System Integration and Re-engineering Using XML/Web Services

Dr. Michael J. Hu
Police Information Technology Organization (PITO)
New King's Beam House, 22 Upper Ground
London SE1 9QY
Telephone: +44 208 358 5402
Email: michael.hu@pito.pnn.police.uk

ABSTRACT
Use of latest technologies, such as the Internet, to enhance the government services to the public in the United Kingdom has been underpinned by the British government's emphasis of several key e-initiatives, including e-policing and integration of criminal justice IT (CJIT) systems. This paper presents an overview of several on-going and planned system integration and re-engineering programmes in the Police Service and CJIT in the UK using XML & Web Services, as well as key findings and lessons learnt from undertaking these programmes.

General Terms
Algorithms, Management, Performance, Reliability, Experimentation, Security, Standardization.

Keywords
XML, Web Services, system integration, system re-engineering.

1. INTRODUCTION
A wide range of IT systems have been developed during the last decades within the fifty-two police forces in the United Kingdom. Many of these systems are currently being upgraded with added functionality, while new systems are continuously being developed and rolled-out. Integrating such a wide range of new and legacy systems in the Police Service, and linking these systems to those in the CJIT communities (including the British Crown Court, the Crown Prosecution Service, National Probation Service, etc.) and other government agencies (e.g., the Home Office), has been undertaken under the guidelines of an national Information System Strategy for the Police Service published in 2002 [1]. Further effort has been made to provide European Union-wide or even global police intelligence & investigation, and counter-crime capabilities. A number of system integration and re-engineering models have been deployed at the national and regional/local levels. This has resulted in a number of different integration strategies and migration paths towards the XML/Web Services based enterprise information architecture that are discussed in this paper.

WWW 2003, May 20-24, 2003, Budapest, Hungary.
ISBN 963-311-355-5.

- **System-to-System Messaging —** The integration of legacy systems using a system-to-system messaging model, has been deployed in several forces including the Metropolitan Police Service (New Scotland Yard). However, the general trend is to move away from the traditional approach of tight coupling between service requestors and information/service providers, and converge to a more flexible enterprise solution using SOAP messaging.

- **Broker-Based Data Request and Response —** Integration via the request-and-response broker based architecture, such as Enterprise Java Beans (EJB), has been deployed for the Police National Computer (PNC). Within the UK, the PNC provides a key centralized information repository containing criminal records etc. PNC responds to enquiry from a wide range of users from police officers on duty, to staff in the European Intelligence Bureau.

- **Data Gateway and Data Centre —** The model of integrating legacy systems via data gateways or meta-data centres, has been used in several regional forces. The concept is to load the data from legacy systems into central or regional data gateways, in which meta-data can be extracted and legacy data schema mapped to a common data (interchange) schema. A "data centre" would further extend such a data gateway and meta-data service to provide a data warehouse and data mining capabilities, for intelligence analysis and investigation for the Police Service. An ongoing project to set up such a data centre, the Cross Regional Information Sharing Project (CRISP), is designated to serve about ten police forces in North West of England.

In this paper, we review each of these system re-engineering models, and their convergence to XML/Web Services based enterprise information architecture for the Police Force in the UK. In Section 2, we first outline the enterprise information architecture, as a foundation for integrating and modernizing the Police IT systems in the United Kingdom. In Section 3, we discuss each key re-engineering model in more detail, as well as the potential options for migrating it to the enterprise architecture. We also present key issues and lessons learnt in implementing these models, as well as our comments and recommendations with regard to the Web Services Architecture that has been proposed recently by the Web Services Architecture Working Group.

2. XML/WEB SERVICES BASED ENTERPRISE ARCHITECTURE FOR THE POLICE SERVICE

It is envisioned that the **XML/Web Services based enterprise information architecture** for the Police Service constitutes a three-tier architecture, as it is shown in Figure 1.

