Sharing Ballistics Data across the European Union

Richard S Wilson
C$_3$RI
Sheffield Hallam University
Sheffield, United Kingdom
r.wilson@shu.ac.uk

Lukasz Jopek
C$_3$RI
Sheffield Hallam University
Sheffield, United Kingdom
l.jopek@shu.ac.uk

Christopher D Bates
C$_3$RI
Sheffield Hallam University
Sheffield, United Kingdom
c.d.bates@shu.ac.uk

Abstract - Across Europe police organisations are using numerous systems, both computerised and manual, to capture information about firearm crimes. The Odyssey platform intends to address this issue by providing police organisations with the ability to access ballistics data from other European law enforcement agencies. The Odyssey platform is a prototype system that has been developed to identify standards for the development of a European wide ballistics information system. In this paper, we outline the investigation tools found within the platform and discuss how these were developed. The prototype has been demonstrated to law enforcement communities across Europe and is in its final stages of development.

Keywords – ballistics; sharing; law enforcement; data mining; Europe.

I. INTRODUCTION

Police organisations across Europe deploy many different systems, both computerised and manual, to record information about crimes which involve the use of firearms. Ballistics crime is relatively rare – just 0.2% of all crimes in the UK involve firearms, but justifiably this is taken very seriously by Law Enforcement Agencies. The disparate systems used today are well suited to resolving crimes when they are committed locally, but problems arise when crossing jurisdictions. Although individual offences are unlikely to cross borders, guns and bullets are moved between European Countries [4]. Exchanging information so that, for example, guns can be tracked or patterns of usage followed has always been difficult. Information is often exchanged via the telephone, email or one-to-one meetings. Before information can be shared, investigators need to have some indication that the ballistic item is linked to another offence.

The Odyssey project addresses this problem by providing users with access to a plethora of information from investigations across Europe. Incidents are evaluated to find ones which are *similar* to the one under investigation. Detectives are then alerted to these similar incidents. The Odyssey platform provides police users with access to information such as expert and witness statements, images and videos, as well as details of bullets, cartridge-cases and firearms. All of this information is presented through directed graphs or a historical timeline view of an investigation. Odyssey improves crime resolution times by facilitating communication between experts.

In this paper we will describe the data structures which underpin Odyssey. We will provide an overview of some of the most recent developments in police information management systems followed by a description of the Odyssey platform prototype. The paper concludes with a discussion about the standards that were identified through the development of the prototype.

II. BEYOND STATE OF THE ART

Internationally, there are a number of crime database systems in use by different law enforcement agencies. Some of these systems include: COPLINK, NABIS (National Ballistics Intelligence Service), HOLMES (Home Office Large Major Enquiry System) 2 and I-24/7.

COPLINK is an information and knowledge management system aimed at capturing, accessing, analysing, visualising and sharing information between United States (US) law enforcement agencies. COPLINK comprises of two components COPLINK Connect (CC) and COPLINK Detect (CD).

CC is designed to integrate disparate heterogonous data sources, including legacy systems, to facilitate information sharing between police departments. CC provides police officers with access to one central data repository, which allows them to carry out four types of independent searches (person, vehicle, incidents or locations). In addition to this, police officers can carry out partial and phonetic based searching, access previous searches and upload images and documents.

CD expands the functionality of CC to automatically find associations within police databases. It is aimed at supporting detectives and crime analysts in finding associations between people, vehicles, incidents and locations. At present the system is able to find associations between individual entities, but is unable to map them onto a geographical map. The strength of an association is determined through the use of co-occurrence analysis and clustering. The system is able to search for meaningful terms in both structured (database tables) and unstructured (witness statements) data [1].

UK police forces have access to a number of independent database systems. These databases are used to record, monitor and manage offences in such areas as sex offences, gun crimes and major incident management.

NABIS provides ballistic examination services, for twenty UK based police forces, through three hubs, which are based in London, Birmingham and Manchester [7].

The database is fundamental to the service that NABIS provides. The NABIS database supports the recovery and analysis of ballistic items. Associations between the ballistic information, people, objects and events are formed to create tactical intelligence. Security is implemented on a per-user basis so that users are only able to access information that is relevant to their role [6].

