer Redeployment Effective Program*. Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.

Harris, Kelly J., and Romesburg, William H. (2002) *Law Enforcement Tech Guide: How to plan, purchase and manage technology (successfully!), A Guide for Executives, Managers and Technologists*. Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.

Holme, Dan. (2011). To BLOB or Not to BLOB. *SharePoint Pro*. Retrieved from http://sharepointpromag.com/sharepoint/blob-or-not-blob.

IACP/COPS Technology Assistance Program (TTAP). (2004). *Records Management Systems*, International Association of Chiefs of Police. 331-379.

Kennedy, K. Heather. (2001). Dictionary of GIS Terminology. Redlands, CA: The ESRI Press.

Law Enforcement Information Technology Standards Council (LEITSC). (2003). *Standard Functional Specifications for Law Enforcement Records Managements Systems Version II*, Bureau of Justice Assistance.

Law Enforcement Information Technology Standards Council (LEITSC) (2006). *A Project Manager's Guide to RMS/CAD System Software Acquisition*, Bureau of Justice Assistance.

Sabidussi, Gert. (1966). The Centrality Index of a Graph. Psychometrika, 31(4), 581-603.

| # | Requirement | Yes/No |
|---|---|---|
| **1** | **Crime Analyst involvement in Procurement and Implementation** | |
| | Connecting to and setting up the mapping infrastructure | |
| | Building and modifying look-up tables | |
| | Determining required fields | |
| **2** | **Open Architecture** | |
| | Standard database conectivity protocols allow access for third party tools | |
| | Does not use unsearchable field types to store images and large unstructured text | |
| **3** | **Data Dictionary** | |
| | Vendor provides an Entity Relationship Diagram (ERD) | |
| | ERD and data dictionary are available to crime analysts | |
| | Field descriptions are in clear language describing the data stored in each field. | |
| | A schema is provided with technical notes for each element | |
| | Schema includes field type, size, primary and child key(s), used and unused fields | |
| | Schema includes calculations, look-up tables, and data dependencies | |
| | Schema includes dates of creation, last modification, termination and source | |
| | Metadata is accessible as an online or network resource in an easily-searchable format | |
| **4** | **Visible Control Source** | |
| | System provides the ability to see the database table and fieldnames associated with controls | |
| | The feature identifies look-up tables necessary to decode stored values | |
| | The feature provides information in a discreet way | |
| **5** | **Importing Legacy Data** | |
| | Legacy data data evaluated in advance of decision to import it into a new system | |
| | Crime analyst involved in evaluation process | |
| **6** | **Importing Computer Aided Dispatch (CAD) Data** | |
| | Agency evaluates CAD and RMS independently | |
| | CAD data is transferred into the RMS system once the CAD event is complete | |
| | All relevant data elements are transferred into the RMS | |
| | Comments are easily understandable, delimited by a carriage return and date/timestamp | |
| | System-generated comments are separated from user-entered comments | |
| | Addresses and associated geographic information is transferred to the RMS | |
| | System identifies master name records that were queried during a CAD event | |
| **7** | **Data Entry** | |
| | Data entry tools are visually pleasing and clearly understandable | |
| | Data entry tools are easy to navigate | |
| | Data entry tools are customizable enough to reflect local procedures and terminology | |
| | Data entry interface should be separate from the interface used to review crime reports | |
| | Mobile data entry tools are designed to prevent inadvertent data submission | |
| | System captures data required for regional and national reporting standards | |
| | Crime charges are automatically ordered based upon national standards | |

| | | |
|---|---|---|
| **8** | **Data Validation** | |
| | Validation rules increase data accuracy without complicating data entry | |
| | Validation controls prevent impossibilities such as future dates of birth, etc. | |
| | System maximizes use of look-up tables rather than free text formats | |
| | Look-up values are based upon national standard codes | |
| | Look-up values allow multiple items to point to the same national code | |
| | Validation rules are balanced against the increased workload they cause for data entry | |
| | Agency is able to adjust these rules to their desired level | |
| | System allows for conditionally required fields | |
| **9** | **Data Quality Control** | |
| | System makes it simple to choose a master record that already exists | |
| | Suggestions are made available while the user is typing | |
| | System makes notification of duplicate master record entries | |
| | Agency has a data quality control unit/person responsible for verifying that data is accurate | |
| **10** | **Historical records** | |
| | System stores historical records of updated fields | |
| | Name | |
| | Vehicle | |
| | Address/Location Records | |
| **11** | **Mapping Data** | |
| | System stores geographic coordinates for every address entered | |
| | System has composite geocoders that use multiple layers including online resources | |
| | System has a bypass function with which users can add a point on the map | |
| | System allows for incidents to have multiple address points, such as with pursuits or transitory crime | |
| **12** | **Data Sharing, Importing, and Linking to Third-Party Data** | |
| | System is built using the National Information Exchange Model | |
| | System is compliant with Global Justice XML Data Model | |
| | Imported records are automatically matched to existing master index records | |
| | System electronically submits reports to regional, state and federal agencies | |
| **13** | **Attachments** | |
| | The process of attaching documents is simple to use and open to appropriate users | |
| | Attached documents are automatically processed using OCR technology | |
| | Recognized text from attachments available for searching | |
| | System allows folder browsing of attached documents and images | |
| | Attachments are stored in non-proprietary file types | |
| **14** | **Data Relationships** | |
| | Systems is built on a relational database management system | |
| | Systems captures all common relationships | |
| | Systems captures outlying links such as emergency contact, employer and school | |
| | Systems allow visual display of relationships with link diagrams | |
| | System exports relationships to a tabular relationship specification file | |

| | | |
|---|---|---|
| **15** | **Data Review** | |
| | Crime Analysts have access to information that is not yet finalized | |
| | Split system allows analysts to see reports in the queue | |
| **16** | **Search Functions** | |
| | System allows users to search against individual fields including narratives | |
| | System includes a single point search tool that searches against all data elements | |
| | Search tools allow for wildcards, single character wildcards, ranges, proximity and phonetic match | |
| | Search tools can perform pattern searches using regular expressions | |
| | System can search all attached documents based upon OCR data | |
| | System can perform federated searches of external databases based upon local record data | |
| **17** | **Crime Analysis Functions** | |
| | Standardized reports include meta data and explanations of what is/is not represented | |
| | System has a reporting tool that allows users to create ad-hoc reports | |
| | Reporting tool allows user to query data from multiple data tables | |
| | Reporting tool allows user to store and retrieve reports | |
| **18** | **Vendor Partnerships** | |
| | RMS vendor partners with a separate company for analytical reporting and mapping functions | |
| **19** | **Data Warehouse** | |
| | System includes replicated data in a data warehouse with denormalized/flat-file tables | |