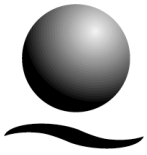
Learnings from the Monetary Penalties Enforcement Project:

Creating a document metaphor for information interchange – XML, schemas and transforms

Draft version 0.C 23rd March 2005



Content created by Quill Australia working with Jonathon Rees of the Department of Justice and the MPEP team

Report supplied by Tracey Rodgers, Project Manager, Monetary Penalties Enforcement Project

Introduction

L'exactitude n'est pas la vérité [1]

Henri Émile Benoît Matisse

Interactions between computer systems have always been conceptually simple yet technically frustrating.

Many factors have hindered the transfer of information. Different operating systems, different database vendors, different communication protocols and different business rules have hindered the meaningful transfer of information.

Now it is all much simpler. The communications protocols are basically agreed. Most systems can establish communication via TCP/IP. HTTPS is a widely accepted method of securing data being transferred. SOAP is widely accepted as the 'envelope' in which data will be transferred.

And now XML provides a means of structuring documents, XML Schemas provide a means of validating the structure and content of those documents and XML Stylesheets provide a means of displaying those documents in a way that conforms to business needs.

In keeping with the above quotation, the ideas presented in this short paper will provide some little assistance with the veracity of information but will enable great strides to be made in the exactitude of information.

Learnings from the MPEP Project

There have been a number of significant learnings arising from the development of the Criminal Justice Schema and associated XML artefacts which we will summarise before entering into the detailed explanation.

- Documents are the most appropriate metaphor for thinking about data that needs to be transferred between systems.
- Documents are best structured using XML.
- Any XML document that is to be transferred to another system should be validated against an XML Schema.
- XML Schemas should incorporate whole-of-government components wherever possible when describing data
- Wherever possible we should move away from the strictures imposed by legacy computer systems and use proper descriptions instead of codes (thus Australia rather than AU, Female rather than F)
- Wherever possible we should 'retire' objects that are no longer valid and record the date of the retirement (for example if legislation is repealed we

[1] *Exactitude is not truth*

should mark it as such and note the date so that no new offences are linked to that legislation)

- When an XML document is used as an authoritative source it should be stored along with a reference to both the XML Schema that was used to validate it and the XSLT stylesheet that was used to display the document.

The document metaphor

Thinking about information transfer in terms of documents provides a more immediately understandable and persistent metaphor for how to resolve the many transfer and subsequent storage issues that currently exist within and between organisations throughout the world[2].

This is a fundamental thinking point. Many people view data transfer in terms of values in fields being transferred between one database and another. This is an incomplete metaphor.

A previous document[3] introduced the metaphor of an envelope with an address to explain the role of SOAP and HTTP in ensuring that messages are packaged and delivered to the right destination. **In this paper we are looking inside the envelope to see what the letter looks like and how it should be constructed.**

Most electronic transactions can be thought of in terms of a business document. These documents have a form and structure that is inherent in the type of document.

When someone walks into a Service Tasmania outlet and pays their registration then some data is generated that forms the basis of a receipt. This can be thought of as a **receipt document**. The electronic information stored about the transaction should contain at least the data needed to re-create the physical receipt document given to the customer. A receipt document usually has the word Receipt as a heading. It then contains data elements such as the person's name, the date of the receipt and the amount.

When the Magistrates Court finds against an offender they produce a paper document called a **Memorandum of Sentence**. The electronic record of this outcome should be in the form of a document that contains sufficient information to re-create that Memorandum of Sentence. A Memorandum of Sentence has a defined structure with the heading **Memorandum of Sentence** at or near the top

---

[2] *Much of the thinking that has contributed to the document metaphor has been the result of discussions with Jonathon Rees of the Department of Justice who developed a framework for this concept a number of years ago and who has been a tireless advocate for the adoption of a more conceptual thinking model across Government.*

[3] *A Preferred Method for Interactions with the FIND system. MPEP document.*

along with certain legislative references. It also contains information about the offender, information about the finding and a set of orders.

When a **Water Licence** is issued to a primary producer, the information contained in the electronic document should also contain sufficient information to re-create the transaction.

Also, when one computer system notifies another that they have received a transaction then that notification can be thought of as a **notification document**.

Validating a document

When a document is created in XML format it should first be checked against an XML Schema to determine that it conforms to the set of business rules that have been agreed upon by the party creating the document and the parties authorised to receive the document.

This is shown in Figure 1 below:

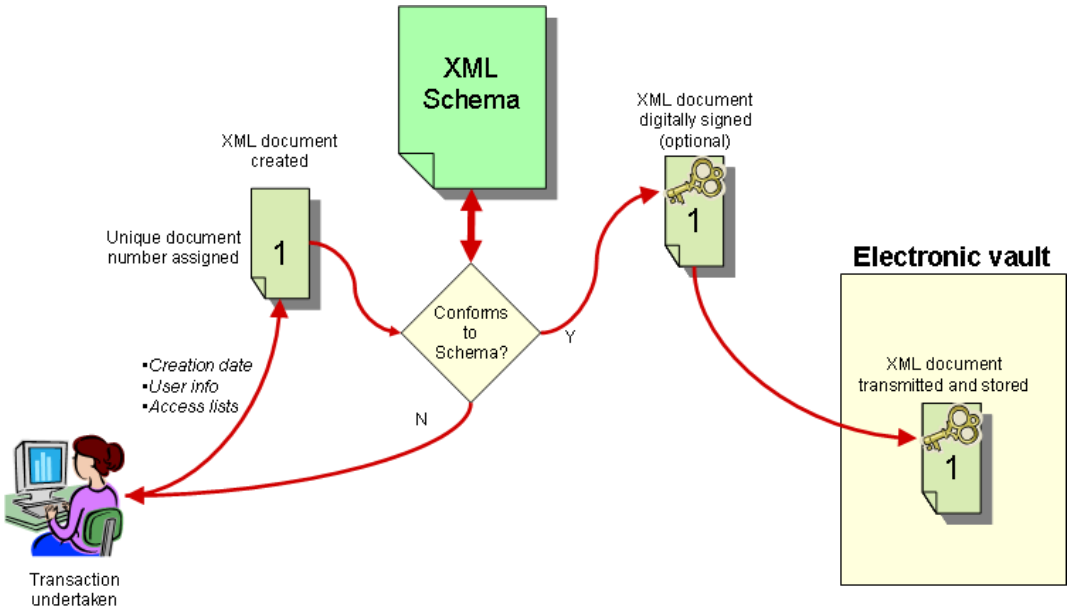


Figure 1

The role of the XML Schema will be explored in more detail later in this document. At the moment it is best to think of a schema as the contract between the business entities that describes what data will be present and what format it will be in.