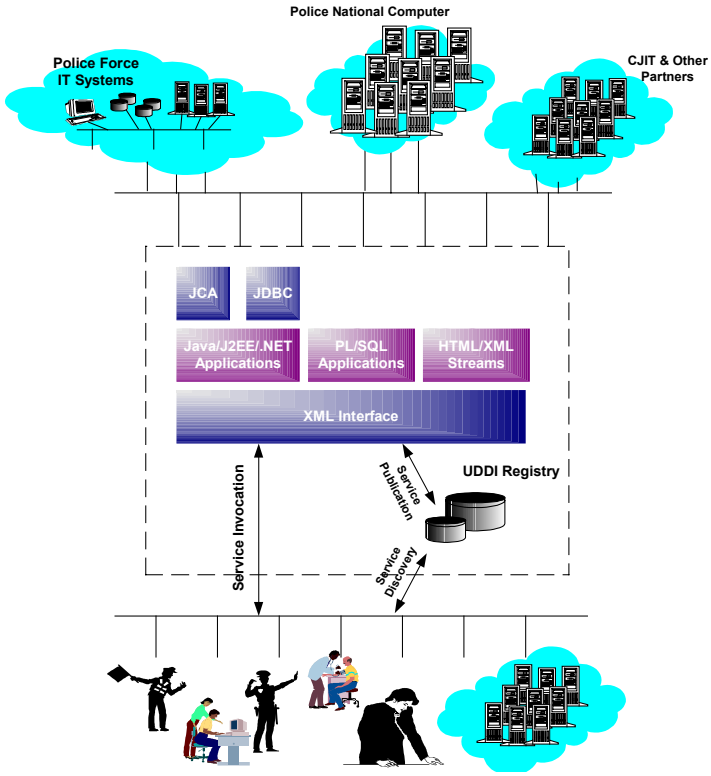

Figure 1 Enterprise information architecture for the Police Service

- **Information and data** stored in local police forces and in the Police National Computer (PNC), as well as those provided by other relevant government agencies (as business partners), form the core information assets for the Police Service. Such data and information are accessed via a wide rage of **enterprise services**.

- **The Police National Network (PNN) and Criminal Justice eXchange (CJX)** provide a secure/network backbone connecting the end users to the enterprise information base. **Other middleware and service components** in the middle tier offer information/service directory, switching and binding, security, and logging services.

- **Police Users** are the primary users accessing the enterprise information base using desktop computers and mobile devices. **The public** will soon be able to access non-confidential information via the Police Portal.

XML has been widely adopted internationally as the lingua franca for data interchange between computer systems; its impact on current and future development in all public and private sectors world-wide is enormous. In the UK, XML has been adopted as the basis for data standard and data interchange in all government ministries and agencies, through the e-Government Interchange Framework (e-GIF) [2]. Web Services, as it is defined by the W3C, is a software system identified by a URI whose public interfaces and bindings are defined and described using XML. As a result, its definition can be discovered by other software systems (as service requestors). These systems may then interact with the Web Services in a manner prescribed by its definition, using XML based messages conveyed by internet protocols [3][4][6].

The XML/Web Services based enterprise architecture provides a loosely coupled (and therefore more flexible) yet robust environment, that complies with the Information System Strategy for the Police Service. With such architecture, service requestors do not have to worry about the location of information assets and enterprise services requested, nor the operating systems/ programming languages/component models used to create or access these enterprise services. It is, therefore, ideal for integrating, modernizing and re-engineering Police IT systems.

3. CONVERGING TO THE ENTERPRISE ARCHITECTURE

While a number of large-scale system integration and modernization programmes have been undertaken or are being planned in the Police Service and in CJIT in the UK, most of them are currently based on the re-engineering models outlined in Section 1. The forward plan is to converge these different models and strategies using the XML/Web Services based enterprise architecture shown in Figure 1. In this section, we discuss two key elements as pre-requisites for such convergence: 1) A standard data structure or data model, and 2) An enterprise-wide component structure or catalogue. Thereafter, we discuss these different re-engineering models in more detail, focusing on their deployment and the key issues that we have encountered.

3.1 ENABLERS FOR CONVERGENCE

In our experience, two key enablers to convergence are: 1) A shared data structure or data model across the entire enterprise; and 2) A unified component structure shared by enterprise services.

