# Principles of XML Development
# for Justice and Public Safety

Draft 0.3

August 28, 2001

**Table of Contents**

# Executive Summary

While this document addresses the principles of XML development for the justice and public safety communities, it sets the stage for confronting similarly compelling issues associated with interoperability and information sharing among local, state, and Federal entities. What had initially appeared as a situation with differences that were totally insurmountable became workable as the practitioners' motivations for their actions became apparent. Cooperation, and then innovation, were used to progress work that generated mutually beneficial results. This successful methodology can be applied elsewhere.

As detailed investigative work centered on the analysis of different XML specifications (or more precisely, implementations of the World Wide Web Consortium's XML Standard), it was obvious that three excellent efforts had been undertaken independently. It was also clear that each practitioner group aimed at satisfying its own mission and functional (operational) requirements. Naturally, therefore, the XML implementations specified for the Rap Sheet, the Regional Information Sharing System (RISS), and the Electronic Court Filing Standard would not be compatible. For example, tagging conventions were different, as well as the data elements themselves (and how they were used).

Principles and procedures were established, and agreed to, by all participants in the "XML reconciliation" process. It was recognized from the outset that the needs of the individual practitioner groups were of the utmost importance during technical discussions. Therefore, one practitioner group will have strong feelings about a particular technical approach in one functional area, but not have a preference for the manner in which other functions are performed via XML. That observation allowed technical compromise, and gave the participants the ability to dissect the whole effort and operate on it in "pieces".

One unforeseen by-product of using principles and procedures for dealing with XML development was the emergence of "best ideas" during technical discussions. With three sets of developers looking at information sharing from a new perspective, there were suggestions to choose approaches that were developed by others as the common approach. Essentially, the best ideas were picked from the three implementation specifications, regardless of who developed them. This concept alone would be attractive to any other information sharing efforts among different practitioner groups.

# 1 Introduction and Background

A number of efforts are underway to develop XML[1]-based data exchange specifications within the Justice and Public Safety (J/PS) community. The Office of Justice Programs (OJP) recognized the need to create interoperability between the emerging specifications and selected three emerging specifications with which to begin the process. The emerging specifications are the "Interstate Criminal History Transmission Specification" developed by the Joint Task Force on Rap Sheet Standardization, the Regional Information Sharing System (RISS) XML Specification developed by the RISS, and the "Electronic Court Filing Proposed Standard" developed by LegalXML. These organizations are attempting to accomplish dissimilar missions with their specifications, which makes this effort somewhat more difficult. However, an

---

[1] XML is the eXtensible Markup Language developed by the World Wide Web Consortium. Information can be found at their website:  http://www.w3c.org

understanding of those missionsnot only enables progress,  it broadens the utility of the proposed standard by making it applicable to multiple taxonomies accomplishing a variety of objectives.

## 1.1   *XML Interstate Criminal History (Rap Sheet) Transmission Specification*

"In 1995, the National Task Force on Increasing the Utility of the Criminal History Record recommended expanded data content, a presentation format (page layout) for the expanded content, and the creation of a transmission format for the interstate sharing of criminal history information.   The National Task Force included representatives from the Federal Bureau of Investigation (FBI), FBI Criminal Justice Information Services Advisory Policy Board (CJIS APB), National Law Enforcement Telecommunications System (NLETS), the National Center for State Courts, SEARCH, the National Consortium for Justice and Statistics. Its members were a diverse array of justice practitioners drawn from the judiciary, prosecution, court administration, local, state, and federal law enforcement, juvenile justice pre-trial services and state criminal records repositories.   In 1996, the Joint Task Force on Rap Sheet Standardization with representation from the FBI CJIS Division, the APB, NLETS, SEARCH and state and local law enforcement agencies, was formed to carry forward the work of the National Task Force by developing a standardized criminal history transmission format.

