



Integration of DNA, fingerprint, and firearm databases into forensic intelligence networks for a real-time case assessment model

Jamie S. Spaulding & Keith B. Morris

To cite this article: Jamie S. Spaulding & Keith B. Morris (2019) Integration of DNA, fingerprint, and firearm databases into forensic intelligence networks for a real-time case assessment model, *Journal of Policing, Intelligence and Counter Terrorism*, 14:1, 39-61, DOI: [10.1080/18335330.2018.1548770](https://doi.org/10.1080/18335330.2018.1548770)

To link to this article: <https://doi.org/10.1080/18335330.2018.1548770>



Published online: 15 Feb 2019.



Submit your article to this journal [↗](#)



View Crossmark data [↗](#)



Integration of DNA, fingerprint, and firearm databases into forensic intelligence networks for a real-time case assessment model

Jamie S. Spaulding and Keith B. Morris

West Virginia University, Morgantown, USA

ABSTRACT

Forensic analysis traditionally supports the investigative process from crime scene to trial on a case-by-case basis which fails to exploit the potential of forensic data. Introduction of forensic data to an intelligence framework enables an analytical assessment which seeks to prevent missed linkages between cases. An intelligence network is constructed of entities interconnected with links to show associations and relationships. Utilisation of criminal case information in an intelligence network enables the investigator to access information from other cases to solve the case at hand given other cases. A three-step import specification model was developed which incorporates information relevant to the incident, evidence recovered as a product of the investigation, analytical results, and performed database searches. Application of the provided import specifications aid in harnessing inter and intra-case linkages for a shift in the forensic landscape to a multi-case global focus while streamlining the process with automated information extraction. The product of forensic intelligence yields the ability to conduct crime analysis enabling the identification and disruption of criminal activity while informing operational decision making. The multi-case approach also has the potential to assist in developing a more uniform approach, nomenclature, and case assessment across the forensic and investigative science community.

ARTICLE HISTORY

Received 1 August 2018
Accepted 15 October 2018

KEYWORDS

Forensic; intelligence;
database; Analyst's
Notebook[®]; case assessment

1. Introduction

Forensic science is utilised in criminal investigations to provide insight for the trier of fact through inference development and analytical results from the evaluation of collected evidence. Currently, the process is dependent on the investigator to recognise and assess evidence for inference development about the circumstance. Jackson and Jones (2009) note that forensic scientists and investigators primarily rely on personal experience and personal opinion for assessment and it can often be vague or unclear how the expert arrived at an opinion. Without understanding how the opinion was reached, it is difficult to rationalise and challenge the opinion within a legal forum. A graphical interface or chart

showing the linkages within a case may alleviate this concern. Additionally, forensic intelligence networks can integrate evidence associated with a case and draw linkages with other cases, evidence, and other entities in the chart. Forensic intelligence is defined as an accurate, timely product from logic case data to inform investigation and/or intelligence processes (Ribaux, Walsh, & Margot, 2006). For this project, forensic intelligence is being integrated with and applied to an intelligence-led policing platform for United States (US) law enforcement agencies. According to the US National Institute of Justice, intelligence led, or proactive policing relates to law enforcement use and analysis of data and patterns to comprehend the nature of crime problems (National Institute of Justice, 2018; Sklansky, 2011). Strategies and tactics are derived by management from these data to prevent, intervene, or mitigate future crime and for operational decision making (Sklansky, 2011).

An intelligence network is a chart which highlights the relationships between entities by arranging data to emphasise associations for the benefit of other entities using more than one, inter-related source (i2 Analyst's Notebook User Guide, 2009). Within a forensic intelligence network, data can be represented in several ways. An entity is any object with a distinct and independent existence (e.g. person, weapon, or vehicle) (i2 Analyst's Notebook User Guide, 2009). Characteristics which describe the object are referred to as attributes (e.g. manufacturer, serial number). Cards are another way to attach information to an entity which is not attributes (e.g. suspect interviews to a case). Links are connections which show associations or relationships between entities (i2 Analyst's Notebook User Guide, 2009). Links can have directions to represent facts (e.g. ownership) or associations (e.g. hierarchy) and can also have diverse types as a descriptor of the link characteristics (e.g. suspect, victim, owner) including the properties it contains and its appearance in visualisations (solid, dashed, or dotted lines). All these data can be graded (assessed) to represent the strength, reliability, accuracy, or degree of belief regarding an entity or association in the network chart. The collation of information is similar to traditional police methods where a major investigation might include the creation of a link chart on a bulletin board which shows people, locations, *etc.*, with some connected by lengths of string. The difference is that a digital interfaced intelligence network is entirely searchable and can uncover linkages automatically. The network should also contain all cases to detect link all entities, an innovation that would not be possible with the bulletin board.

Forensic intelligence supports other investigations through the collection, interpretation, and dissemination of information. Entry of information into a network facilitates linkage of entities through identities. Each entity must have a unique identity which establishes the distinct and independent existence of that entity. The development and use of identities must be consistent because if the same identity enters the network later, these two entities are automatically merged. Another way linkages are formed are through specified relationships by the user (e.g. incident → victim). These linkages, in and between cases, help facilitate information for future investigations by identifying series, relationships with the new case at hand, and individuals involved. According to Williams (2008), 'the true utility of forensic science lies in intelligence-led policing as information to direct ongoing criminal investigations and police intervention rather than as props in a criminal trial'.

1.1. Rationale and need within forensic community

Crimes are solved through the assessment of linkages drawn between an incident, evidence, and persons involved. This is referred to as case assessment and interpretation; primarily achieved through personal experience and expert opinion (Cook, Evett, Jackson, Jones, & Lambert, 1998). Interpreting evidence entails the drawing of rational and balanced inferences from observations, test results, and measurements. Currently, interpretation is done in isolation within a single case approach, which lacks consideration of the greater environment and landscape (Bruenisholz et al., 2016). Integration of forensic intelligence systems and databases into a single cohesive framework is an optimal strategy to improve the current case assessment process and manage the schism that exists between science and law enforcement.

In 2007 it was reported that the work of a forensic laboratory concerns evidence more than for the courts and should also provide insight to the investigation process (Kopp, 2007). Laboratory results are often able to elicit valuable information which can assist the investigation or can be used as intelligence for future investigation numerous advantages to developing a digital repository of case information:

- Collated information yields case assessment in a global and historical context
- Investigators can access cases outside of their own experience
- Provides a reference for recidivists and repeat offenders (crime series)
- A searchable network is far more efficient than notebooks for report generation
- Graphical tool with linkages and grades for trial preparation
- Illustration for juries about the context of the case and the interpretation of evidence
- A heads-up view of incidents which have occurred for determination of potential leads to follow up

The potential reduction in time for the discovery of key information within the case is another advantage of forensic intelligence. When at the crime scene, critical decisions are made which influence the downstream investigation. With active communication of observations to an analyst, the analyst could enter the case/details into the network and if any linkages are drawn, relay any *a priori* information to the investigator on scene. This information may assist in recognising potential evidence/features, relationships between pieces of evidence, inform assessments of the scene and support inferences of case linkage. Identification of crime patterns has the potential to close more cases by using information/evidence from multiple scenes to identify a perpetrator. Additionally, without linkage of cases in the same pattern, an individual may be implicated or convicted for a single crime when in fact that individual should have been identified as a person of interest or potential perpetrator for an entire series of incidents.

Currently, forensic science uses several major database systems which are searched for hits in casework. The three most prominent forensic databases in the US are: the Combined DNA Index System (CODIS) (Federal Bureau of Investigation, 2018a) used for searching of DNA profiles nationally; Next Generation Identification (NGI) (Federal Bureau of Investigation, 2018b) for fingerprints, palm prints, and biometric recognition; and the National Integrated Ballistic Information Network (NIBIN) (Bureau of Alcohol, Tobacco, Firearms, and Explosives, 2018) which is a repository of spent bullets and cartridge cases from

crime scenes and confiscated firearms. Each of the databases stores samples and search algorithms help to identify case samples by linking them to candidate samples in the database. Although effective, these databases are separated. A lack of integration between forensic identification databases and the other components of the justice system responsible for following up on results is perhaps the biggest weakness (Bieber, 2006). Ribaux, Roux, and Crispino (2016) add that forensic investigators tend to reason in terms of solving potential serial problems, while studies remain focused on the chance to extract a profile from traces, on a case-by-case approach, and on searching for hits in databases, whatever their relevancy. A streamlined approach to exploit forensic evidence is necessary because it regularly contributes to solve investigations. Ribaux et al. (2016) also note that the evidence is used much more throughout the course of the investigation than it is during court proceedings. The forensic investigation must be considered both as an information process and as a rationalisation for future problem-solving. Forensic intelligence networks also offer full searchability for case details (*modus operandi*, location, time, etc.) as do current databases for evidence characteristics. In addition, the intelligence network can integrate and accommodate the search results of those databases for assessment and interpretation.