Corporate Data Model (CorDM)

Establishing a standard data model as a baseline for legacy systems, new applications, and emerging enterprise information solutions, is the first step towards the XML/Web Services based enterprise architecture. The data model provides the vocabulary (e.g., data types etc.), based on which the XML schema for inter-service messaging can be developed. This makes it possible for

the enterprise service providers to announce their services using the Web Services Description Language (WSDL) [4].

Secondly, a standard data model provides the standardized reference point for system integration and re-engineering. In other words, it provides a data-layer convergence point for legacy systems and new enterprise information solutions for the Police Service, while the former are undergoing the modernization and data migration process.

PITO has been leading the definition and implementation of a Corporate Data Model (CorDM) that provides common data standards for the information used in the Police Service and CJIT community in the UK [5]. One of the key inputs has been the UK government's the e-Government Interoperability Framework (e-GIF) published by the Office of e-Envoy [2]. The CorDM also formulates the basis for common message schema used for inter-service data exchange and transmission using Simple Object Access Protocol (SOAP) [3][4].

One of the key lessons learnt in a pilot project in the Metropolitan Police Service (MPS) is that adopting CorDM for service publication by legacy systems, and using it as a base for message schema for SOAP messaging, is absolutely key to the success of integrating these legacy systems, and moving towards the XML/Web Services based enterprise architecture [10]. Another lesson learnt in the project was that mapping the legacy data format(s) to the standard message schema was a lengthy and extremely difficult task. As the result, in the earlier stages of the project, publication of MPS existing services was not implemented in full, and mapping the legacy data formats to the standard message schema was not possible though originally planned, due to time constraints and technical difficulties. This resulted in a partial implementation of MPS' Web Services, in which SOAP messages were exchanged between one legacy system (as service requestor) and another (as service provider) in the format of system-to-system messaging, or XML-RPC (remote process call). As the message format is proprietary between these two systems, and binding is constrained to two specific port (types), it drastically limited the potential of the system architecture.

Corporate Component Catalogue for the Police Service

To fulfill the vision of providing independent and distributed enterprise services available over the Internet, which are usable on any platform from any development language, and are reusable/sharable among different services, a standard component structure or catalogue is needed.

In the UK, we are at the inception and planning stage of developing a Corporate Component Catalogue for the Police Service. In preparation, several local or regional forces and the Police National Computer (PNC) have been piloting and testing some components that are developed either in house, or by third parties [10].

3.2 CONVERGENCE PATHS

Convergence of different system integration and re-engineering models that are currently deployed in the Police Service in the UK, to the XML/Web Services based enterprise architecture, is considered an important step towards the Information System Strategy for the Police Service. Different migration strategies may be adopted by each individual police force as part of its system re-engineering and modernization process, taking into consideration its existing system architecture, legacy application portfolio and on-going service development.

1. **From system-to-system messaging to Web Services**

   Transforming system-to-system based messaging architecture to the XML/Web Services based enterprise architecture constitutes the challenge that many police forces face in the UK. The key is to upgrade the messaging component from the traditional XML RPC to the more loosely coupled SOAP messaging, by standardizing the message format(s) and introducing a more flexible service invocation process using WSDL and SOAP. Key steps in such a transformation process can be summarized as follows:

   - Service publication using WSDL. This includes definition of enterprise services as a set of network endpoints or ports (which associate an address with a specific binding), and definition of messages using data types defined in the CorDM. Messages are grouped into operations, which are further mapped into a portType with a binding. Two WSDL primitives, *request-response* and *solicit-response*, are currently applied in the pilot project in the Metropolitan Police Service. The former is used for service request and request acknowledgement, and the latter for service delivery. The other two primitives, i.e. *one way primitive* and *notification primitive*, will be added when other messaging exchange patterns (MEPs) are implemented.
   - Extending SOAP messaging, and making full use of its potential as a simple and lightweight mechanism for exchanging structured and typed information between peers in a decentralized and distributed environment.

   It is noted that the features included in the basic and extended Web Services Architecture published by the Web Services Architecture Working Group are not adequate for our implementation. This will be discussed in more detail in Section 4.