Before and during transmission of the data from one computer system to another, various security devices can be employed to ensure that it is not available for malicious viewing or editing. The document can be digitally signed, it can be transmitted over secure transmission channels and it can even be encrypted if high levels of security are required.

Changing a document

When the receiving application interprets the data in the XML document and makes the necessary changes to its relevant databases, an additional action should be taken in future. A copy of the XML document that was used as the 'authority' to make these changes should be stored for future reference[4]. This improves the level of auditability of the data and the systems.

The authoritative source or the copy of the authoritative source should also be stored with references to the version of the XML Schema that was used to validate the document and, if a viewable version of the document was used, then a reference to the XML Transform (Stylesheet) document should also be stored.

This implies that all changes to both schemas and stylesheets should be made under strict version control so that when a schema or stylesheet is updated, the former version of the document is stored along with the date that it became deactivated.

If a mistake was made then the original document should not be altered. The entire document is re-issued and the original should be kept and flagged as de-activated and the date that it was de-activated should be associated with the document. The correcting document should be linked via a unique document identifier to the document it replaces.

The diagram below assumes a check against the schema as per Figure 2.

[4] *In a sophisticated implementation, only the URI of the authoritative document would need to be stored.*

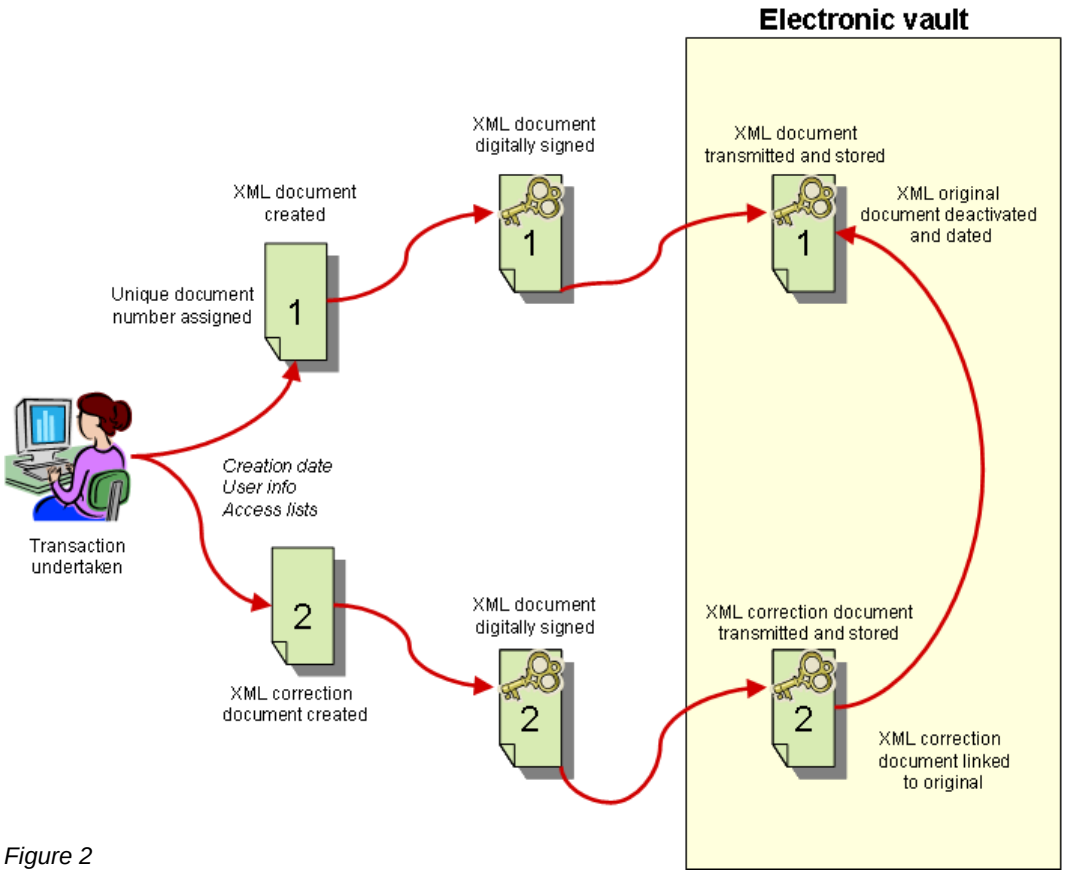


Figure 2

By storing an electronic record about such transactions, the transactions themselves can later be validated and audited and there is also a complete record of what happened and when and by whom[5].

Creating a persistent artefact

If data are transferred between two systems using traditional methods then artefacts relating to that transfer are not created and maintained with auditable persistence.

However, there is now a better way.

Creating an XML document that contains the data that needs to be transferred produces an artefact that is:

- Valid
- Persistent
- Auditable

[5] *It is often the case in legacy systems that audit data is difficult to retrieve and in many cases does not provide a complete picture of the transaction.*

- Signable
- An authoritative source

Each of these items will now be briefly addressed.

Valid

The document conforms to the requirements of the XML Schema that the business entities have agreed to.

Persistent

Because the XML document is stored, then it remains as a record of the transaction at a point in time. When the document is stored it should also have as attributes information relating to the user who created the transaction, the date and time of the transaction and information about the schema used to validate it and the transform used to display it if necessary. This means that it can then be recreated at some time in the future.

Auditable

Because the document contains all the information relating to the transaction and because it can be viewed in a document format it is easy at any point in time to audit the document.

Signable

A key feature of XML technology is that a robust standard is in place that allows XML documents to be digitally signed. This allows a check to be carried out in the future to determine if the contents of the document have been changed in any way. The signing of documents means that we can, with confidence, use the electronic record as an authoritative source.