2. Materials and methods

A database of 29 criminal cases was created using adjudicated cases from a local police agency. The case types included: embezzlement, forgery, fraud, and theft. All case details (date, time, location, persons involved, evidence collected, etc.) were extracted from the case files into Microsoft® Excel® spreadsheets.

In addition, based upon the cases received, the format was used to simulate 51 additional cases to increase crime types and forensic evidence including burglary (fingerprints); shooting incident (firearms); sexual assault (DNA); and other (e.g. auto theft, narcotics, robbery, and unexplained death). The cases were simulated to contain articles of evidence which were then entered into three databases. A simulated DNA database resembling CODIS, a fingerprint database in AFIX Tracker®, and the Integrated Ballistics Identification System® (IBIS®) were used to obtain search results for the simulated cases. The IBIS® is the US Bureau of Alcohol, Tobacco, and Firearms' platform for NIBIN and searching firearm related evidence (Bureau of Alcohol, Tobacco, Firearms, and Explosives, 2018).

2.1. Simulation of criminal cases

The simulation of criminal case details was primarily accomplished using R (The R Project for Statistical Computing, 2018) via RStudio® (RStudio Team, 2018). A random name generator was used to simulate personal information about victims, suspects, and investigating officers. The `randomNames` (Betebenner, 2017) package was used to generate random names (first and last), date of birth, and sex. For each of the persons, a social security number (US Identification Number) was randomly generated using the sample function to select nine-digit integers without replacement. A repository of all simulated persons was created (Database of Persons). The random number generator (sample) was also used to create street numbers for incident locations. These numbers were assigned to a

list of Chicago street names (collected from Chicago Data Portal (City of Chicago, 2012)) for the address of each simulated incident. A separate repository of all generated locations was created (Database of Locations). Case information was also generated. A random integer (1–365) was generated to provide the month and day of the incident, all incidents were randomly assigned one of the last three years (2015–2018). Using this date in conjunction with random integers a case number was created: ‘year-month-random5’ (e.g. 2018-01-12345). Incident time was also generated for each case. Lists of potential evidence items and common locations where that item would be expected were manually created and assigned to each case number. An example of evidence assigned to a case is given in Table 1.

Using the reports obtained from the police agency as a template, cases were compiled. Persons were selected from the repository (Database of Persons) and assigned roles as victims, suspects, and investigating officers for each case. Cases were assigned one victim and the number of suspects was random (1–3). A location was also assigned to each case for the crime scene. The case information was extracted into two separate Microsoft® Excel® spreadsheets. The first contained the case details and information regarding the incident: location; date and time; type (e.g. burglary); organisation/business involved; and victim/suspect(s) or person(s) involved; and identifying information (name, social security no., date of birth, sex, ethnicity, etc.). Another spreadsheet contained all of the evidence collected or involved in the incident and identifying information (manufacturer, model, serial c.) about the evidence. This information is typically what would be contained within an investigation report.

2.2. Database preparation

Three forensic databases were used to integrate information into the intelligence network. DNA profiles were simulated and compiled into a database to represent the CODIS database. AFIX Tracker® was the automated fingerprint identification system (AFIS) used. The IBIS® was used for firearm identification data.

2.2.1. Creation of a DNA database

The NIST database (Hill, Duewer, Kline, Coble, & Butler, 2013) of allele frequencies for the Caucasian subset ($N = 361$) was used to simulate DNA profiles. Of these loci, only the initial 13 CODIS loci and Amelogenin were used. After possible alleles were determined for each locus, these alleles were sorted by increasing frequency. Two random numbers between zero and one were generated (using **runif**) and allele calls were assigned to the bin in which the random numbers fell. An example is given in Figure 1.

Table 1. List of evidence associated with a simulated case.

Tag #	Type	Item	Location found	Manufacturer	Model	Serial number
A	Fingerprints	Fingerprint Lift	Residence – Front Door			
B	DNA	Swab	Sexual Assault Kit			
C	Firearms	Hi-Point C9 Handgun	Residence – Kitchen: NW Corner	Hi-Point	C9	P1234567
D	Trace	Blue Fibre	Residence – Kitchen: On Victim			
E	Firearms	Cartridge Case	Residence – Kitchen: NW Corner			

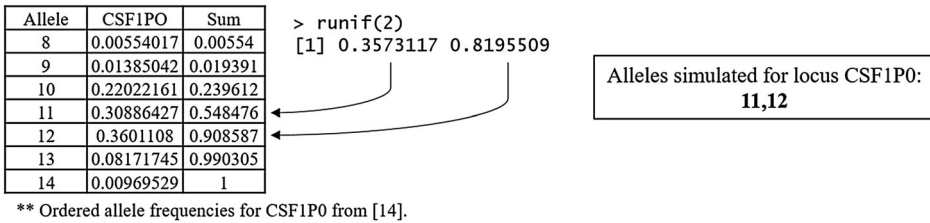


Figure 1. Simulation of alleles at a locus using NIST Caucasian frequencies.

Simulation of alleles was completed for each locus. The created DNA profile string is a concatenation of all simulated alleles. Each allele is two characters in length (e.g. ‘08’ and ‘17’) except where there is a decimal value (e.g. ‘31.2’). In this case, the character after the period belongs to the two characters preceding the period. If a profile is homozygous at a locus, the allele was repeated for consistency (e.g. 1212). Additionally, the loci were generated and used in a fixed order (shown below). As an example, the following profile string was created:

XY121217181213091012131112121413152831.22126091108091717

AMEL/CSF1PO/D3S1358/D5S818/D7S820/D8S1179/D13S317/D16S539/D18S51/D21S11/FGA/TH01/TPOX/VWA

No duplicates were found for any of the created profiles. Additionally, the DNA profile simulation was evaluated to ensure that the results were realistic. The script was run to simulate 100,000 DNA profiles. The allele frequencies of the simulated results were then compared to the frequencies of the Caucasian database. Performance can be seen in Table 2.

Profiles were simulated and assigned to each person in the repository. The R script for the simulation of DNA profiles can be made available from the authors upon request.

2.2.2. Fingerprint database

A sample database within AFIX Tracker® was used for fingerprint comparisons. The database consisted of a tenprint database and a latent database. The simulated information from the database of persons was populated into the personal detail fields. This associated each simulated person with a set of fingerprints. Based upon a master list of latent matches for the tenprint database, latent print images were assigned as evidence to

Table 2. Evaluation of the CSF1PO locus simulation. ‘CSF1PO’ is the allele accepted NIST frequency and ‘Database’ is the simulation frequency. Difference and % difference were used to compare the values (accuracy of the simulation).

Allele	CSF1PO	Sum	Database	Difference	% Difference
8	0.00554	0.00554	0.00565	-0.00010	-1.89
9	0.01385	0.01939	0.01394	-8.45845E-05	-0.61
10	0.22022	0.23961	0.22092	-0.00069	-0.31
11	0.30886	0.54848	0.30797	0.00090	0.29
12	0.36011	0.90859	0.36014	-2.91967E-05	-0.01
13	0.08172	0.99030	0.08144	0.00028	0.35
14	0.00970	1.0000	0.00999	-0.00030	-3.09

cases. Each latent fingerprint assigned to a case was searched and the matching fingerprint was identified mirroring the actions of an analyst using AFIX Tracker® or another AFIS.

2.2.3. Firearms database

A research database of .380 ACP caliber cartridge cases from an IBIS® was used for firearms data. Cartridge case search matches were extracted from the database and compiled into Microsoft® Excel® spreadsheets. Cartridge cases and the information that had been entered into the IBIS® were assigned to simulated cases as evidence and firearm registrations were simulated to persons involved in the cases.

2.2.4. Summary of simulated cases and databases

Table 3 outlines the breakdown of simulated cases by type of evidence/database involved. There are cases which had one type and cases which use combinations of evidence to evaluate the performance of importing case information. Combinations of evidence from databases also allowed for the detection of interference or errors in multiple imports. Overlap of evidence types was also used to show possibilities that could be encountered with casework involving major cases.