2. **Web Services: registration, publication, and discovery**

   The UK Police National Computer (PNC) is currently using the request-and-response broker based architecture to provide a wide range of centralized services to its users and business partners. Its convergence to the XML/Web Services based enterprise architecture is currently being defined and tested within the PNC Modernization Programme. Key activities within this programme are discussed below:

The first challenge is to set up and provide a standard PNC registry, using the Universal Description, Discovery and Integration (UDDI) standard [3][4][6][7], and to publish PNC's services and interfaces. As it is shown in Figure 2, the PNC UDDI Registry includes the entities/information stored in PNC Data Servers and a range of distributed enterprise services that PNC offers to its users and partners, as well as those offered by its partners. The PNC UDDI Registry provides the standard UDDI service functions including:

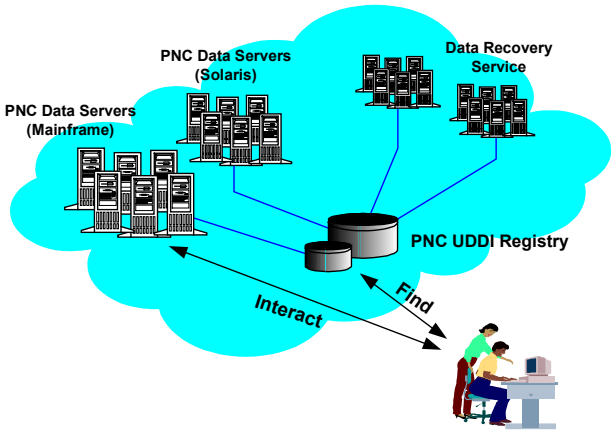

Figure 2 Police National Computer (PNC)

- **Publish** PNC registers its current services, adds new services, and issues its entity list to its partners and users;
- **Find** The PNC users (as service requestors) interrogate the PNC UDDI Registry, and discover the Web Services provided in the PNC, or by its partners;
- **Interact** The PNC UDDI Registry specifies how a service requestor connects to and interacts with the Web Services provided in PNC, and the associated access control and security requirement.

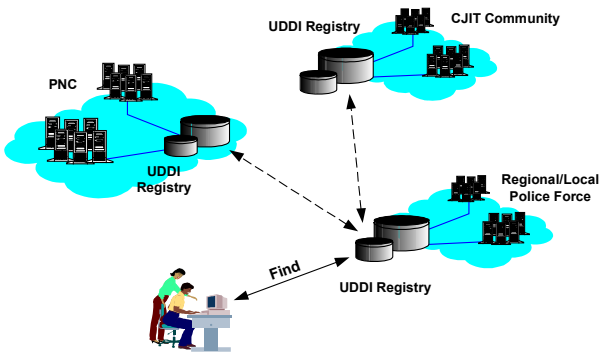

Figure 3 UDDI Registry: A hierarchical directory service

There has been some discussion in the PNC Modernization Programme about redefining the Police National Computer as the Enterprise Nerve Centre (ENC) for the Police Service and the CJIT communities and as the main gateway of these communities to access other e-government initiatives in the UK. We are also currently reviewing PNC's strategic position in the context of XML/Web Services based enterprise architecture for the Police Service, in which centralized, federated, and distributed enterprise services co-exist with one another.

This implies that different UDDI registries may emerge in the near future, possibly in different hierarchical levels as suggested in Figure 3. A local or regional UDDI registry may serve one or more local or regional police forces, whereas PNC UDDI Registry becomes an enterprise-level service registrar/service directory for the entire Police Service in UK. Moreover, it may become a key UDDI registry node in the UDDI Service Cloud, interconnecting other government agencies and the CJIT community.

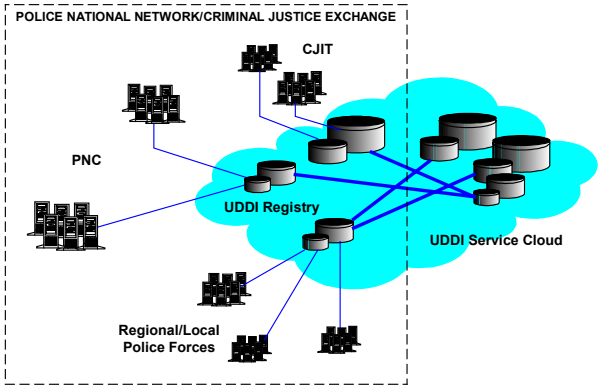

Figure 4 Secured UDDI Registries

Furthermore, these UDDI registries (including PNC UDDI Registry, regional/local UDDI registries, etc.) will become part of Internal Enterprise Application Integration UDDI, a secured UDDI Service Cloud, sitting inside the secured/confidential environment of Police National Network (PNN) and Criminal Justice Exchange (CJX), as it is shown in Figure 4.