HOLMES2 is an information management system which assists police forces in the investigation of serious crimes, such as serial murders, large scale fraud and major disasters. HOLMES2 lets police forces share information and identify links between independent incidents. HOLMES2 is an Oracle based database that resides on a UNIX system [9].

Across Europe, Interpol has implemented the I-24/7 global police communications system that allows 188 member countries to share information about criminals and criminal activities. I-24/7 provides member countries with 24 hour access to a vast array of police information. Such information is related to suspected terrorists, wanted persons, fingerprints, DNA profiles and stolen vehicles. In addition to this, the I-24/7 system provides each member country with access to other member's national databases [4].

When ballistics crimes are investigated, recovered items such as guns, bullets or cartridge-cases can be compared by forensic specialists. Test-fired bullets are examined for a range of marks made as they pass down the barrel of the gun [2]. By comparing the marks on different bullets a trained examiner can determine if two bullets come from the same weapon. The process of examining bullets under a microscope is time-consuming and difficult. A number of systems such as IBIS, Papillon and EVOFINDER have been built to automate the evaluation process; however, these systems do not inter-operate [11]. A bullet scan from one manufacturer's system cannot be used on anothers.

The Odyssey project is helping to define standards for sharing ballistics data between systems across Europe. Figure 1 provides some indication of the different ballistic matching systems in place across Europe.
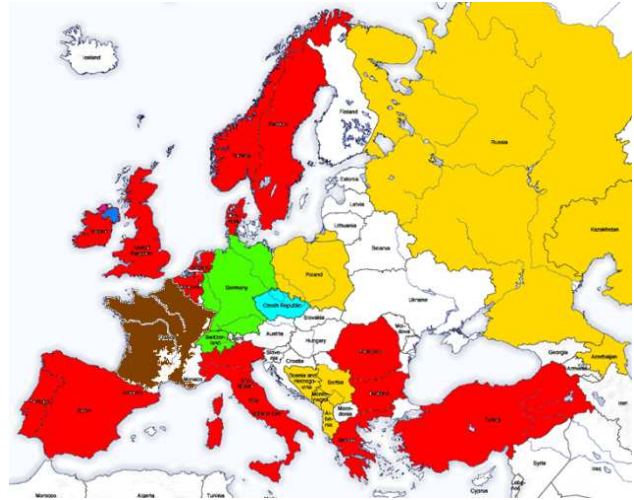

Figure 1. Ballistic systems in place across Europe

This is in contrast to the United States of America (USA), where IBIS has first mover advantage and has developed a centralised IBIS system. Furthermore, Odyssey is different to NABIS and the I-24/7 database currently in place in the UK and across Europe, as the NABIS database is specifically designed to manage the examination of ballistic items. The I-24/7 database has a European-wide database which largely retains information related to the individual. The I-24/7 system doesn't integrate data from ballistic systems due to the heterogeneous nature of data. Odyssey is different to other US and European police systems, as Odyssey seamlessly combines relational querying and data mining results from multiple different domains.

III. INTELLIGENT SEARCHING

Odyssey retains data within two main databases - a local database and a central database called CEON. The local database is maintained by the individual police organisation, whilst CEON holds ballistics data uploaded by police organisations across Europe. The data within the local database is replicated within CEON through an XML transfer structure, which is updated on a daily basis. The central database is interrogated, using supervised and unsupervised data mining techniques, to find associations with other ballistic incidents that have been committed across Europe. It is anticipated that the central database will handle at least six hundred thousand new ballistic incidents per year. This is based on the average number of firearm offences committed in the UK between 1999 and 2009, multiplied by the number of member states [3]. Each police organisation with a related ballistic incident is alerted to the fact that an association has been found. An alert is generated when a potential match is identified in the central database. An automated message is then sent back to the user to inform them that a potential match has been found. This message is delivered to an e-mail type inbox within the GUI (Graphical User Interface). In addition to this, the platform

allows individual police organisations to restrict access at the individual ballistic incident level to their data.