2.3. IBM® i2 Analyst's Notebook®

IBM® i2® Analyst's Notebook® was the software used for data analysis and investigation. The software is a visual intelligence analysis environment that enables the user to quickly gather, organise, analyse, and visualise data from several sources (criminal cases within the database) (i2 Analyst's Notebook User Guide, 2009). A single cohesive and integrated network with case associations was developed from the case data. Analyst's Notebook® has a variety of entities that are built into the programme regarding crime, cyber-crime, locations, and weapons. Examples for evidence types and forensic analysis are: DNA, fingerprints, footprints, controlled substances, firearms, and cartridge cases.

2.4. Import specifications

To develop an intelligence network that can accommodate and reproducibly depict a case, import specifications were created to interpret the spreadsheets of case information and streamline the process. An import specification is a structure defined by the user which defines how Analyst's Notebook® interprets a text or spreadsheet file from an external source for conversion into chart items (including entities, attributes, and instances) during an import procedure (i2 Analyst's Notebook User Guide, 2009). The model in this

Table 3. Number of cases using each database. The databases were searched, and results were integrated into the intelligence network.

Database	Total cases
AFIX®	10
DNA	11
IBIS®	10
AFIX® & DNA	5
AFIX® & IBIS®	5
DNA & IBIS®	4
AFIX®, DNA, & IBIS®	6

project consists of three steps of import specifications to integrate case information into a forensic intelligence network.

2.5. Generation of a forensic intelligence network

The first import specification begins the import of case information into Analyst’s Notebook®. Firstly, two import specifications were created to incorporate the criminal case information (i.e. case description, potential suspects, victims). One specification imports cases where the victim was a person, while the other imports information where the victim was an organisation or business. The distinction between these import specifications is the icon depicting the entity of the victim and the details (attributes) associated with the victim (person vs. organisation). Sorting, searching, and network analytics (e.g. hotspot analysis) are made easier with different icons and the icon difference helps distinguish the crime incidents. The import specification design (individual import) is shown in Figure 2.

In Figure 2, the victims are represented as a silhouette consisting of a black outline. Suspects or persons of interest are presented as a generic icon for an individual. The difference in icon is solely to aid the user in differentiating persons involved in the case (male/female icons could also be used here). Below all icons are attributes associated with that entity. The following attributes are shown (can be customised): date of birth, race/ethnicity, social security number, and gender. Also shown are the links between the incident icon and entities. In the import specification, these are shown as ‘(same as Type)’ meaning that the link type corresponds to the designation as a potential suspect, a victim, or someone involved as noted by a column in the spreadsheet; that title is then displayed on the graphical network. The spreadsheet is raw data and the import specification interprets, charts, and identifies linkages within the data. As the information is imported, the software populates fields (entity names, attributes, cards) from assigned columns in the spreadsheet.

2.5.1. Network design for consistency

The specification shown in Figure 2 will import, populate, and link three potential suspects and two victims, but can easily be expanded to accommodate more individuals.

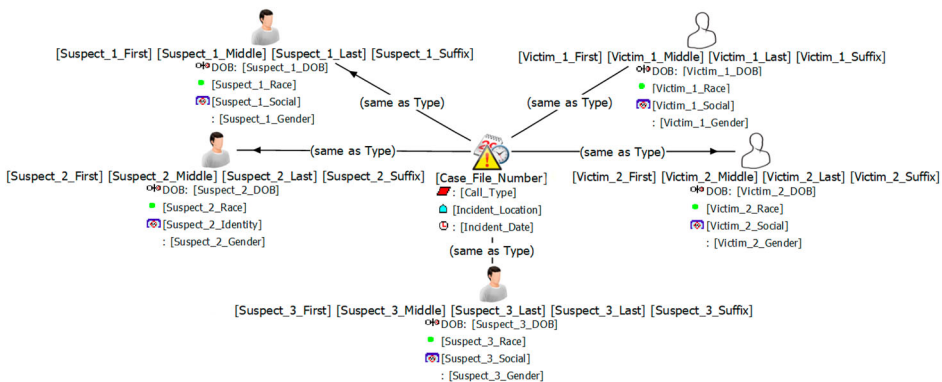


Figure 2. Image captured of the import specification for a case with two persons as victims. The linkage of each individual and article of evidence to the incident is shown.

Each individual has an underlying unique identity in the following format: 'First_Middle_Last_Suffix_SSN' (e.g. John_Adam_Smith_123456789). Organisations were assigned the identity of 'Name_FullAddress'. This format for identity and name must be consistent and is case sensitive. Analyst's Notebook® uses the identities to create linkages and duplicate identities are merged into a single entity/icon. This is useful for repeat offenders because the entry (or import) of that individual into the network a second time will merge with the previous entry, thus linking the cases/incidents. The name visible in the network (with the icon) is 'First_Middle_Last_Suffix' which is a concatenation of columns in the spreadsheet containing the first, middle, last names, and a suffix (if applicable) instead of displaying the identities on screen. Each incident also has a consistent identity: 'Department-CaseNumber'. Consistency is necessary for successful and accurate linkages. The design of the import specification model is aimed to facilitate consistency with Microsoft® Excel® spreadsheets which populate fields and a standardised entry for the system. For this project: all capital letters were used (Analyst's Notebook® is case sensitive); no spaces (underscores instead), and no dashes or spaces for social security numbers (e.g. 123456789). Attention must also be paid to correct spelling. A standard method for data entry into the system is also ideal as it ensures reproducibility. Additionally, a standardised method prevents the development of false links. Accurate time and date information allow for the software to yield a chronology if necessary or create heat maps at time intervals if desired.

Use of these specifications is shown here as the first step yet can also be used when individuals of interest are identified downstream in an investigation. The case file number provides the basis for all the information to be tied together for the continual development of the network.

2.6. Inclusion of evidence articles

The second step of import specifications adds evidence collected on scene or associated with the case into the network. The specification is built to automatically link evidence to the given case. If available, attributes are also assigned to the entity (evidence) and included: evidence tag number, manufacturer, model number, and serial number. This is also customisable depending on the protocols of the agency. The link between the entity and the case was labelled with the role that the entity played in the case (e.g. collected from scene, stolen, recovered from suspect). The evidence can be imported as it is collected (imported sequentially one at a time) or after leaving the crime scene (mass import) and the import specification integrates the items into the network. The import specification to import articles of evidence is shown in Figure 3.

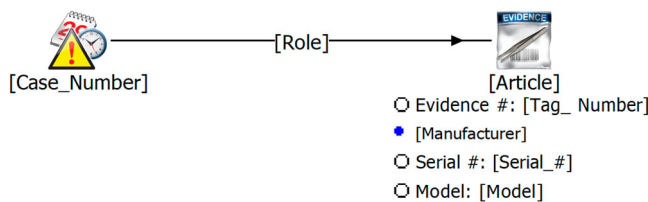


Figure 3. Import specification design for inclusion of evidence articles for an incident.

This import specification also has underlying identities. The case identity is the same as the previous import step which links the evidence to the case. The evidence is given the following identity from columns in the spreadsheet: 'Case#_Article_Tag#' (i.e. 2018-11-1111_SerratedKnife_B). After the first two import steps, the case(s) is/are imported into the network.

2.7. Integration of forensic analytical result or database searches

The third import step integrates any analytical test results or forensic database searches that are performed on collected evidence. Using the DNA, fingerprint, and firearm databases; matches were made, and these results were extracted for entry into the intelligence network. Import specifications were designed for DNA matches, fingerprint identifications, firearms evidence, and cartridge case evidence.

The import of DNA matches did not require any extraction of database information because the profiles were simulated and output as spreadsheets. Names and case numbers were assigned to the profiles and this served as the database search result. The import specification for DNA profiles is shown in Figure 4. This import specification can link up to ten cases based upon one case. Observable in Figure 4 is the incident (Case_File_Number) and item of evidence (Subject_Object) from the previous steps. Added is an analytical link to a DNA profile recovered from the item of evidence, also linked to the person (also added if not in network). Any other cases which matched the profile are also charted as leads to follow up on. If that case is in the network, then the two cases become linked through the person.