The PNC UDDI Registry can also play a key role in the system/service re-engineering process in PNC, when: 1) Some legacy data or systems may be relocated from the centralized PNC mainframe to some regional data centres, or even to local forces; 2) Some legacy data or systems may be replaced by upgraded enterprise services; 3) New enterprise services are introduced and rolled out, either in PNC, or somewhere else, but registered in the PNC UDDI Registry.

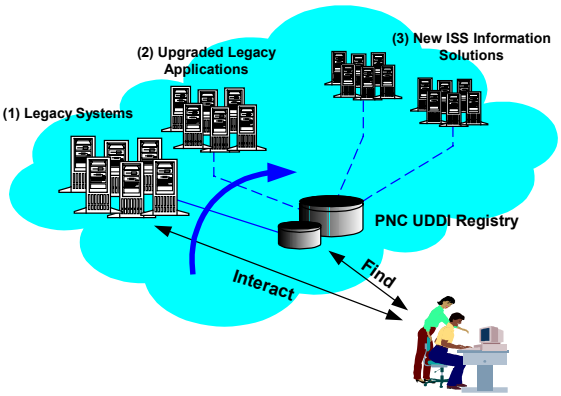

Figure 5 System re-engineering and management using PNC UDDI Registry

Figure 5 shows a conceptual view of system re-engineering scenarios using PNC UDDI Registry, migrating PNC legacy systems [(1) in Figure 5] to some interim solutions (i.e. upgraded or wrapped version of these systems, as shown in (2), and finally to the new enterprise information solutions (3). During such a migration process, it is only necessary to update the information stored in the PNC UDDI Registry. This simplifies security and increases system reliability as it prevents conflicting changes at different operator nodes.

3. From data gateway to XML/Web Services based enterprise architecture

The concept of using regional or centralized data centres for integrating and re-engineering enterprise services in the Police Service has been successfully applied several local and regional police forces. Figure 6 shows the conceptual architecture of such a system, in which the distributed enterprise entities are pulled (regularly) into a regional or central data centres, where data is extracted and transformed before being loaded into structured object bases. The resulting object bases then serve one or more of the following purposes:

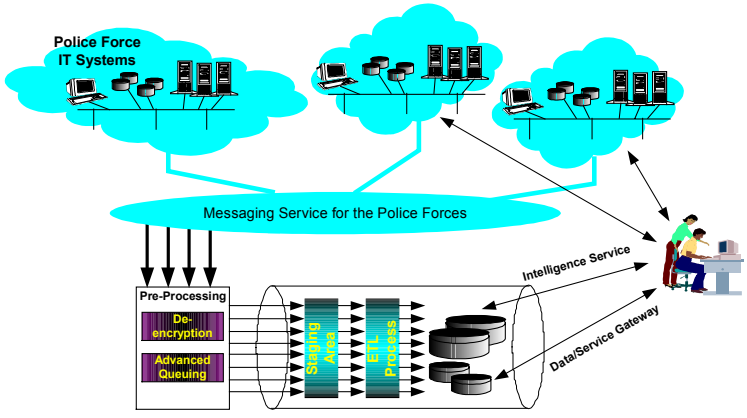

Figure 6 Centralized and regional data centres

a) To serve as a data gateway and/or meta-data centre, in which an entry in the object base points to the physical location of a specific entity in a local police force, or in the Police National Computer. In other words, the data centre provides some of the functionality of an UDDI registry, by providing the location information of the requested service/entity to service requesters (Figure 6);

b) To facilitate intra-legacy application data/information exchange by providing a standard data structure (which in most cases conforms to CorDM);

c) To act as an intermediate stage for system re-engineering, in the same way as using UDDI Registry as an intermediate for system re-engineering (Figure 5).

d) To provide intelligence and investigation services for the police forces, using data marts and data warehouse features to process and mine the entities and abstracted data stored in the object bases.