This document can be stored in a database for later retrieval and checking. But the important thing is that a persistent artefact has been created which can be stored in a retrievable location. This retrievable location could be a URI (or IRI)[6] or a specially designed 'electronic vault' that provides levels of security access appropriate to the sensitivity of each document.

An authoritative source

A valid, electronically signed and properly stored XML document is as much a valid source as a paper copy of the document that has been signed. As such, organisations are rapidly moving to electronic originals as authoritative sources of the document.

[6] *Uniform Resource Identifier and Internationalized Resource Identifier*

Case Study: The Tasmanian Criminal Justice Schema

The Tasmanian Criminal Justice Schema has arisen out of work by Jonathon Rees and others in the Department of Justice in the development of a Criminal Justice Framework.

When the Monetary Penalties Enforcement Project was established and it became clear that significant changes would need to be made to the DoJ's CRIMES System in order to transfer robust data to and from the new FIND System (that was to be implemented to enforce monetary penalties), the two parties agreed that it would be of mutual benefit in the short term and of much wider benefit in the medium term to carefully define a schema that captured all of the business processes required in the transfer of data between these systems.

It was also agreed to use national and international standards in defining types within the schema wherever possible.

A draft of this schema is now being used by developers of both systems and the draft has also been socialised with a number of other stakeholders as well as with agencies who currently or who are likely to maintain systems that send or receive data to compliant systems.

The design of the schema started at the conceptual level. Before a line of XML was generated, the business processes of the courts were mapped in a great deal of detail and the artefacts received by and generated by the Courts were analysed carefully.

As a result of this analysis a high level view of the major types of documents relevant to the court process was described via a graphical view of the schema as shown in Figure 3 below.

The decision was made to attach a persistent identifier to every document produced by the courts and that these documents could be classified into one of the following areas.

- Pre-sentence document
- Sentencing document
- Post sentence application or notification
- Post sentencing order

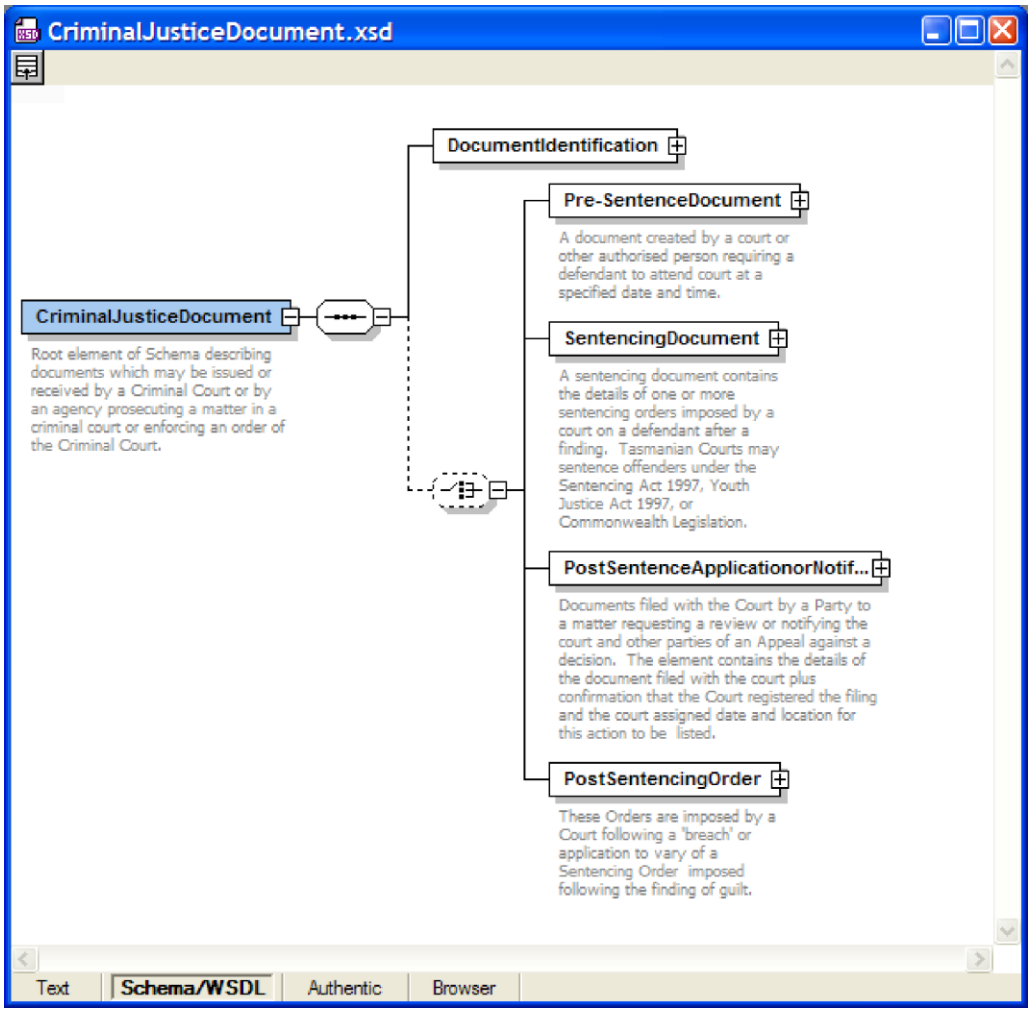


Figure 3

Notice in Figure 3 above that each element displayed in the schema is accompanied by some level of documentation. This has turned out to be a significant benefit of the project. The schema is clearly documented thus giving all parties a significant level of documentation.

One of the four document types that were identified was the Sentencing Document. Sentencing documents are, in turn, broken down into smaller components.

Each document will have some general data associated with it. For example a sentence will be associated with some legislative reference such as the *Sentencing Act 1997*. The document to be produced will have a title. The hearing will have started and finished on a particular day and have been held in a particular court in front of a judicial officer. The defendant needs to be uniquely identified and so on.

Therefore a structure of a sentencing document can be built up as shown in Figure 4 below.