The Joint Task Force on Rap Sheet Standardization has accomplished three major objectives:

- Developed an XML based standardized criminal history transmission format

- Developed a standard presentation format utilizing the XML transmission format

- Developed a concept of operations which combines criminal histories from multiple sources into a single criminal history

The Interstate Criminal History Transmission Specification provides a method by which an authorized user who requests an interstate criminal history record, regardless of the request method:

- Will always receive the same set of information

- Will always receive a single record for multi-source interstate criminal histories, in which the criminal justice event cycles are presented in date order

- Upon request will receive the record in computer-readable format for use in filling display screens, data entry screens or databases, or for editing or state-specific presentation formats

- Upon request will receive the record in any of a variety of available presentation formats, tailored for a specific printer or mobile digital unit, for example.

- Upon request will receive the record at an approved destination whether or not it is served by an intrastate law enforcement network."[2]

---

[2] Joint Task Force on Rap Sheet Standardization, "Interstate Criminal History Transmission Specification – XML Version 2.0," April 2001.

## 1.2    Regional Information Sharing System XML Specification

The Regional Information Sharing Systems (RISS) Program is composed of six regional centers that share intelligence and coordinate efforts against criminal networks that operate in many locations across jurisdictional lines.  Each RISS center has from 520 to over 1,300 member agencies.  The vast majority of member agencies are at the municipal and county levels, but more than 304 state agencies and 685 federal agencies are also members.  The Drug Enforcement Administration; Federal Bureau of Investigation; Internal Revenue Service; Secret Service; Customs; and the Bureau of Alcohol, Tobacco and Firearms are among the federal agencies participating in the RISS Program.

Typical targets of RISS activities are drug trafficking, violent crime, and gang activity, and organized criminal activities. Each of the centers, however, selects its own target crimes and the range of services provided to member agencies.[3]

## 1.3    Court Filing XML Specification

The Court Filing XML Specification provides a mechanism by which documents can be electronically filed with a court system.  The developers listed the following assumptions and requirements:

"1) The three-tier application model is assumed, including three cooperating applications: the application on the users desk or the Electronic Filing Provider (EFP), i.e., the client; the Electronic Filing Manager (EFM), i.e., the server; and the Case Management System (CMS).

2) [The court filing specification], with a few exceptions, does not specify how any of the three cooperating applications will function.  Rather, it defines the data elements and data tags for use in such applications.

3) Any EFM application shall have the capability to return an electronic acknowledgment to a filer.

4) The DTD [document type definition] is intended to support both the filing of a pleading initiating a new case and filings in existing cases.

5) The DTD is designed to include all data elements needed by any court.  However, any court may limit, as a matter of local policy, the data it will accept in an electronic filing. In particular, a court may refuse to accept:
· more than one filing in a single envelope,
· a pleading that initiates a case,
· payments associated with a filing, and
· hypertext links to document residing elsewhere.

6) This [specification] differentiates between filing and docketing of received pleadings, orders, or notices.

7) There is no state maintained at the server between requests for a particular user connection.

8) There shall exist a method by which the user identity can be verified.

---

[3] Regional Information Sharing System website, http://www.iir.com/riss/

9) Extraneous data is discouraged. That is, only the information required for successful completion of the transactions should be included.

10) Certain XML document elements are assumed to have their content protected from public view. Other protected data shall be marked as such explicitly.

11) Tags have been chosen which provide a neutral name where the definition of what some would consider a more appropriate term could vary. For example, in some courts the "filer" is the attorney, in others the "filer" is the party to the case.

12) The DTD is designed to support the inclusion of multiple filings in a single submission for the same case.

13) There is no intent to require multiple packets per individual task. For example, a user asking the system to provide a list of all cases s/he is associated with in order to choose the case for which a pleading is going to be made would require a single exchange of packets: one packet to make the request, the second to supply the answer."[4]

## 1.4    The Role of OJP and the Global Initiative

The Global initiative is "an effort to advise the Attorney General regarding the concurrent planning and design efforts now underway in the area of justice systems integration, as well as to advise [OJP and] the Attorney General regarding the new and developing technologies that bear upon justice information sharing." In fulfillment of this mission, the Global Advisory Committee created five working groups, including the Infrastructure/Standards Working Group (I/SWG), "which is responsible for 1) conducting an examination of justice information management and infrastructure activities with the goal of determining how each activity relates to Global, and 2) reviewing the major standards-related efforts currently underway and providing recommendations to remedy the breakdowns in interoperability." It is the latter responsibility through which I/SWG provides significant benefits to the Justice community with regards to information sharing by working closely with OJP (and its technical support staff), and is also the reason OJP is interested in developing as much commonality as practical among the various information sharing efforts throughout the Justice and Public Safety (J/PS) communities.