Extracting intelligence using the GUI is achieved in two stages. First, the user *defines the search item* then they *browse the results,* which can then lead to further searches. Figure 2, below, shows a search that has been carried out through the Odyssey GUI.
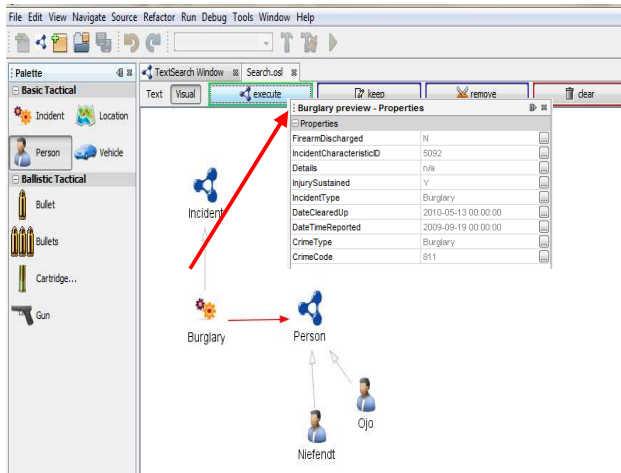

Figure 2. Odyssey GUI

First, the user selects a *basic or ballistic tactical item* and refines their search in the *search properties*. The user then has the option to add further tactical items and link them to the results of the previous search. Searches are refined, in the platform, through the search properties which converts the query into OSL (Odyssey Semantic Language). Querying with OSL allows the user to access information directly from the database and combines it with intelligence from the data mining backend. Below is an example of a simple OSL query generated through the GUI.

QUERY person WHERE weapon HAS VALUE Sig Sauer P238 AND country HAS VALUE France

Expanding the query to determine similar ballistic incidents by using the data mining backend and within a confidence limit is expressed as follows:

QUERY person WHERE weapon HAS VALUE Sig Sauer P238 AND country HAS VALUE France WITH CONFIDENCE > 0.7

Having identified the high level information, through the GUI, the user is able to drill down into it. This data is contained within the Odyssey Evidence Repository (OER). The OER provides access to some of the physical artefacts collected through the investigation process, including: images, videos and expert and witness statements.

In addition to this, Odyssey also provides historical data views of related incidents. These views are presented to the user as a timeline. This promotes transparency as the user is

able to identify any updates that have occurred as the investigation has progressed, which includes the identification of additional ballistic incidents and changes made to the underlying information.

The intelligent searching services found within Odyssey's arsenal are vital to facilitating the resolution of ballistic incidents committed across Europe.
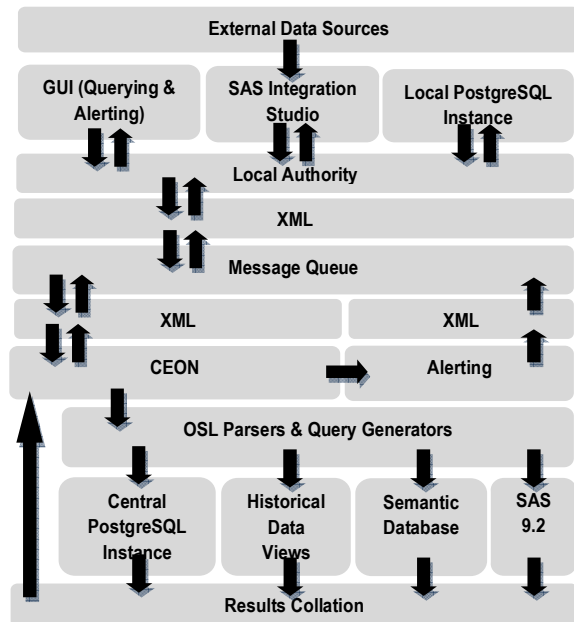
IV. ODYSSEY HUB

Figure 3. Architecture of the Odyssey Platform

The Odyssey platform is pragmatic software that can run on any standard off the shelf system. It was developed through the integration and repurposing of open source and common off the shelf software applications. These included Java, NetBeans, PostgreSQL, Antlr and SAS. Whilst the platform currently utilities SAS for the integration and mining of data the software could quite easily be replaced with open sourced software such as Python and WEKA. Figure 3 above provides an overview of the architecture of the Odyssey platform.