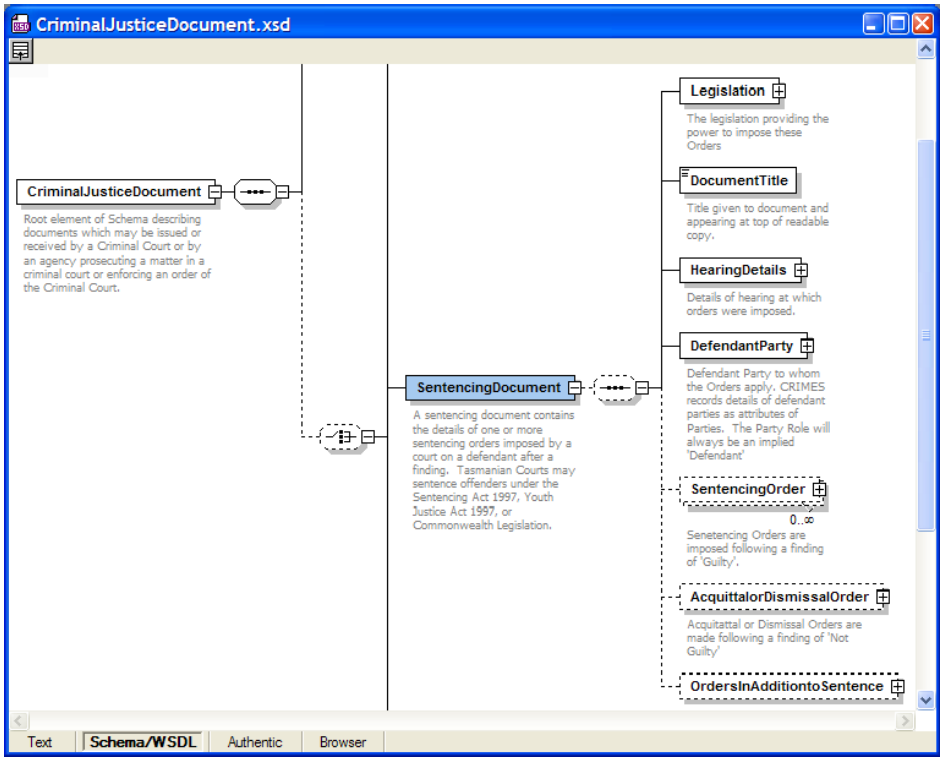


Figure 4

When the Legislation element is expanded we see a new feature of a schema as shown in Figure 5 below.

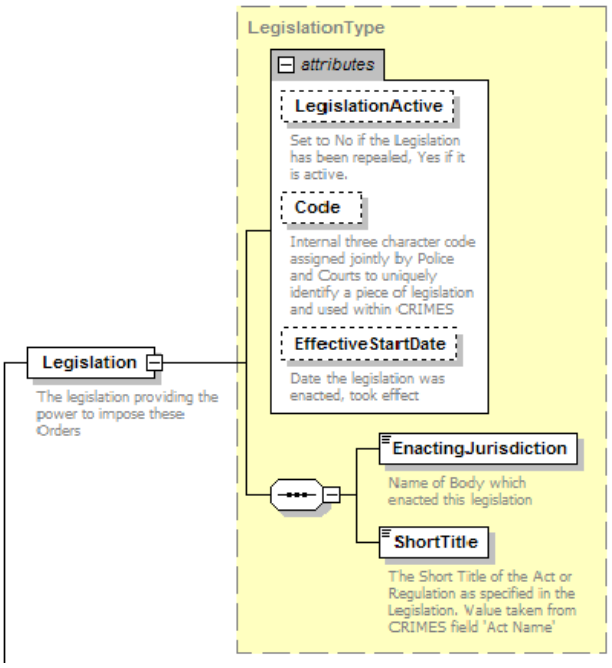


Figure 5

Here we have some items that will occur in a number of places and which will always be structured in the same way. We have therefore declared LegislationType to be a global type which can be used anywhere in the schema. Notice that there are also some attributes associated with this type as legislation is introduced to become active from a particular date and may become inactive as the result of the introduction of new legislation or having been repealed.

A similar reuse benefit occurs with identity information. In the domain within which this schema has been created there are three major entity types – these are natural persons, organisations and the 'joint' entity (where vehicle registration has been enacted in joint names).

If the entity is a natural person then their identity should be established through their name, sex, birth date and address. Other data may also need to be recorded.

This is shown in Figure 6 below.