In meeting their responsibilities, the I/SWG has assembled a group that provides an outstandingly broad representation of users in the Justice community. This representation virtually guarantees that recommendations generated by the I/SWG will address the widest possible range of the justice community's requirements. On the other hand, the OJP, in fulfilling its responsibility of providing a suite of interoperability standards for J/PS information sharing, has assembled a team of technical experts knowledgeable in adapting user requirements into technical specifications and standards. It is envisioned that, with support from the technical experts, the I/SWG, and its member organizations, can recommend a suite of information sharing interoperability standards to the Global Advisory Committee. In turn, Global can provide recommended standards to the OJP and the Attorney General for adoption as "the standard solution" for the information sharing interoperability issues confronting the justice community. Such a standard might be, for example, a data dictionary for use by the whole of the J/PS community.

---

[4] LegalXML JTC/Court Filing Working Group, "Electronic Court Filing Proposed Standard," January 4, 2001.

# 2 Principles for Development of XML Specifications for the Justice and Public Safety Community

At the OJP XML Technology Working Group meeting held June 7, 2001 in Denver, Colorado, developers of the Rap Sheet Transmission Specification, the RISS Specification, and the Court Filing specification agreed to work under a set of guiding principles during their efforts to achieve effective information sharing between their systems. These principles, listed below, begin very broad and general, and then proceed to more specific guidance.

➢ Any XML specification developed should be guided by the principles put forth by the World Wide Web Consortium (w3c)

➢ Internal system representation is not constrained by these guiding principles or the associated data element definitions. The information contained in these documents simply provides a baseline for exchange of information.

➢ XML Specifications should be over-inclusive by specifying those elements that may be required by fewer than all participants and making those elements optional.

➢ XML Specifications should be extensible.

➢ Wherever possible previously developed solutions should be adopted or extended.

➢ It is essential to be considerate of the operational requirements of the participants, as a failure to meet those requirements will result in an unaccepted specification.

➢ International implications of XML specifications should be considered, and international standards should be used as guides where possible.

➢ XML specifications should be broad enough to accommodate jurisdictional differences.

➢ When a participant has expertise in a particular area, that expertise should be utilized and the expert's requirements should take precedence.

➢ When operational requirements dictate differences in specificity, provide a mapping from the more specific elements to the less specific elements.

➢ When no participant has a substantive interest in a facet of a specification, the simplest solution (from a technical standpoint) will be chosen. If no technically superior solution presents itself, a participant will be expected to volunteer a solution to be quickly accepted by consensus.

➢ It is the responsibility of each group to insure that all system-specific features are removed prior to transmission to another group.

➢ Certain complex elements are sufficiently independent and driven by group business rules such that they cannot be used by more that one organization. In such cases the shareable simple elements contained within the complex element are defined.

➢ Where necessary, it is acceptable to use a more specific tag name (e.g., arrestDate), provided it includes the data model of its more generic counterpart (e.g., date).

➢ The data element *<number>* shall not be used.

Specific recommendations based on these guidelines are provided in the appendix.

# 3   Procedures for Development of XML Specifications for the Justice and Public Safety Community

This section describes the process and procedures used to achieve success in bringing the three aforementioned specifications closer to interoperability.

1) Identify requirements each participant is attempting to meet and the goals they are trying to accomplish.  Ensure all participants have at least a moderate understanding of each other's needs.

2) Identify similar information being shared by participants, and the differences and similarities between tag names.

3) Identify and resolve non-substantive differences (e.g., tag capitalization, naming conventions)

4) Identify and resolve those substantive differences that can be resolved quickly.  (e.g., tag names for person name elements)

5) Identify those substantive differences that are difficult to resolve.  Where possible, resolve them.  Where resolution is not possible (usually due to differing requirements) ensure that there is no tag-name overlap and document the differences.