External data, from European police organisations, is extracted and transformed using SAS Integration Studio. This data is then loaded into the local authority database, which is a replication of the central database. These databases were developed using PostgreSQL and are based on the database structures currently in place at NABIS, Europol and CiFEx (Centre for Information on Firearms and Explosives) – CiFEx are ballistic experts in the UK. This helped to focus the development of the databases towards ballistic incidents. This allows the user to manage their own data and helps to insure the smooth transfer and integration of data between the local and central databases. There are over fifty tables in the database that are linked using Object Identifiers (OIDs). An OID is automatically generated when data is uploaded into the database, which is unique across

the entire instance of the database. Referential integrity between items is not maintained by the databases. This causes scalability problems as associations between data items have to be made outside of the database. This issue is further compounded by the predicted size of the database, given the number of ballistic incidents per year. Indeed, as the platform stores images, videos and statements, the anticipated size of the database is expected to grow into the terabytes. A binary items table was therefore implemented within the database which works as a file system that points to the stored images, videos and statements.

Data is uploaded into the local component through the GUI. This is achieved through a number of automatically generated SQL insert statements, which is then speeded up through the use of OIDs. It is at this point where the user is able to specify the initial sharing permissions. This is done through the GUI, by selecting the *ALLOW* or *DENY* option followed by the *user specified options*. The *user specified options* allow the user to grant or restrict access at five levels: country, region, police organisation, department and user. These options are then changed into numerical codes that are later translated, along with the insert statements, into OSL. These codes are then retained within the permissions table of the authorisation database. Data is loaded into the local database through the GUI, see Figure 4.
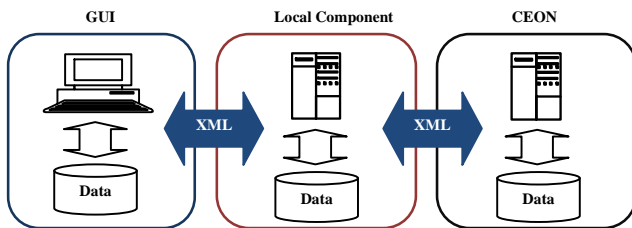

Figure 4. Odyssey XML Transfer Structure

The GUI translates the OSL into SQL which is used to pass data to the local database as XML. The data is then transferred to CEON using the same XML structure, where it is interrogated for associations. Through the use of SAS integration studio, we have demonstrated that it is possible to extract, transform and load data from ballistic and incident systems into a relational database structure.

When a match is identified an alert is then triggered and sent back to the local component. If permission is granted to view the data, the related data is also sent back to the local users through the XML structure. Following on the data is then retained in the cache of the local component where it is serialised within the platform. The platform also allows the user to modify their data, which is done through the translation of SQL update or delete statements into OSL. The process of uploading and returning a result (alert and data) is completed in twenty four hours. The user is restricted from directly changing the data in the central database due to batch data mining and user processing. A change to the central database would occur within twenty four hours of a user modifying the local database. When the

update has occurred the users receive an alert that asks them to re-run their query.

The shared data is based on the options specified by the user, which is retained in the permissions table within the authorisation database. The shared data is sent back to the local component through the XML structure where it is serialised within the platform. It is presented to the user as a related result within the GUI. Figure 5 below shows the results of an Odyssey search.
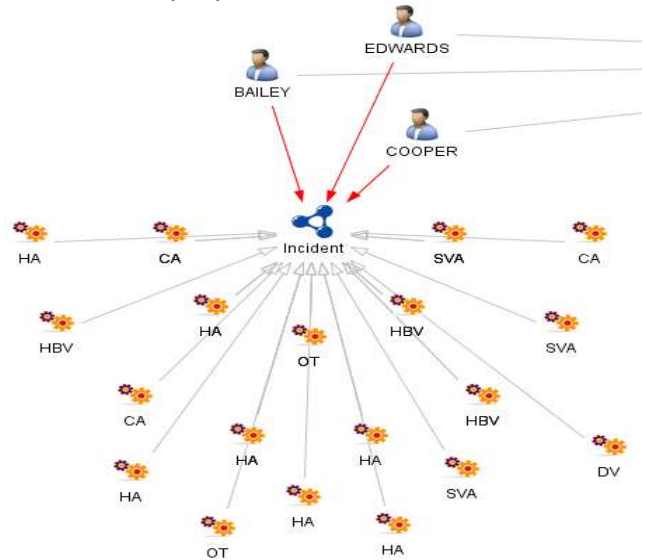

Figure 5. Search Result

This shows the related crimes committed by Edwards, Cooper and Bailey.