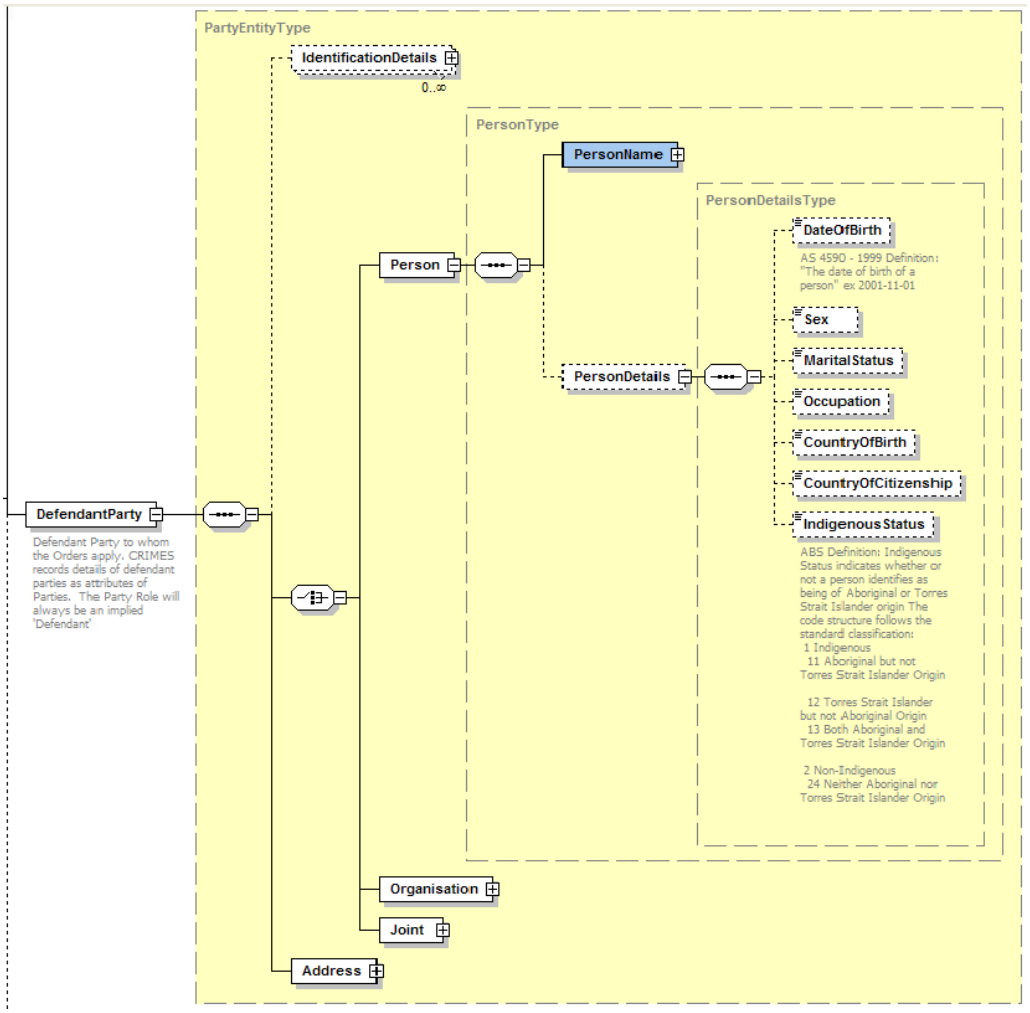


Figure 6

Further validation can be added to a schema by assigning restrictions to elements. So, for example, the DateOfBirth element shown in Figure 6 could be restricted is shown in Figure 7 below.

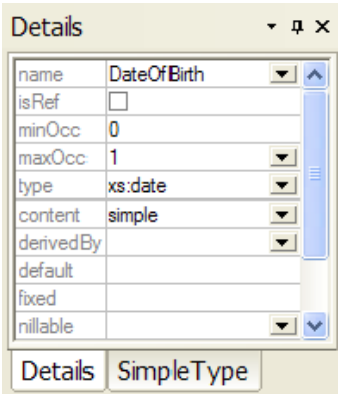


Figure 7

Notice that the element has been constrained to occur a maximum of one time and its type must a date type.

In Figure 8 below the Sex field is restricted so that only the values Male, Female and Unknown are possible.

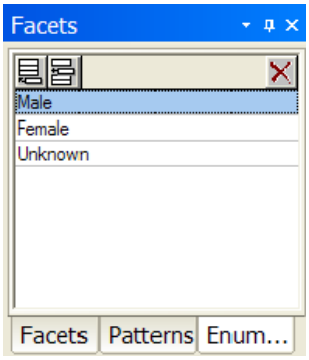


Figure 8

And this means that one person can't send a code of M while another sends Male.

The same applies to organisations. In the section of the schema shown in Figure 9, there are elements to hold the name of the organisation, the name type code to qualify the type of name being sent, as well as an organisation code that is conformant with those specified in AS4590.

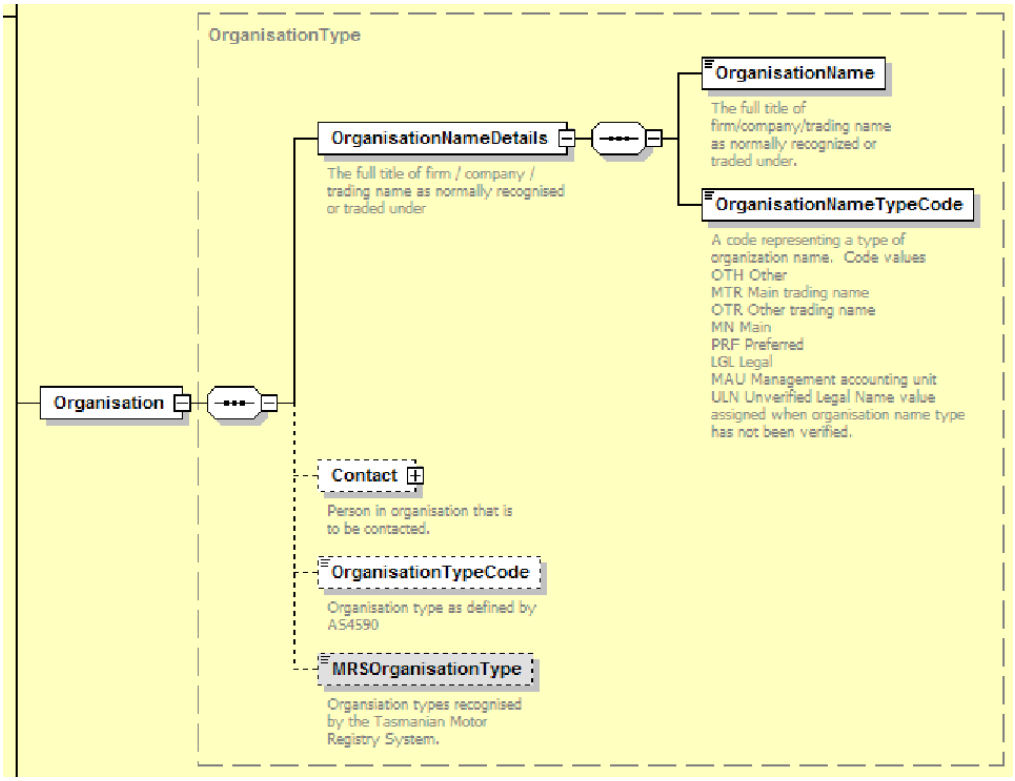


Figure 9

The short tour through the Criminal Justice Schema shown above should provide business managers and IT professionals with an understanding of how MPEP's FIND System will define and validate the documents that will flow into and from the new system.

It is quite clear that there are many opportunities for reuse. The work that has been done to define this schema could be used by many other agencies to implement a more rigorous approach to the definition of important whole-of-government data.

The XML Cornerstones

The approach adopted by MPEP and the Department of Justice has now been explained. The reasons why this approach was implemented is now addressed in relatively simple technical language.