Fingerprint information from the AFIX Tracker® database required extraction prior to import. AFIX Tracker® populates a Microsoft® Access® database containing all information and records upon searches/use. R was used to connect to the database, extract, transform, and export the data. The RODBC package (Ripley & Lapsley, 2017) enables Open Database

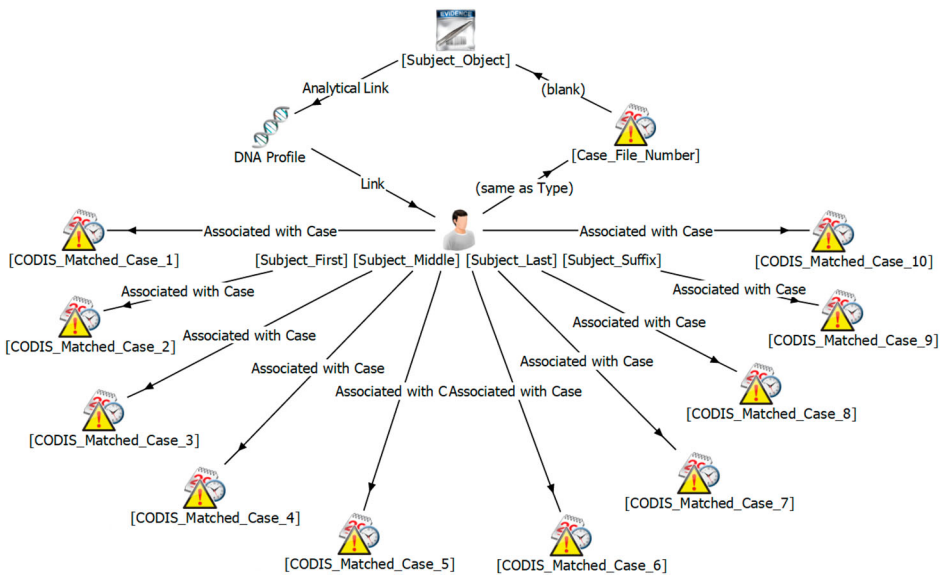


Figure 4. Import specification design for inclusion of DNA evidence search results for an incident.

Connectivity (ODBC) which aims to provide a common API for access to SQL-based database management systems (Ripley, 2017). First, a channel was established to the Microsoft® Access® database using **odbcConnectAccess**. Next, a SQL query was made to the database for each table (**sqlQuery** function), which retrieved the results. These tables were then transformed and compiled into a single data frame containing: biographical information, submitted latents, latent cases, search jobs, and match results. Finally, the script of commands used to extract the information from the AFIX Tracker® database was made executable for automation of extracting the information (Meissner, 2015). The executable file (.Rexec) was programmed into the Windows® framework to run every five minutes (can be changed depending on computation capability or need) (Meissner, 2015). The R script for the extraction of fingerprint information from AFIX Tracker® can be made available from the authors upon request. Minor alteration will enable this script to connect with Oracle®, SQL®, and other types of Microsoft® databases and allow for the transformation and integration into a unified input for Analyst’s Notebook®. ActiveX® (IBM, 2018) is an IBM® add-on which provides interoperability with other software and flexibility for building external applications to communicate with Analyst’s Notebook®.

The new extracted and compiled data frame contains all information needed to import the search/match result into Analyst’s Notebook®. The import specification for the extracted fingerprint results is shown in Figure 5. Observable is the incident (LC.Case_ - Number) and item of evidence (L.Location) from the previous steps. Added is an analytical link to a fingerprint entity recovered from the item of evidence, also linked to the person (also added if not in network). Since this fingerprint is the link between the person it would have this same link in any other cases; an indirect link would be made from this case to another through the fingerprint entity.

The import of firearm database information did not require any extraction because the data was output from the IBIS® as spreadsheets. Names and case numbers were assigned to the firearms in the database and this served as the database search result. Two import specifications were developed for firearms; the first integrates recovered firearms into the intelligence network whereas the second integrates collected cartridge cases and the matching firearm as entities. The import specifications developed for firearm and cartridge case comparison results are shown in Figure 6. Shown in both are the incident (CaseID_ - Sample) and the items of evidence (handgun and cartridge case) from the previous steps. In the recovered firearm specification, a link between the recovered evidence and the test fired sample in the IBIS® database is drawn. Any other cases that firearm has been linked to

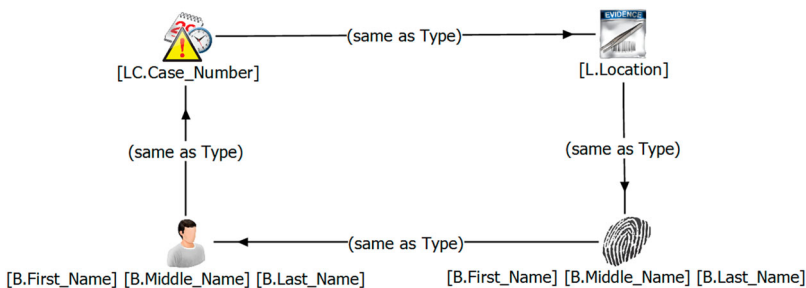


Figure 5. Import specification design for inclusion of fingerprint evidence search results for an incident.

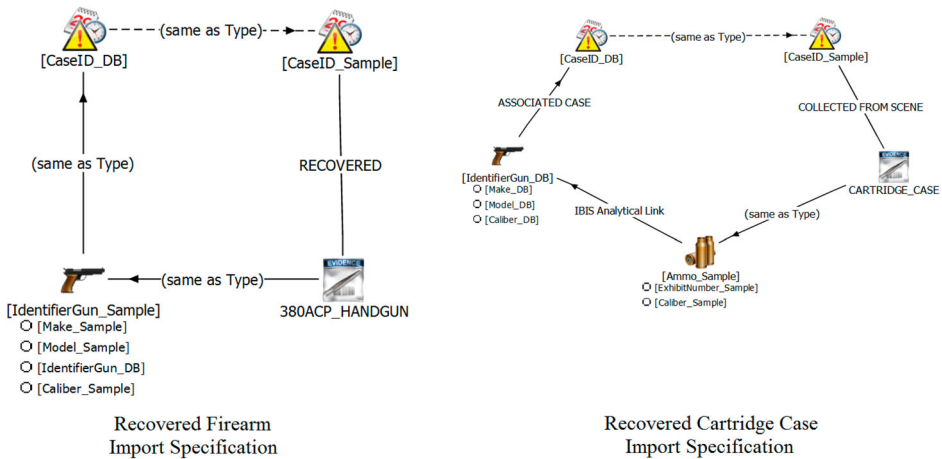


Figure 6. Shown are the import specification designs for inclusion of firearm evidence results (recovered firearm and recovered cartridge case).

are also charted (CaseID_DB), and an unconfirmed (dashed) link is drawn to the case at hand for the investigator to follow up. The recovered cartridge case import specification has a slight difference. A cartridge case entity is built in between the evidence submission and the firearm in the database (hit from system), this firearm is then associated with any cases where it was used. Again, an unconfirmed link is made from any cases the firearm which discharged the cartridge and this case.

2.8. Evaluation of network and system

Evaluation of the import specification model and intelligence network was done through an assessment of a few key features. The simulation of criminal incidents and case files allowed for knowledge of ground truth enabling determination of whether or not appropriate links exist and associations were made. Fifty cases formed a training set used to design import specifications. The remainder of the cases formed a test set which was imported and manually evaluated for accuracy of charting and linkage development. This assessment was completed to ensure all appropriate links were developed for each case. Use of the import specifications for both mass import (all cases) at the same time as well as an individual import into the network yielded the same resulting network. No issues were detected in either import sequence. Articles of evidence were found to be appropriately linked to results and no duplicate links were detected. Entities with the same identities were merged leading to links being drawn between cases yielding a functioning network.

