

Global Justice XML Data Model: current challenges and efforts to resolve them [DRAFT]

Newton Howard, Sergey Kanareykin

Center for Advanced Defense Studies (CADS)
Cyber Security Policy & Research Institute (CSPRI)
Dept. of Computer Science, The George Washington University
802 22nd St. NW, Phillips Hall
Washington, DC 20052

{newton,sergey}@c4ads.org

Abstract – The release of the new version of the Global Justice XML Data Model offers new opportunities in justice information sharing. This development has broad applications for the Homeland Security area as it enables easier cooperation of information sources and users, and provides further opportunities for knowledge management in key security areas. A number of issues could potentially hinder the wide adoption of this emerging standard. These issues are related to performance, security, interoperability, and cooperation at national and international levels. This paper discusses these concerns and presents the ongoing and planned efforts by several institutions to address them.

1 Introduction

As part of the US Department of Justice initiative for Global Justice Information Sharing (Global) and through the efforts of the Global Advisory Committee (GAC), the adoption of XML as a technology standard for document-centric exchange of information has been a center of gravity in the world of law enforcement and justice [8]. Deborah Daniels, assistant attorney general in the Office of Justice Programs (OJP), noted that “emerging technologies like XML are a core component of our strategy to give state and local governments new tactics and methods to help them respond to the security challenges of a post September 11 era,” [9]. The GAC, a part of the Office of Justice Programs (OJP) and an adviser of US highest-ranking law enforcement officer - the U.S. Attorney General, has an impact on more than 1.2 million justice professionals. Believing

that the efficient sharing of data among justice entities is at the very heart of modern public safety and law enforcement, Global is a group of more than thirty independent organizations, spanning the spectrum of law enforcement, judicial, correctional, and other related bodies [8].

Efforts by members of Global to find a common exchange format have led to the development of the Global Justice XML Data Model (GJXDM) [1] and its recent public release. The work of the SEARCH consortium in deriving the Justice Information Exchange (JIEM) methodology, the XML Structure Task Force (XSTF), and Georgia Tech Research Institute (GTRI) in conjunction with Integrated Justice Information Systems (IJIS) Institute resulted in a national consensual model that hundreds of agencies are now in the process of adopting and implementing. The recently released version of GJXDM is

being used in more than 50 information sharing efforts – including the national AMBER Alert program already in operation [9].

The release of the Global Justice XML Data Model version 3.0 gives an opportunity to exchange information among law enforcement and justice information systems in ways never before possible. John Wandelt, a senior research scientist with the Georgia Tech Research Institute (GTRI), has noted that “[b]y providing a common language and vocabulary, the XML initiative allows agencies to efficiently share data while continuing to maintain their own data and operate their own computer systems.” [9].

GJXDM is a large XML-based data model and data dictionary, which contains 2754 data components consisting of 545 types and 2209 properties, covering 20,000 data objects – an average number of unique data objects that are used in all branches and departments of a typical state government [10]. These data components are organized into different categories (e.g., activity, person, property, location) with 45% majority of activity-related components. Person-related data components come second with 21% of the total, leaving the third place to property-related components.

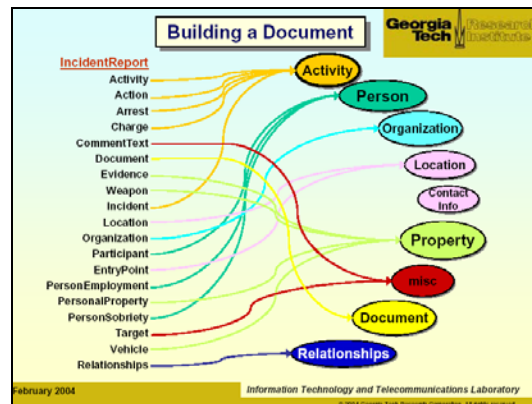


Figure 1 - Document building [10]

The object-oriented mechanisms built into the model allow for re-use and extension of components, which is promising to greatly facilitate the adoption of this technology at different levels and in different jurisdictions (see Fig. 1 that demonstrates how typical documents are constructed). Built from existing data models, dictionaries, processes, and document specifications [10], the requirements-based architecture “avoids the cost and compatibility issues that would be involved in trying to develop a single unified national network. It also provides a foundation that individual agencies can use to develop compatible systems without having to re-invent key elements”, as noted by Wandelt [9].

To the extent that the data dictionary and structure in the GJXDM are accepted throughout the developer and user community, it will become much easier for information to be sent, transformed, and applied in disparate information systems. This model is clearly a breakthrough in justice information exchanges [1]. Obviously, the development of this model has implications that go beyond law enforcement information sharing. Recent efforts to improve information access

and sharing have been greatly motivated by the Homeland Security initiatives and will contribute to the mission of the Department of Homeland Security.