There are many advantages to using an XML document as the primary source of information being generated by computer systems.

- An XML document is structured
- An XML document can be signed so that we can tell whether it is a true original
- An XML document has a creation date
- An XML document can contain a variable number of elements depending on the circumstances
- An XML document can be 'transformed' by a stylesheet to produce readily understandable output
- An XML document can be 'validated' against another XML document called a schema to ensure that it conforms to a set of agreed rules about its structure and its content.

First it is necessary to understand the role of each of the XML artefacts we have mentioned and what they are capable of delivering – and more importantly, what they are not capable of delivering.

An **XML document** is a collection of XML tags called elements that have a few rules associated with them.

Consider the XML document in Figure 10 below:

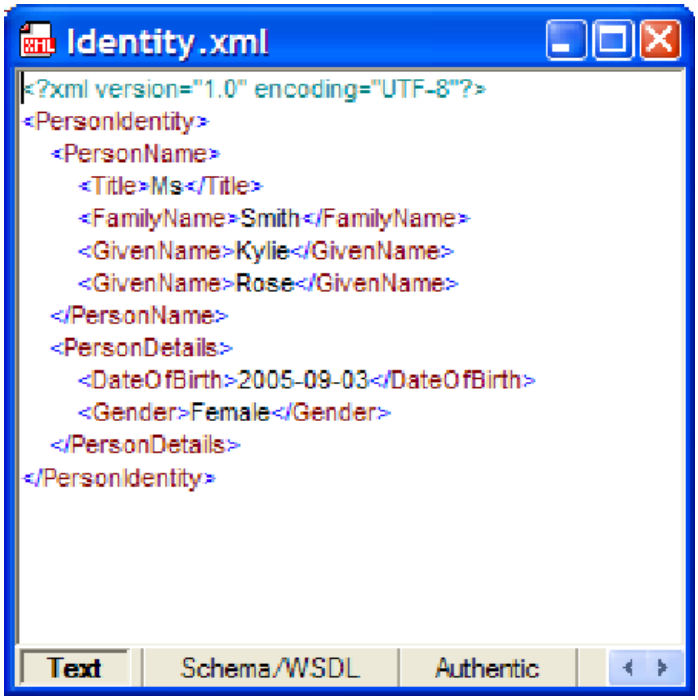


Figure 10

This is a 'well-formed' XML document that obeys the rules defined for XML. Notice that the data is enclosed by *tags* to create *elements*. Thus the Gender element is:

```
<Gender>Female</Gender>
```

The inherent structure of an XML document makes it easy for computer software to locate data items within it.

But how do we get it to obey some business rules that two organisations might want to impose for data moving between their systems?

This is where an **XML schema** comes in.

A schema defines:

- the elements that must be present in a particular XML document
- the order in which those elements must appear
- any restrictions on the data that an element can contain.

A sample schema for the XML document that we have created is shown in Figure 11 below:

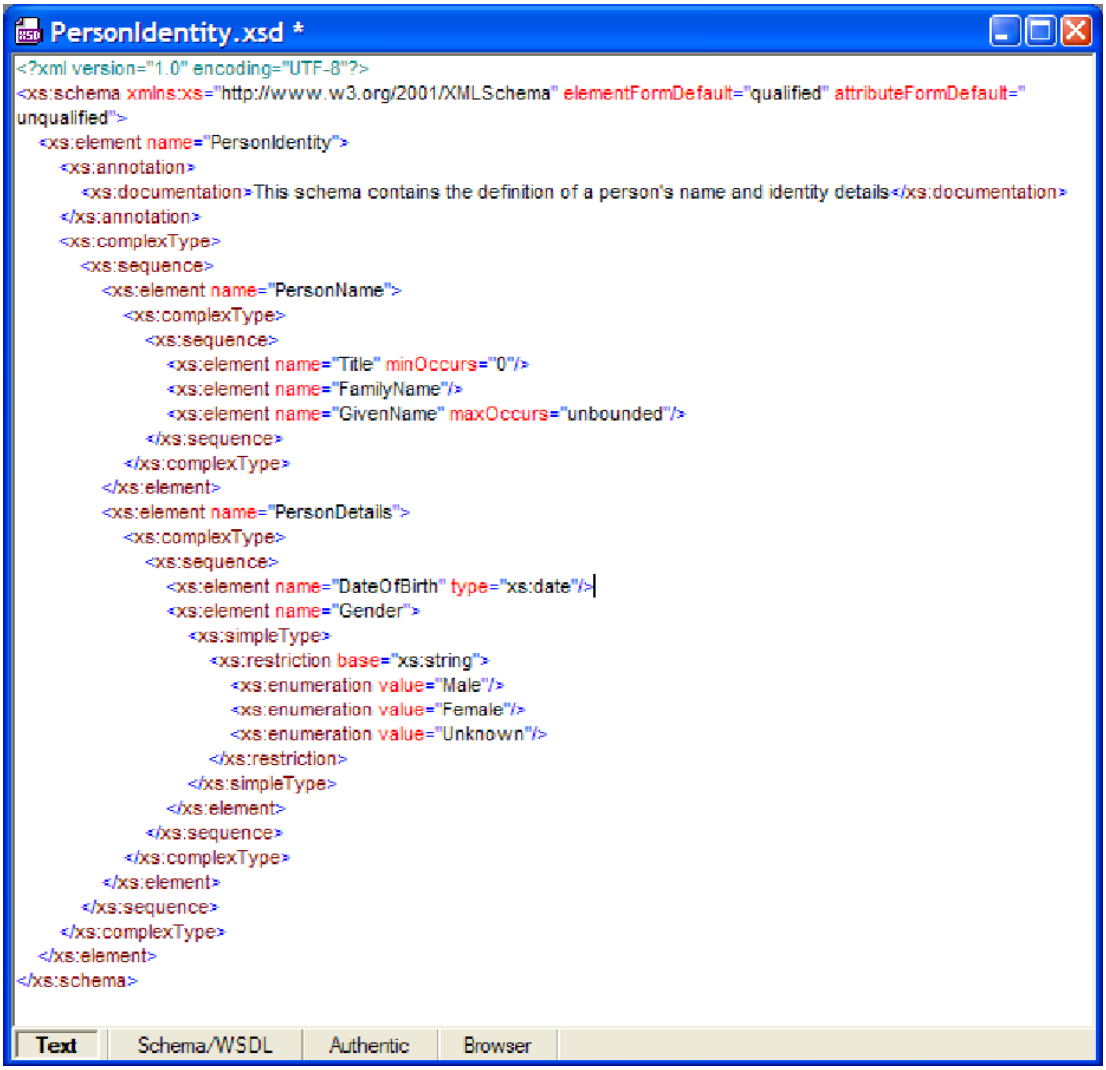


Figure 11

The schema contains the 'rules' the XML document must adhere to. Notice that the PersonName element contains three elements (Title, FamilyName and GivenName), however the Title element can occur 0 times which means it is not compulsory. Notice also that the DataOfBirth element will be checked to ensure that there is a date present and the Gender element is restricted to containing only the values Male, Female or Unknown.

The problem with the XML shown in Figures 10 and 11 is that it is highly 'coded' and not particularly suitable for people without severely twisted minds to read.

This is where the third of the three key XML structures comes into play. An **XML Transform or Stylesheet** is a document that can be used to transform the XML document into a more readable format.

A sample transform is shown in Figure 12 below.

Figure 12

This is starting to get very 'nerdish' but essentially the code is ensuring that a heading of **Applicant details** will appear (notice the <h2> tag) and then a table with some table headings (notice the <th> tags).

The stylesheet then selects the values of the data and displays them in the body of the table. The code is also able to cope with multiple names if they are present. Notice also that data can be ignored in the transform. In the transform results shown in Figure 14 below, the Given Name is not displayed to show that the transform can select whatever data is required for the business need.

When the XML document is linked to the XML Stylesheet an extra line is added as shown in Figure 13 below.

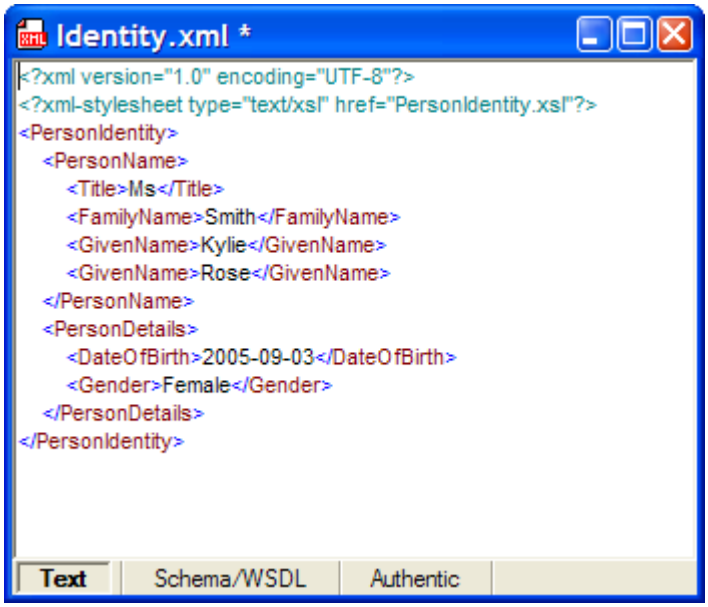


Figure 13

The XML document is now linked to a stylesheet called PersonIdentity.xsl.

When we view this XML document in a browser the output is now transformed by the stylesheet so that it is presented in a table as defined in the stylesheet as shown in Figure 14.

Figure 14

By creating a different stylesheet then different output could be generated.

Some more information on the interactions between a schema and an XML document are provided later in this document.

In summary, the three cornerstones of XML technology are:

- XML documents that provide structure for data
- XML Schemas which validate the structure and content of XML documents
- XML Stylesheets which transform the data in XML documents into other formats.

Validating XML documents using a schema

In the previous section we created an XML document and an XML Schema. Now we are going to use a tool called XMLSpy to show how a schema can be used to validate the structure and content of an XML document.

First, we will load the XML document into XMLSpy and then choose the schema we have already created to link the document to the schema as shown in Figure 15.

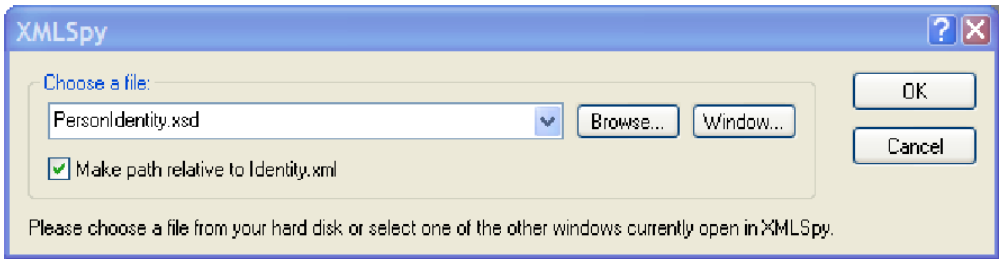


Figure 15

In Figure 16 you can see the link to the schema called PersonalIdentity.xsd.

Figure 16

In XMLSpy you can select an option to ‘Check for well-formedness’ – this checks to see if your XML document obeys the rules of XML.

Figure 17 shows that XMLSpy thinks that the document is well-formed.

Figure 17

You can also check to see if the content of the document is valid according to the schema that it is linked to as shown in Figure 18.

Figure 18

We now know that the XML document obeys the rules set out in the schema. But what if we break some of these rules? Notice in Figure 19 that we have not included a gender in the data.