3. Results

An example of a case in a forensic intelligence network using the case information and evidence import specifications (first two steps) is shown in Figure 7. Both cases were received from the police agency and have been anonymised. All evidence, the victims

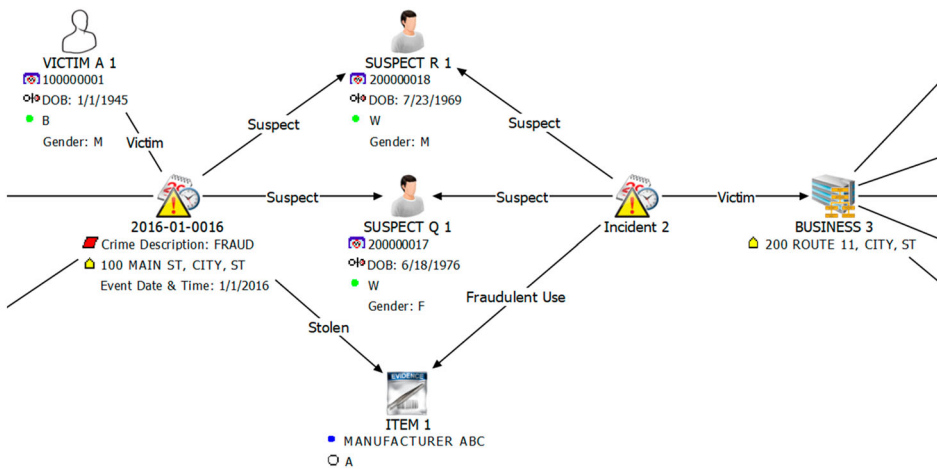


Figure 7. Example of a linkage that was developed in the network between two suspects and two cases.

(person and organisation), and the suspects were charted successfully with linkages to each case. Two separate cases were found to be linked through both individuals and an item of evidence. In one of the cases, the two perpetrators were accused of larceny; the individuals stole a wallet from the victim and were identified by the accused in the report. A later incident of fraud was reported where the two individuals used a credit card alleged to be in the wallet stolen during the prior incident. Figure 7 shows the linkage drawn by Analyst’s Notebook® between the two cases. Note that different import specifications were used for these two incidents (different victim types and separate times) and the persons were merged based upon the underlying identities.

An example of information for intelligence-led policing is shown in Figure 8. The example represents a hot spot of several incidents which occurred at one organisation. The network shows all reported incidents and clustering about the business is evident. Another benefit is the crime type attribute for each incident: fraud and larceny appear common here. The other cases of embezzlement and forgery may be an alert to an investigator if a pattern were to develop.

Two simulated cases using search results from both the DNA and AFIX Tracker® database are shown in Figure 9. In the later incident (2011-07-2758) two pieces of evidence were recovered (latent fingerprint and DNA profile) from the stolen vehicle. A witness implicated the suspect (Dickens). After known DNA samples and fingerprints were taken from Dickens, she was linked through both DNA and fingerprint evidence to incident 2011-07-2758. This search also yielded a hit to a latent fingerprint from a prior auto theft (2010-08-2493) which was also charted and linked to suspect Dickens. The tenprint record for Dickens was not in the database at the search time of the first incident and upon entry, for the latter case, a link was created between both cases in the intelligence network.

DNA, fingerprint, and firearm evidence are combined and shown in Figure 10. The fingerprint evidence consisted of a series of latent fingerprints lifted from the rear master bedroom window (Evidence C1-C3) and two latent fingerprints were also

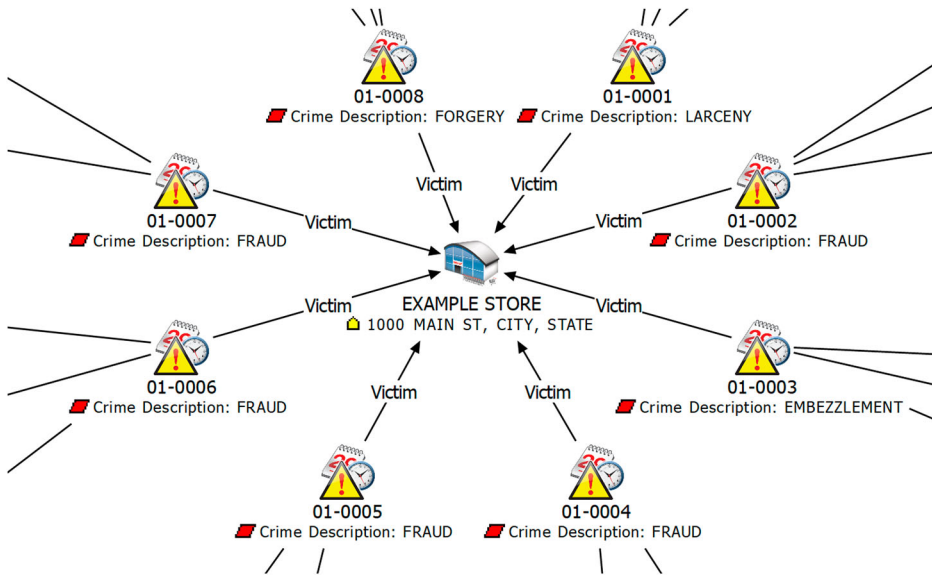


Figure 8. Example hot spot shown at an organisation with several reported incidents. Links leading outward from incidents are evidence/persons associated with each incident.

processed and developed from a coin holder (Evidence D) collected from the scene. All five latent fingerprints were identified as being left by suspect Gomez. Additionally, a touch DNA swab was collected from the scene and a DNA profile was developed, also matching Gomez. The final piece of evidence collected from the scene was a cartridge case which was identified as being fired from a Ruger® handgun as shown. Note that the database search revealed the firearm was associated with a different incident (also charted). The

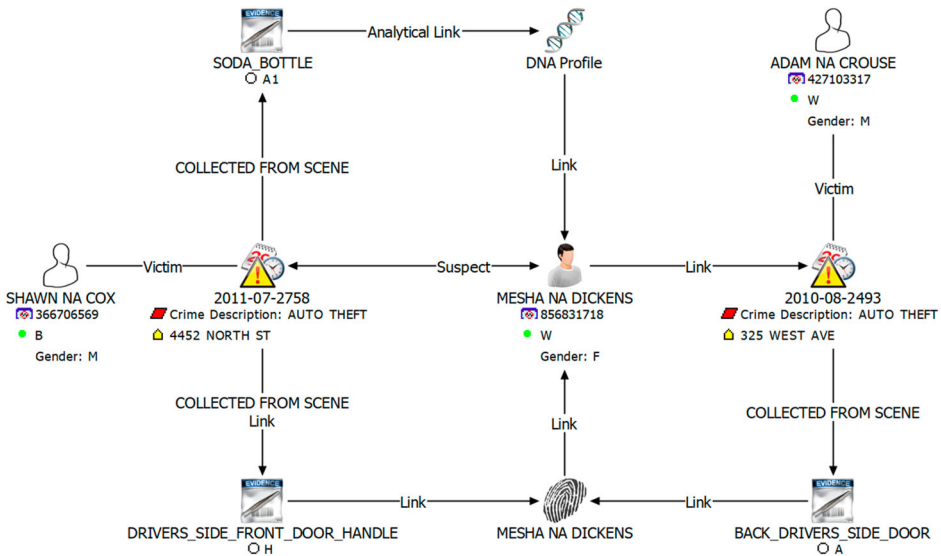


Figure 9. Identification of a linkage between two cases based upon a database hit in the later incident.

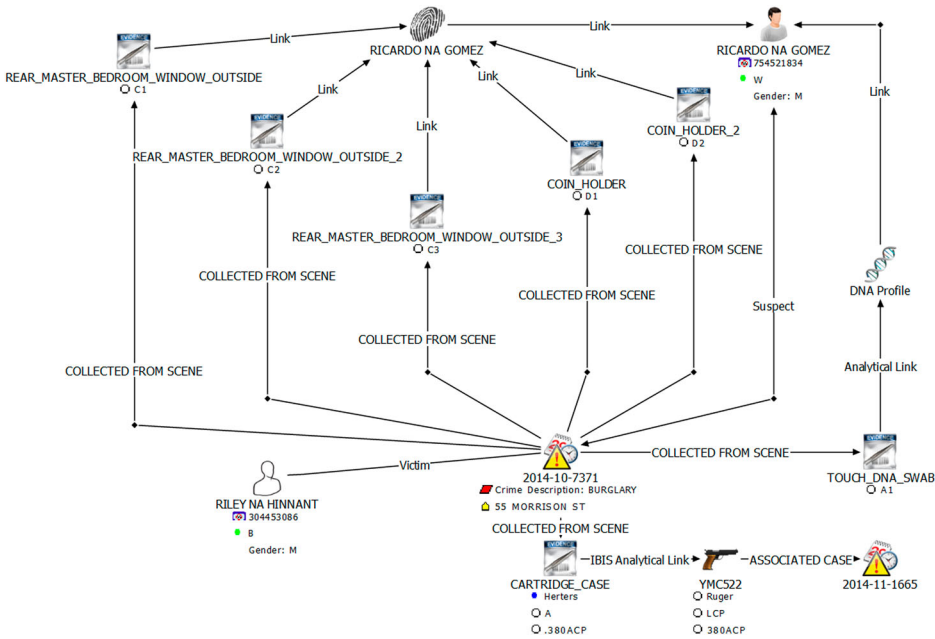


Figure 10. A case example where evidence was integrated from all three databases.

proposed intelligence model in Analyst’s Notebook® enables case assessment with all evidence presented and linkages developed. Figure 10 shows that fingerprints from different sources and DNA collected on a swab match the suspect. Furthermore, the investigators can look into the firearm linkage with an associated case potentially linking the suspect to both incidents with more evidence.