Querying the database is carried out through using the GUI, or by creating OSL with the built in assisted functions. Querying with the GUI, the user drags a tactical item from the tool bar and specifies the search options in the search properties. The user specified options are transformed into OSL using the backend semantic engine. OSL is a grammar that was developed in Antlr, which translates the search into SQL and is specifically designed to hide the underlying structure of the database from the user. The semantic language was created through understanding how police investigators communicate and think during investigations and by capturing domain specific knowledge about the meaning of police language. Below is an example of OSL and its translation into SQL:

QUERY ballistic incident WHERE weapon_manufacturer HAS VALUE Sig Sauer AND victim_gender HAS VALUE female

```
SELECT *
FROM odyssey.ballistic_incident
LEFT JOIN ballistic_incident_has_recovered_firearm ON
(ballistic_incident_has_recovered_firearm.recovered_firear
m_oid = ballistic_incident.oid)
LEFT JOIN ballistic_incident_has_recovered_firearm ON
```

```
(ballistic_incident_has_recovered_firearm.recovered_firear
m_oid = recovered_firearm.oid)
LEFT JOIN ballistic_incident_has_case ON
(ballistic_incident_has_case.ballistic_incident_case_oid =
ballistic_incident.oid)
LEFT JOIN ballistic_incident_has_case ON
(ballistic_incident_has_case.ballistic_incident_case_oid =
case.oid)
WHERE case.gender_of_victim = "female" AND
recovered_firearm.manufacturer = "Sig Sauer";
```

If the user specifics a confidence limit then Odyssey will return the results ranked in descending order of confidence. The confidence is calculated on a daily basis, by applying prebuilt algorithms to a merged copy of the central database tables the Odyssey data mining mart (DMM). The data is merged using SAS Data Integration Studio and is scored using prebuilt algorithms that were created in SAS Enterprise Miner. The score is returned to the ballistic incident table by matching the OID from the DMM with the OID in the ballistic incidents table. The ballistic incidents table is then sorted in descending order of score. The following is an example of an OSL query with an expressed confidence limit, which is translated into SQL:

```
QUERY ballistic incident WHERE weapon_manufacturer
HAS VALUE Sig Sauer AND victim_gender HAS VALUE
female WITH CONFIDENCE > 0.7
```

```
SELECT *
FROM odyssey.ballistic_incident
LEFT JOIN ballistic_incident_has_recovered_firearm ON
(ballistic_incident_has_recovered_firearm.recovered_firear
m_oid = ballistic_incident.oid)
LEFT JOIN ballistic_incident_has_recovered_firearm ON
(ballistic_incident_has_recovered_firearm.recovered_firear
m_oid = recovered_firearm.oid)
LEFT JOIN ballistic_incident_has_case ON
(ballistic_incident_has_case.ballistic_incident_case_oid =
ballistic_incident.oid)
LEFT JOIN ballistic_incident_has_case ON
(ballistic_incident_has_case.ballistic_incident_case_oid =
case.oid)
WHERE case.gender_of_victim = "female" AND
recovered_firearm.manufacturer = "Sig Sauer" AND
ballistic_incident.score > 0.7;
```

Documents are stored in a separate PostgreSQL database as blobs. Using local compatible software, police experts are able to share information across the platform. It helps to facilitate communication between experts, as they are able to access documents from other law enforcement agencies, along with their contact details. Such communication is believed to be vital in helping to resolve crimes that have been carried out in different locations [8].