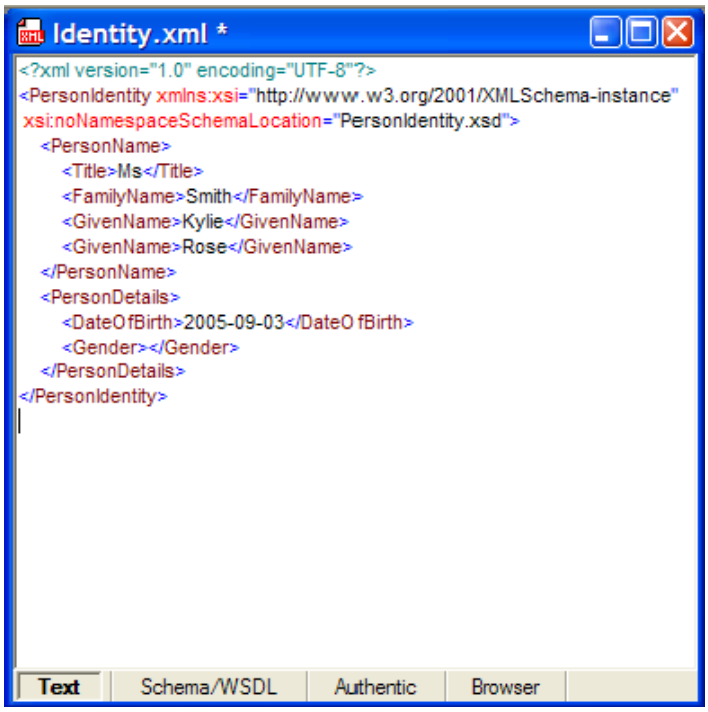


Figure 19

Now when we check for well-formedness it will still be fine however when we check it against the schema the message shown in Figure 20 is displayed.

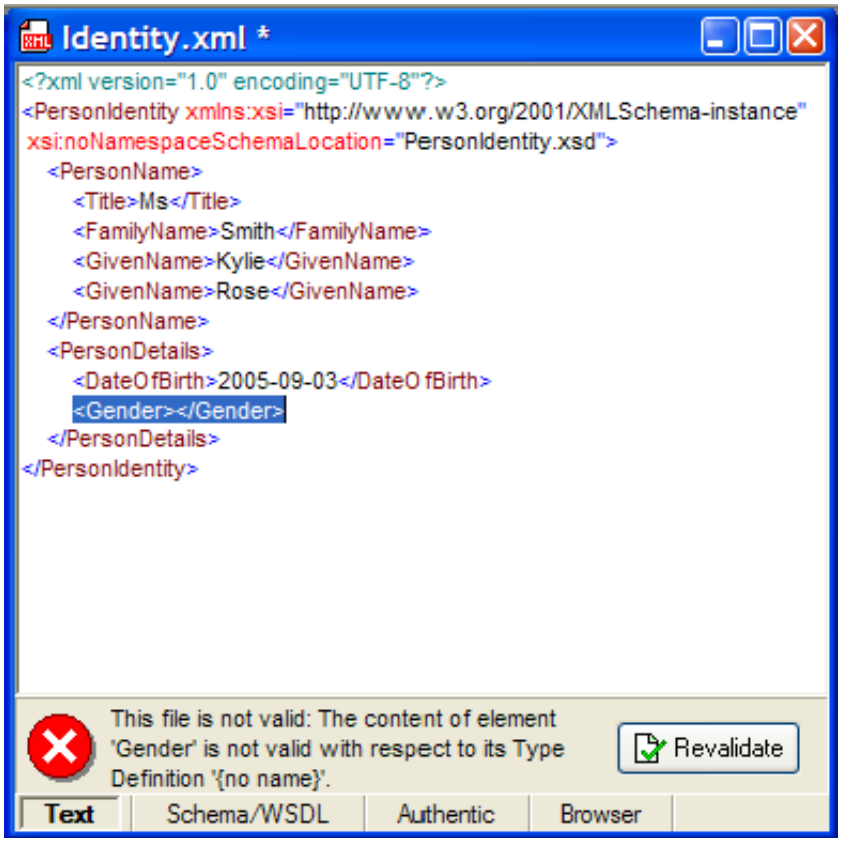


Figure 20

Now, when an XML document is being transferred from one computer system to another, an error such as the one above can be detected by software and an agreed process can be put in place to ensure that the error is corrected through human intervention before another attempt is made.

Appendix A: Monetary Penalties Enforcement Project Overview

The effectiveness of the process for the collection and enforcement of monetary penalties is a long-standing and complex issue.

The Auditor-General has commented on the level of outstanding fines and fees in each of his reports on Government Departments and Public Bodies since 1992-93.

Based on the Auditor General's report, the amount of fines and fees due to the consolidated fund has increased by $13.4 million or 620% from June 1992 to June 2000. The amount due to other parties/agencies has increased by $5.9 million or 300% from June 1995 to June 2000.

In his report the Auditor-General included statistics on the collection rate (the proportion of fines imposed collected within a given period). This showed that the collection rate for court fines reduced from 60% in 1995-96 to a low of 52% in 1998-99. In the same period the collection rate for Infringement Notices reduced from 80% in 1995-96 to 71% in 1998-99.

There are several consequences of a falling collection rate:

- The first is loss of revenue;
- The second is the impact of an ineffective penalty system; and
- The third is the potential cost of attempting to make the existing enforcement system operate effectively.

This project is the third stage of a three-stage process for reforming the collection and enforcement of monetary penalties in Tasmania.

Stage One was the agreement by the Chief Magistrate and Secretary of the Department of Justice on a framework for the collection and enforcement of monetary penalties. The core of this framework was documenting the outline for a new Monetary Penalties Enforcement Bill. The proposed Bill will implement a range of new and more effective mechanisms for the enforcement of unpaid court fines, infringement notices, restitution and compensation orders, and orders for court costs.

Stage Two was the development of a comprehensive Business Case for the implementation for the proposed procedures (MPEP2). The Business Case presented to Government costed options for the reform of the infringement and fine process. The Business Case was successful in its bid for funding.

Stage Three, described within the MPEP3 Project Brief, is the phased implementation of the new processes, new system support and the enactment of the legislative framework for enforcement of monetary penalties and the establishment of the Monetary Penalties Enforcement Service (MPES)

The project is being overseen by a Steering Committee, which is a Sub-Committee of the Inter Agency Steering Committee. The Project Manager (who reports to the

Steering Committee) is Tracey Rodgers of DoJ and the Business Owner is Seth Hills also of DoJ.

The implementation of the project is run by a project team which consists of the Project Manager, the Business Owner, a representative of the Department of Police and Public Safety, the Project Officer and the Business and Technical Consultants.