In the analysis of a sexual assault kit multiple contributors are expected. The import specification is designed to chart all contributors and their developed DNA profile. The case is shown in Figure 11 shows both male and female DNA profiles developed from the kit; one matching the victim, another matching a person of interest labelled as suspect.

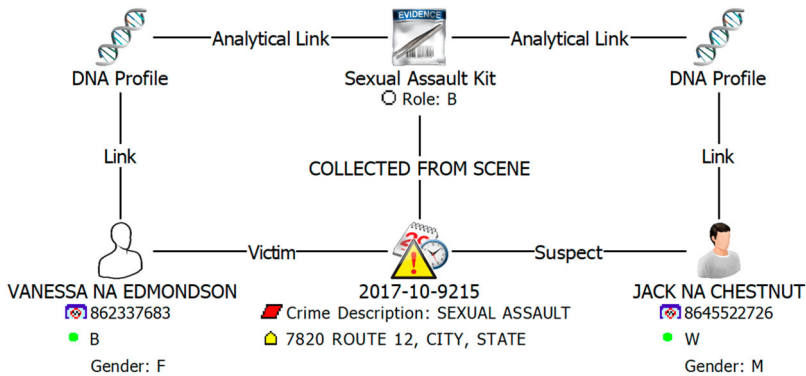


Figure 11. Case example where two profiles were developed through examination of a SAFE kit. One profile was that of the victim and the other belong to a male subject in the database who then became a suspect for the case.

4. Discussion

An effective and efficient investigative service can be achieved through a contemporary model of case assessment and interpretation using a forensic intelligence framework. Developing a comprehensive understanding and outlook on criminal activity is critical to developing and applying effective police capability. Advantages of such action are complete interpretations with the use of forensic analytical results, resource distribution, and informed policing/security actions.

The three-step import specification model developed in this project combines both closed and new cases into an intelligence network intended to aid future investigations through linkage development. The model was created to add an analytical assessment to investigations by exploiting all case information. The first import step utilises two import specifications to enter information about the case into the network. This step is separate because this is the sort of information that can be preliminarily gathered and imported while attending the crime scene. Any resultant linkages provide information to the investigator informing their decision making and potentially identifying evidence of interest. Secondly, the evidence is imported into the network and associated with the cases. The specification for evidence import is separate so it can be rerun numerous times as evidence is discovered or located. The separation of the two import steps provides a less computationally intensive process than the repeated merging of several entities. Finally, step three compiles and draws linkages from analytical results. This step is downstream in the investigation after the evidence has been submitted to a laboratory and processed.

The potential of the forensic intelligence network can be seen in the ability to uncover relationships between items of evidence. Intelligence analysis serves to remedy investigative problems by providing investigators with information about incidents by combining data and applying analytical techniques. Relationships between entities and forensic intelligence analysis for each are given in Figure 12.

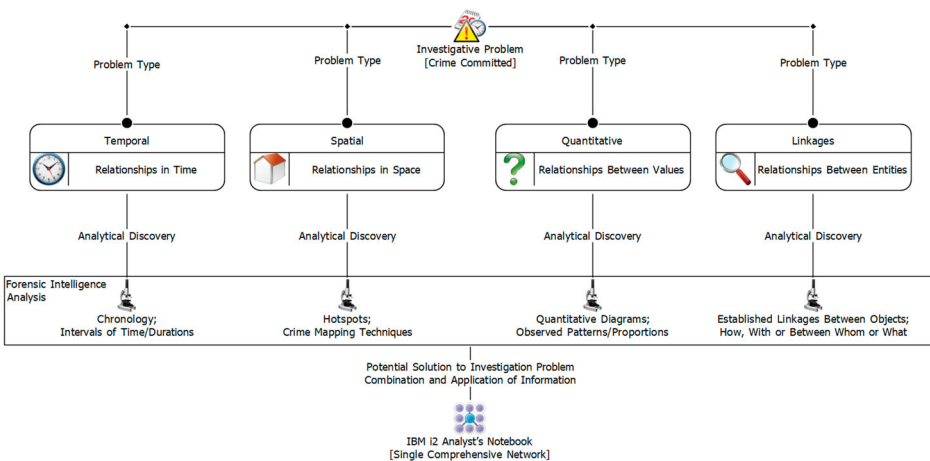


Figure 12. Chart depicting different relationships existent between evidence and the implications of forensic intelligence analysis. Figure adapted from (Rossy & Ribaux, 2014).

Several relationships among persons and evidence are exhibited in Section 3. A fraudulent scheme with the potential to identify the *modus operandi* of the perpetrator(s) is given in Figure 7. For example, a couple (depicted) could steal credit cards from certain target location and then fraudulently use the card at other locations (spatial/temporal relationship). The couple could also make purchase of similar value (quantitative relationship). Additionally, hot spot or cluster analysis with the intelligence framework may yield new *a priori* information for an agency (i.e. Figure 8). The information may suggest that criminal tendencies are changing or that the individual involved may not be a regular customer. Analyst's Notebook® also features the ability to link with satellite imagery for cluster plotting of data as an alternate spatial visualisation of activity.

4.1. Enabling near real-time case assessment

Several avenues of case assessment and interpretation are available through the application of this model. Decisions must be made, both strategically and within individual cases, on the course of action that should be taken. In forensic science, these decisions are made routinely when they are considering which items should be examined. How these professionals make such decisions and whether they have any formal guidance to do so are key considerations (Jackson & Jones, 2009). Suggested is a need for more transparent, repeatable, and objective methods to assess criminal cases to develop safe, robust opinions. The import specifications within this intelligence network model are a fixed mechanism; information is merely added into the spreadsheets. Additional changes within the network can be tracked to develop a history of actions for reproducibility purposes. Representing this reasoning process is of the utmost importance because the inferential process is a complicated one (Robertson, 1990).

A priori information on the crime scene can greatly aid the investigators in their decision-making processes. Database import remotely would highlight new entities upon the execution of the import for review. Direct linkage facilitated by Analyst's Notebook® yields immediate association and recognition by the user. That would enable the investigator to consider this case within the context of other cases. Additionally, entry of witness statements as cards on scene can develop linkages. For example, if a witness claims to have seen John Smith near the scene at the time of incident, a search can be made in that case for John Smith which encompasses all notes and entities. Links or leads can be entered as avenues to follow up.

Near real-time case, assessment is also fostered through the automated extraction of database results. The execution of scripts by the Windows® framework on a loop makes the extraction of results seamless, this execution process was evaluated using a programmed five-minute loop in the Windows framework®. The developed script for AFIX Tracker® result import solely extracts and prepares the new results from the database to reduce the computational load for a given import sequence. Any interval of time can be scheduled into the framework for a loop cycle, at the need and discretion of the agency. For further implementation of the methodology, IBM® has a command prompt utility, Series Importer®, which opens Analyst's Notebook®, if not already open, and automatically runs designated import specification(s) to load data into the chart without manual user intervention (IBM, 2018). A batch file (.bat) can call the Series Import utility as many times as desired which any extracts and charts new entries (IBM, 2018). Upon

import, the new entities are highlighted for the user to observe any linkages and draw conclusions. This restricts the workload for implementation of the model to preparation of case details in a spreadsheet (import specification stage 1). In the US, numerous agencies have transitioned to entirely digital platforms for case information, export from these report management systems (RMS) can also be achieved to ease concerns over the time and effort consumption. Additionally, most agencies within the US have dedicated crime analysis sections which would accommodate the impact this model would have on work flow.