Historical data are presented to the user as a timeline of events, which shows any updates to the incidents and any associated incidents that the user has the authority to view. An update is defined as the user deleting, updating or inserting data into an existing ballistic incident. The timeline of historical events is created through using another PostgreSQL database, which stores the delta along with the author and date/time of the change. The updated item is identified in the historical database by the OID of the initial item, combined with an SQL timestamp. The approach adopted by the platform is quite unique as other historical databases that use OIDs with timestamps have retained the historical data in the same database. The main argument to support this is that consistency and speed of data retrieval is maintained when the historical data is stored in the same database [10]. Whilst speed of retrieval and consistency of data items is important to the platform, the initial aim is to return the latest view of the ballistic incident to the user. The speed in which historical data is returned to the user is however improved, through the use of the OID and timestamp, as the requested data is easily ranked through sorting these values in ascending order.

V.   TOWARDS COMPLETION

Further development of the platform will focus on the automatic identification of key words in expert and witness statements – the Odyssey Statement Miner (OSM). This will present the user with a list of key words, ranked in order of occurrence, from the associated statements. We expect that this will help the police experts to assess the usefulness of the documentation.

The OSM will be developed using SAS Text Miner in conjunction with a bespoke java programme, which removes any conjunctive adverbs from the statements. SAS Text Miner will be used to identify any words that occur more than three times within a statement. From this a dictionary of words, along with the frequency of occurrence will be created. This will then be presented to the user as a list of words in descending order of frequency. It is hoped that this will help the user to decide whether or not to view the statement.

It is noticeable and disappointing that no EU-wide standards exist for secure police data systems. Odyssey has been able to demonstrate that widely available open software can be repurposed easily to build such systems. The project will be recommending a list of such technologies.

VI.   CONCLUSION

This paper provides an overview of some of the most current crime information management systems in place in the USA, UK and across Europe. It highlights that there is a need for a platform which combines data from the different ballistic systems currently in place across Europe. An overview of the functions found within the prototype are also discussed, which focuses upon accessing, loading,

sharing and querying of data. The development of the investigation tools within the platform are also explained and future work in relation to the completion of the prototype is outlined. The paper concludes with a brief discussion regarding the standards that will be recommended at the end of the project.

REFERENCES

[1] Chen, H., Zeng, D., Atabakhsh, H., Wyzga, W., and Shroeder, J. (2003), *COPLINK Managing Law Enforcement Data and Knowledge,* Communications of the ACM, January 2003. Vol. 46, No.1. pp. 28-34.

[2] Bundeskriminalamt. (2004), Firearm Type Determination, Available [online]: https://www.forensic-firearms.bund.de/ [17/06/2010].

[3] Gun Control Network. (no date), Firearm Offences – England and Wales, Available [online]: http://www.gun-control-network.org/GF05.htm [17/06/2010].

[4] Interpol. (no date,) Connecting Police: I-24/7, Interpol, Available [online]: http://www.interpol.int/Public/ICPO/FactSheets/GI03.pdf [17/06/2010].

[5] Leon, F.P. (2005) 'Automated comparison of firearm bullets', Forensic Science International, Vol. 156, February, pp. 40-50.

[6] National Ballistics Intelligence Services. (no date), *Database*, Available [online]: http://www.nabis.police.uk/database.asp [16/06/2010].

[7] Sims, C. (2010), National Ballistics Intelligence Service Update Report, West Midlands Police Authority, Available [online]: http://www.west-midlands-pa.gov.uk/documents/committees/public/2010/12_PerfandOps_22April2010_National_Ballistics_Report.pdf , [16/06/2010].

[8] Travis, J. (1998), 'Informal Information Sharing Among Police Agencies', National Institute of Justice, December 1998.

[9] Unisys. (2007), What is Holmes 2?, Unisys, Available [online]: http://www.holmes2.com/holmes2/whatish2/ [17/06/2010].

[10] Van Oosterom, P. (no date), *'Maintaining Consistent Topology including Historical Data in a Large Spatial Database'*, Cadastre Netherlands. pp. 327-336.

[11] Yates, S., Jopek, L. Johnson Mitchell, S., and Wilson, R. (2009), *'Semantic Interoperability between Ballistic Systems through the Application of Ontology'*, IADIS WWW/ Internet Conference. Vol. 2 pp. 153-157.