Migrating the data gateways/data centres to the XML/Web Services based enterprise architecture requires separation of the above features and functionality, and replacing and upgrading the service using UDDI, WSDL and SOAP messaging (Figure 7), as described in more detail as follows:

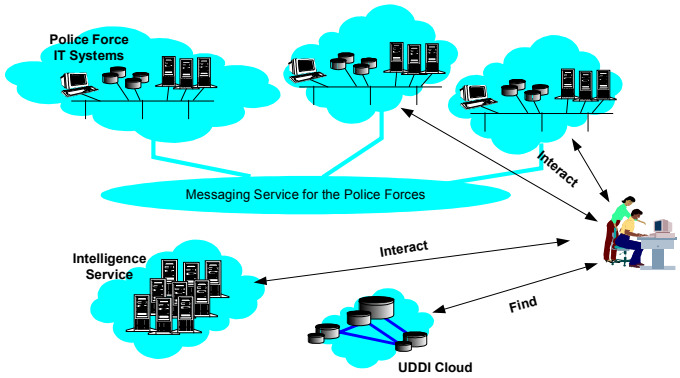

Figure 7 Enterprise services for the Police Service

- Replacing and upgrading a) with UDDI registries, and extending the services by introducing advanced search and query capabilities. By converging the data gateway or meta-data centre service provided by different object bases, to a standardized UDDI information model (which conforms to CorDM), the service/entity directory service could be extended beyond the local and regional boundary, and become part of the enterprise directory. The rich services provided by the UDDI, such as the service type registration, white pages, yellow pages, and green pages, will be valuable for service discovery.
- It would be ideal to automatically populate a UDDI Registry in the enterprise, and automatically update the registered entities, as what has been implemented in the object bases. This will be discussed in more detail in Section 4.
- Feature b) can be replaced using the WSDL and SOAP messaging.

- Feature c) is still valid, for the same reason as is summarized on earlier in the paper (and Figure 5).
- Feature d) should remain as a valuable enterprise service, providing intelligence analysis and data mining functionality to the service requesters in the police forces. It remains as an open question whether it is necessary to pull the data into the object bases (as data warehouses) on a regular basis, so that data mining can be executed efficiently on such "snap-shot" data, or it is achievable to pull the distributed data into a data mart or data warehouse at run-time using SOAP messaging. This may largely be dependent on the scalability of SOAP messaging, which has not been fully addressed in the Web Services Architecture Working Group.

4. WEB SERVICES ARCHITECTURE

The Web Services Architecture Working Group (WSA WG) has published the Web Services reference architecture, which includes the basic architecture and extended architecture [3][4][6]. The Web Services Architecture places into relationship various components and technologies that comprise a Web Services "stack" or functional implementation. While valid implementations may include subsets or parts of the stack, they must at least provide the components defined in the basic architecture. Components and technologies that extend the basic architecture are represented within the extended architecture.

- **Basic Architecture**

  The basic Web Services Architecture defines an interaction between software agents as an exchange of messages between service requesters and service providers. It includes Web Services technologies capable of:

  - Exchanging messages
  - Describing Web services
  - Publishing and discovering Web Service descriptions

- **Extended Architecture**

  The extended Web Services architecture incorporates additional features and functionality by extending the technologies and components defined within the basic Web Services Architecture. A partial list of these features includes:

  - Asynchronous messaging
  - Attachment
  - Caching
  - Messaging exchange pattern (MEP)
  - Reliable message
  - Message authentication
  - Message confidentiality
  - Message integrity
  - Message routing
  - Management messages

Some of these features, such as asynchronous SOAP messaging have been implemented in the various projects in the police forces such as the Metropolitan Police Service. Reliable SOAP messaging was implemented in our pilot projects with acknowledgement from the service providers, and using caching at the messaging components. Other advanced features, such as several messaging exchange patterns, are planned to be implemented in later stages of these projects.