4.2. Impact on criminal justice system and law enforcement

Forensic intelligence analysis aids in identifying relationships or patterns in crime events. A transition from reacting or responding to crimes toward the anticipation of crime events is needed. Major scale contribution of information is without question ideal for the decision-making process. Value from forensic case data and forensic intelligence is maximised if it is a consolidated, multidisciplinary product (Bruenisholz et al., 2016). The true value of the information is the product that can be given to administration for informed operational decision making. Intelligence-led policing needs to be driven by information gathered during forensic analysis and investigations. Forensic intelligence analysis is shown relative to and integrated with the traditional criminal investigation process in Figure 13.

Other implications of forensic intelligence are the potential to close more cases. The potential to link cold cases in the network through similarities (e.g. *modus operandi*), new evidence, or stimulate a lead to refocus the investigation are all provided through the intelligence network. An example of such circumstance was provided in Figure 9.

Integration of information in the intelligence network aids interpretation through linkages among databases which are typically isolated. Figure 10 shows how all of the information can be compiled to assess the evidence. Furthermore, graded linkages can be statistically based (e.g. likelihood ratio, match score, Random Match Probability) showing the degree of belief in a linkage further facilitating case assessment. There is also potential for excerpts of the intelligence network to be presented to juries during

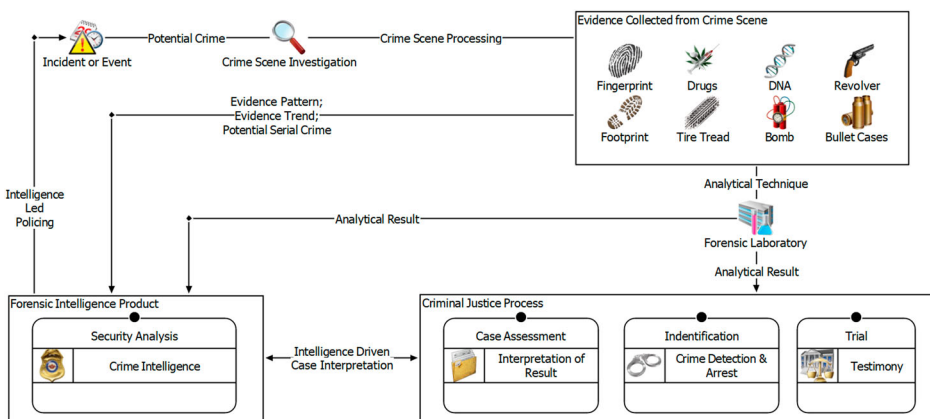


Figure 13. The investigative process (model) showing the role of forensic intelligence. Based on the information presented by Ribaux et al. (2010).

court proceedings. This could aid the explanation of the evidence in the case, and the strength of the evidence linkages (if graded).

4.2.1. Facilitating a working relationship

Sheptycki (2004) presented several organisational pathologies that impair the functionality of intelligence systems. These breakdowns occur because of the occupational subcultures within agencies which lead to information hoarding among agencies and institutional friction (Sheptycki, 2004). Difficulties for moving information across bureaucratic boundaries are expected and will remain until the value is observed. Collaboration is critical to advancing the system currently in place and fostering an effective criminal justice system. Adopting problem-oriented or intelligence-led policing offers the possibility to enhance value of forensic operations while preserving privacy. The methodology presented posits that the formal and structured use of data enables the processing of information as a first step. Implementation of such a system can occur within an agency and still provide intelligence for informed decision making and policing. Once agencies evaluate these results and experience the benefits intelligence-driven policing can provide, attention will be given to overcoming the existing structural challenges yielding collaboration and sharing between agencies. Development of a working relationship only seeks to induce cooperation in solving complex investigations, many of which often involve separate agencies. Organisation will give a more informed effort and is achievable through the drawing of structured information from databases and transforming the law enforcement databases into a single unified input as presented. Collaboration between law enforcement, labs, and their systems also has the potential to prevent cases which have reached resolution from being analysed saving both time and resources.

Communication or interaction between the scientist, investigator, and defense team may facilitate the greatest assistance to the court. This interaction between parties and sharing of information may be well served with the intelligence model. Communication and coordinated efforts will be critical as the criminal justice system detects and tackles new forms of criminal activity.

4.3. Limitations and future directions

The import specification model developed through the project has a few limiting factors. The import specifications were developed using provided adjudicated and simulated case files. The information received was limited by what was kept on record by the agency. Additionally, access to examinations/analysis of articles of evidence were not provided. Spreadsheets used to import the data would have to be modified to accommodate an agency. An important consideration is that the linkages are primarily developed through detection of common elements/entities. The model presented would function with great efficiency in localised areas where a low number of offenders engage in the majority of criminal activity. If this is true, the targeting of police actions on a specific prolific group has the potential to drastically reduce the number of offenses and substantially influence the general level of criminality. If untrue, less linkages would be developed, however, linkage development will still occur in any environment if two entities of the same identity are introduced. Another limitation of this project is that the DNA database had to be simulated because access to CODIS was not achievable.

Another limitation exists in the scalability of the model presented. IBM has noted that an import of data in excess of 20,000 entities and/or links may result in a Microsoft.NET Framework Unhandled Exception error (IBM Support, 2015). The error results from the import process being internally limited to a 2GB maximum amount of physical memory (RAM) (IBM Support, 2015). This error is associated with a single import, not an overall chart size. Using this recommendation, a series of phone calls were simulated to assess the capacity of Analyst's Notebook®. The simulated call data included the following information: outgoing call number, dialed call number, and duration of call, all of which were randomly generated. The data was imported into Analyst's Notebook® resulting in linkages between the phone numbers (e.g. 'Outgoing_Number' → 'Dialed_Number'). A sequence of ten imports consisting of 20,000 call records was made into a single chart and the elapsed time for the import and charting of the records was recorded. Phone data was used as a test because generating phone data, or random number sequences can be readily accomplished. Furthermore, the phone data is similar to a match result from the forensic databases shown within the model; charting of a new entity (e.g. DNA profile icon) and a linkage between two entities (e.g. 'John_Smith' → 'DNA_Profile'). Benchmarks for the phone call import and resulting chart information is provided in Table 4.

To assess whether the entities or links are more computationally intensive, an import with only 1000 entities and 20,000 linkages was run. This resulted in an import duration of seven seconds compared to the 26 s with the same number of links and 20,000 entities (Sequence 1). This is likely due to the smaller amount of entities being ideally positioned on the chart area. IBM® has noted that charts with many items that contain a very large amount of data can affect performance of the software. However, it was found that nearly 400,000 entities and 200,000 linkages can be charted in under seven minutes. As a remedy, IBM® has developed the iBase® package (IBM, 2018) to facilitate the processing of large-scale data. Massive charts, entities, and links can all be stored in a database instead of a flat file which offers basic database functionalities and eases computational capacity for the data (IBM, 2018). Considering the information from Table 4 and the iBase® package (IBM, 2018), the three-step methodology presented appears scalable to accommodate entire databases used in investigative work.

A final concern exists in the legal policy and procedures that govern the application of information sharing necessary for forensic intelligence. Before an agency invests resources in the development of forensic intelligence systems, data integrity needs to be established with attention paid to the governance and legitimacy of the system (McCartney, 2015).

Table 4. Sample scalability benchmarks using a series of ten imports in sequence of simulated phone call data. Total entities and links within the charts are provided with the duration of the import for each of the imports performed and total import time.

Sequence	Total entities	Total links	Duration (s)	Total import time (s)
1	40,000	20,000	26	26
2	80,000	40,000	28	54
3	120,000	60,000	32	86
4	160,000	80,000	34	120
5	200,000	100,000	39	159
6	240,000	120,000	40	199
7	279,999	140,000	45	244
8	319,996	160,000	50	294
9	359,960	180,000	55	349
10	399,943	200,000	56	405

Domestic law enforcement agencies serve as producers, consumers, and administrators of forensic information/intelligence and are expected to expand their technological capacity to gather and disseminate forensic information and intelligence. With political expectations to expand policing operations, traditional parameters restraining law enforcement information sharing are increasingly inadequate. Oversight management of the information flows is critical to the development of standard operating procedures. Another concern is that the citizens lack the ability to know of, understand, and challenge exchanges of their data (McCartney, 2015). Additional legal issues pertaining to forensic intelligence systems are: ownership of data; storage of information and destruction requirements; legality of the intended use of forensic intelligence; evidentiary validity of data in court; human rights and privacy; oversight bodies; and information sharing (McCartney, 2015). No single approach can be made to overcome the legal challenges suggested, and a jurisdiction by jurisdiction approach would be necessary for implementation. Rossy et al. (Rossy, loset, Dessimoz, & Ribaux, 2013) recommend remote access to Laboratory Information Management Systems (LIMS) for investigators as it would add efficiency to the investigative processes. Depending on jurisdiction legislation, legal restrictions may pose a challenge to the development of a comprehensive intelligence framework as proposed and the associated crime intelligence database. Numerous works (Bruenisholz et al., 2016; Joyal, 2012; Ribaux et al., 2016; Rossy et al., 2013) cite IT systems and global data integration as a critical factor for true development and progression of forensic intelligence.