6) Develop a plan (with tasks, goals, and objectives) to be accomplished over a defined schedule.

The Appendix contains recommended guidelines for XML specifications.  Agreed upon simple and complex data elements are provided in a separate document.

# Appendix – Recommendations for Development of XML Specifications for the Justice and Public Safety Community

## *A.1 Tag Names*

Tag names will have the following properties:

Tag names will be descriptive of the data they contain and/or represent.

Tag names in defined specifications will be strictly alphanumeric.
The exception to this recommendation is in the case of tag names developed as extensions to a specification (see A.2 below).

The first letter of the tag name will be lower case (e.g., *name* rather than *Name*).

The second and subsequent words of a tag name will be capitalized (e.g., *firstName* rather than *firstname*).

The same tag name will not be used to represent more than one type of data. (e.g., do not use the tag *type* to represent both a type of weapon and a type of vehicle. Use *weaponType* and *vehicleType.*)

Tag names will be fully spelled out (rather than appearing as abbreviations and acronyms) except in the cases indicated below (in A.4).

## *A.2 Inclusivity and Extensibility*

Where practical, the tag name inclusion policy will be quite liberal to enable a specification to meet the broadest user requirements possible. That is, it is generally recommended that if some (or one), but not all, participants need a tag name, the tag name should be included in the specification and designated "optional". This will facilitate resolution of conflicts in the future. See A.5 Additional Requirements.

In spite of a liberal tag-name inclusion policy, there will be cases where the specified tag names do not fulfill a user's requirements. In order to accommodate such cases, all specifications will provide explicit instructions on extending the specification. While many of the details of this are left to the specification developer, it is recommended that tag names for extensions be of the following format: tag name, underscore, extender identifier (e.g., *neededData_LegalXML*).

## *A.3 Data Format*

It is recommended that the data format be UTC-8, unless an application requires UTC-16.

## *A.4 List of Accepted Abbreviations and Acronyms for Use in Tag Names*

Because there are some instances where an entity is more recognized by its abbreviation or acronym than if the name were spelled out (e.g., IBM vs. International Business Machines), there are exemptions to the "no abbreviations or acronyms" rule stated above. These accepted abbreviations and acronyms will be collected and defined here so that there is one source for those needing clarification.

Abbreviations and acronyms should be all lower case if the tag name begins with the abbreviation or acronym and all uppercase if it is the second or subsequent word (e.g., *oriNumber* and *personID*).

The following table comprises the list of accepted abbreviations and acronyms for use in tag names.

| Abbreviation/ Acronym | Definition |
|---|---|
| BCC | Blind Carbon Copy a message to this recipient |
| CC | Carbon Copy a message to this recipient |
| CFR | Code of Federal Regulations (Federal laws).  Sometimes used to indicate compliance with a specific regulation (e.g., *cfr28* might be used to indicate compliance or non-compliance with CFR part 28). |
| DNA | (deoxyribonucleic acid) Genetic fingerprint information |
| FBI | Federal Bureau of Investigation |
| HREF | Hypertext Reference – this attribute specifies the location of a Web resource, thus defining a link between the current element (the source anchor) and the destination anchor defined by this attribute |
| ID | identification/identifier |
| MIME | Multipurpose Internet Mail Extension – a specification for formatting non-ASCII messages so that they can be sent over the Internet |
| NCIC | The FBI's National Crime Information Center |
| ORI | The Originating Agency code assigned to a law enforcement agency.  This is assigned by a system operator (e.g., database or network operator) and is used to track usage of that system. |
| URI | Universal Resource Identifier (address, similar to a web address) where a described resource can be located. |
| VIN | Vehicle Identification Number (applied by manufacturer) |

### A.5 Additional Requirements

When a developer has requirements for additional elements to be included in a complex data element, those additional elements may be added, e.g., Court Filing when specifying the name of a person requires the following additional elements, full name, designation, and name alias.