A number of issues have been identified in our projects, which we believe have not been fully addressed by WSA WG. We summarize these issues as follows, as comments to the Web Services Architecture published by WSA WG [3], and as recommendations to the Web Services software developers/ manufacturers:

Messaging Exchange Pattern (MEP) An MEP is a specialized form of feature that describes a generalized pattern of message exchange between two services [3][4][8][9]. We found that the MEPs published by WSA WG are not comprehensive enough to cover some of the messaging patterns in our implementations. For instance, issues emerged in one of our projects that some or all of the following MEPs would be needed for the advanced features designed in the system:

- One-to-many messaging
- Broadcasting and multi-casting messaging
- One-request-multiple-responses messaging (ORMR) or multi-request-one-response (MROR) messaging
- Cascading messaging

We believe that some of these, called the *composite MEPs*, could be constructed by using a number of *primitive MEPs* or other composite MEPs. It is suggested that WSA WG look into the issues. If it is allowed to construct a MEP from other MEPs, rules should be specified as part of the Web Services Architecture.

Message authentication, confidentiality, security, and request/access control There exists a comprehensive security and confidentiality marking scheme covering all the enterprise services/networks in the Police Service, as well as all the entities and information stored in the Police IT systems. We believe that SOAP does not currently provide adequate message authentication and confidentiality features. We also identified that the security issue of Web Services needs to be addressed. Furthermore, request/access control should be enforced by the service provider before responding to a service request, ensuring that the confidential services and information would only be sent to the adequately "vetted" service requestors. Such an issue becomes even more challenging if other MEPs such as cascading messaging are permitted in the enterprise architecture.

Private UDDI registries We foresee two types of UDDI deployment scenarios in the Police Service and the CJIT community: 1) Internal Enterprise Application Integration UDDI - a secured UDDI Service Cloud inside the Police intranet [i.e., the

Police National Network (PNN) and Criminal Justice Exchange (CJX)], which allows applications in different police forces, PNC, and other "vetted" partners to publish and find services. 2) Portal UDDI - the UDDI registry inside the PNN/CJX, but it allows external users to find operations on the registry that are accessible to the public [7].

Populating and updating UDDI registries As the enterprise services for the police forces could be highly time-critical, populating and updating the UDDI registries automatically needs to be implemented, by either "push" or "pull" deployment strategies.

5. CONCLUSION

In this paper, we present the XML/Web Services based enterprise architecture for the Police Service in the United Kingdom. We also discuss the main system integration and re-engineering models that have been deployed in the UK, as well as their convergence paths by different regional and local forces. Many of our programmes are still on-going, in the Police National Computer (PNC) Modernization Programme and in several pioneering police forces including the Metropolitan Police Service. We have summarized our experiences and forward plan in the paper, as well as the key issues yet to be addressed, hopefully as lessons to be learnt by others, and as useful comments to the WSA WG and software developers.

6. REFERENCES

[1] Association of Chief Police Officers (ACPO), United Kingdom, *ACPO Information Systems Strategy*, version 2.0, January 2002

[2] UK Office of the e-Envoy, *e-Government Interoperability Framework*, version 4.0, April 2002. It is also available at www.govtalk.gov.uk

[3] World Wide Web Consortium (W3C), Web Services Working Group, *Web Services Architecture*, 14 November 2002.

[4] World Wide Web Consortium (W3C), Web Services Working Group, *Web Services Glossary*, 14 November 2002

[5] Corporate Data Model Project Information is available on the PITO Web site at: www.pito.org.uk

[6] Walsh A. (Ed), *UDDI, SOAP, and WSKL – The Web Services Specification Reference Book*, Prentice Hall, 2002.

[7] Cauldwell, P. etc. *XML and Web Services*, Wrox Press Ltd, 2001.

[8] Oracle, *Oracle 9i Application Server (Oracle9iAS),* January 2003

[9] BEA Systems, *BEA WebLogic Enterprise Platform and Roadmap*, September 2002

[10] Due to the confidentiality of many of our projects, details have to been withdrawn from some of internally published papers while these projects are referenced in this paper. Those who may be interested are welcome to contact the author for more detailed information.