5. Conclusion

A paradigm shift is necessary to meet the growing information and intelligence needs of investigatory work. A new model encompassing a forensic intelligence framework was developed as a solution to reduce linkage blindness and enhance the analytical nature of criminal investigations. The proposed model combines a three-step import specification sequence that guides incorporation of information into the network. Subsequently; strategic, tactical, and operational efforts/actions generate information to drive informed investigations through link, pattern, and trend discovery. Instead of treating each case individually, a multi-case focus and more holistic approach based on the study of crime is recommended. Forensic outcomes have a great potential to detect crime series and the implementation of a forensic intelligence framework fully exploits that potential with a shift away from traditional case assessment.

The model presented in this article exhibits the effectiveness of import specifications for the introduction of case information, items of evidence, and analytical results into a forensic intelligence structure. Forensic intelligence serves to enable a contemporary case assessment methodology through the inclusion of all information related to and uncovered throughout a criminal investigation. Interpretation and presentation of crime can exist within the network enabling assessment within the landscape of crime in the network. Import specifications also allow for consistency between case information integration in a user-friendly package. Refinement of the model may be necessary to accommodate potential information as necessary, and for different agencies. Graphical depiction of incidents within a global perspective of all agency case adds value to the case assessment and interpretation. The multi-case approach offered by the proposed model has

great potential to assist in developing a more uniform approach and nomenclature across the forensic and investigative science community.

Acknowledgments

The authors wish to acknowledge Chief J. C. Corkrean for his assistance with this project. There are no funding sources to declare for contribution to this project. Additionally, the authors use or discussion of IBM® i2® Analyst's Notebook® does not constitute endorsement or sponsorship for the use of the software. The software was available through the IBM® Academic Initiative and therefore used for the project.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Betebenner, D. (2017). *Package 'randomNames'*. Retrieved from <ftp://cran.r-project.org/pub/R/web/packages/randomNames/randomNames.pdf>
- Bieber, F. R. (2006). Turning base hits into earned runs: Improving the effectiveness of forensic DNA data bank programs. *The Journal of Law, Medicine & Ethics*, 34(2), 222–233.
- Bruenisholz, E., Prakash, S., Ross, A., Morelato, M., O'Malley, T., Raymond, M. A., ... Walsh, S. (2016). The intelligent Use of forensic data: An introduction to the principles. *Forensic Science Policy & Management: An International Journal*, 7(1–2), 21–29.
- Bureau of Alcohol, Tobacco, Firearms, and Explosives. (2018). *National Integrated Ballistic Information Network (NIBIN)*. Retrieved from <https://www.atf.gov/firearms/national-integrated-ballistic-information-network-nibin>
- City of Chicago. (2012). *Chicago Street Names*. Retrieved from <https://data.cityofchicago.org/Transportation/Chicago-Street-Names/i6bp-fvbx>
- Cook, R., Evett, I., Jackson, G., Jones, P., & Lambert, J. (1998). A model for case assessment and interpretation. *Science & Justice*, 38(3), 151–156.
- Federal Bureau of Investigation. (2018a). *CODIS-NDIS Statistics*. Retrieved from <https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet>
- Federal Bureau of Investigation. (2018b). *Next Generation Identification (NGI)*. Retrieved from <https://www.fbi.gov/services/cjis/fingerprints-and-other-biometrics/ngi>
- Hill, C. R., Duewer, D. L., Kline, M. C., Coble, M. D., & Butler, J. M. (2013). U.S. Population data for 29 autosomal STR loci. *Forensic Science International: Genetics*, 7(3), e82–e83.
- i2 Analyst's Notebook User Guide. (2009). i2 Limited.
- IBM. (2018). *IBM Knowledge Center – about ActiveX Technology*. Retrieved from https://www.ibm.com/support/knowledgecenter/en/SS6V3G_5.3.1/com.ibm.help.gswapplintug.doc/GSW_About_ActiveX_Technology.html
- IBM. (2018). *IBM Knowledge Center – Example batch import*. Retrieved from https://www.ibm.com/support/knowledgecenter/SSXVXZ_2.2.0/com.ibm.i2.anb.doc/series_batch_import.html
- IBM. (2018, October 7). *IBM i2 iBase – overview – United States*. Retrieved from <https://www.ibm.com/us-en/marketplace/data-management>
- IBM Support. (2015, November 5). *IBM importing large amounts of data into IBM i2 Analyst's Notebook can cause a Unhandled Exception error – United States [CT741]*. Retrieved from <http://www.ibm.com/support>, <https://www-304.ibm.com/support/docview.wss?uid=swg21903175>
- Jackson, G., & Jones, P. J. (2009). Case assessment and interpretation. *Wiley Encyclopedia of Forensic Science*.

- Joyal, R. G. (2012). *State fusion centers their effectiveness in information sharing and intelligence analysis*. TX, El Paso: LFB Scholarly Publishing LLC.
- Kopp, I. (2007). *Review of resource needs in the Forensic Science Laboratory and the Wider Scientific Context in Ireland*.
- McCartney, C. (2015). Forensic data exchange: Ensuring integrity. *Australian Journal of Forensic Sciences*, 47(1), 36–48.
- Meissner, P. (2015). *Making R Files Executable (under Windows)*. Retrieved from <http://www.r-datacollection.com/blog/making-r-files-executable/>
- National Institute of Justice. (2018). *Predictive policing*. Retrieved from <https://www.nij.gov/topics/law-enforcement/strategies/predictive-policing/Pages/welcome.aspx>
- Ribaux, O., Baylon, A., Roux, C., Delémont, O., Lock, E., Zingg, C., & Margot, P. (2010). Intelligence-led crime scene processing. Part I: Forensic intelligence. *Forensic Science International*, 195(1–3), 10–16.
- Ribaux, O., Roux, C., & Crispino, F. (2016). Expressing the value of forensic science in policing. *Australian Journal of Forensic Sciences*, 49, 1–13.
- Ribaux, O., Walsh, S. J., & Margot, P. (2006). The contribution of forensic science to crime analysis and investigation: Forensic intelligence. *Forensic Science International*, 156(2–3), 171–181.
- Ripley, B. (2017). *ODBC connectivity*. Retrieved from <https://cran.r-project.org/web/packages/RODBC/vignettes/RODBC.pdf>
- Ripley, B., & Lapsley, M. (2017). *Package 'RODBC'*. Retrieved from <https://cran.r-project.org/web/packages/RODBC/RODBC.pdf>
- Robertson, B. (1990). John Henry Wigmore and Arthur Allan Thomas: An Example of Wigmoreian Analysis. *Victoria U. Wellington L. Rev.* 20, 181.
- Rosy, Q., Ioset, S., Dessimoz, D., & Ribaux, O. (2013). Integrating forensic information in a crime intelligence database. *Forensic Science International*, 230(1), 137–146.
- Rosy, Q., & Ribaux, O. (2014). A collaborative approach for incorporating forensic case data into crime investigation using criminal intelligence analysis and visualisation. *Science & Justice*, 54(2), 146–153.
- The R Project for Statistical Computing. (2018). *The R project*. Retrieved from <https://www.r-project.org/>
- RStudio Team. (2018). *RStudio: Integrated Development for R*. Retrieved from <https://www.rstudio.com/>
- Sheptycki, J. (2004). Organizational pathologies in police intelligence systems: Some contributions to the lexicon of intelligence-led policing. *European Journal of Criminology*, 1(2), 307–332.
- Sklansky, D. (2011). *The persistent pull of police professionalism*. Retrieved from <https://www.ncjrs.gov/pdffiles1/nij/232676.pdf>
- Williams, R. (2008). Policing and forensic science. In T. Newburn (Ed.), *Handbook of policing* (2nd ed, pp. 760–793). London: Willan.