However, the wide adoption of this new standard requires that a number of challenges are resolved and the developers and managers are provided with the information necessary for a successful implementation. This paper outlines several currently existing concerns surrounding the GJXDM, and presents the related efforts undertaken by IJIS Institute, Cyber Security Policy and Research Institute, and Center for Advanced Defense Studies.

2 Application performance concerns

The GJXDM is a very complex model that employs recent XML technologies, including, but not limited to: the W3C XML Schema language, the DoD 5015.2-STD Design Criteria Standard for E-RMS Applications, and the Intelligence Community Metadata Language (ICML). The overhead coming from using XML technologies (as compared with legacy data exchange systems) is well known. The use of the most recent advancements in the XML area, while bringing a wide array of features, also comes at a price of limited tool support because not all vendors have developed efficient XML Schema-compliant applications and tools. Furthermore, the all-inclusive character of the model makes GJXDM a very large XML schema, which has its implications as well. A thorough research effort is required to determine the performance bottlenecks in the processing tools and find ways of optimal use of the GJXDM.

In Spring 2004, the Office of Justice Programs (OJP) in the Department of Justice awarded a grant to the IJIS Institute to conduct performance testing of the initial production release to provide industry and government developers and administrators with insights into the impact of using this important new tool. The IJIS Institute created a partnership with research organizations within The George Washington University to conduct the performance testing under the leadership of a project committee consisting of representatives of IJIS Institute companies and other organizations having specific interest in the test results.

This project, currently in completion phase, implements a variety of data exchange scenarios using GJXDM and takes measurements of the time, memory, and network resources required for processing at different steps of XML data exchange transactions, e.g., validation or transformation. The testing is performed using Web Services on J2EE/Linux and .NET/Windows 2000. While the final test results will be made available towards the end of summer 2004, preliminary test results already identify several important directions for future research.

In particular, the size and complexity of the GJXDM XML Schema files have a very significant effect on the validation time. For some schema variations and subsets, validation length can approach values clearly preventing real-life implementation. The time spent parsing the GJXDM schema dominates validation time, which can be optimized by using different ways of organizing the

schema modules. Therefore, there is a pressing need to investigate and develop further the modular composition of the GJXDM schema.

Furthermore, there are significant differences in the behavior of the XML toolsets used on the two tested platforms – J2EE and .NET. The .NET platform displays significantly slower validation times for the full 3.0 release GJXDM-conforming XML files. Representatives of IJIS Institute and GWU/CADS research team are working together with industry representatives, in particular Microsoft Corp., to analyze the test effort results, address the questions raised, and make recommendations to the practitioner community.

3 Security concerns

While the choice of XML as the means of exchanging information has been made, there is a distinct lack of widely adopted models and standards for applying security to the exchange of XML information. The World Wide Web consortium has yet to finally endorse the standards that will help apply security to the transmission of XML documents. Much more work is needed to identify the measures that must be applied in the law enforcement and justice environment to ensure the security of data exchange. This will require a comprehensive look at the security systems within which the GJXDM will be used, and development of new security requirements on multiple levels.

Traditional network and host security: Since XML data exchanges in most cases are developed as additional services that interface with existing applications and use existing networks,

GJXDM implementers should use the introduction of this new technology into their systems as an call to review their security architectures. One of the commonly cited reasons to do so is the fact that Web Service operate over commonly used HTTP ports and potentially allow data retrieval and remote code execution over channels that are not watched by standard firewalls [14]. The review efforts should result in adopting new solutions and standards for comprehensive XML data security, which involves authentication and point-to-point encryption.

XML and Web Services security standards: XML and Web Services are becoming the primary vehicles of GJXDD development [11]. While the security standards specific to these technologies can be viewed strictly within the context of the above-mentioned comprehensive security policy, they can and should be evaluated on their own because of their specific focus. Evaluation of the leading XML and Web Services security solutions initiatives is essential. In this area, special attention should be given to OASIS Security Assertion Markup Language (SAML) [5] specification. SAML intends to distribute authentication and authorization information across platforms, organizations, and vendors, simplifying single sign-on and access control. As this standard is nearing completion, it is commonly viewed as the most possible candidate for wide adoption, and should be the primary candidate for evaluation.

Schema- and application-level security: This area of XML information exchange security is frequently overlooked in favor of the traditional

security concerns, but in the long run may become a major concern. With the increase in the number and complexity of GJXDD-encoded data exchanges, it will become increasingly harder to ensure consistency of the data and relationships between referenced objects. A ‘rogue host’ within a law enforcement facility is not a complete impossibility, and the potential damage from subtle and hard-to-trace changes in, for example, criminal records is immense. Development and testing strict constraint schemas for GJXDM validation is one of possible ways to counter this threat. At the very least, the implementers will need clear guidelines on checking ‘real-life’ GJXDM exchange data. The challenging task of developing validation guidelines and applications should be given a high priority.

4 Interoperability concerns

The adoption and successful use of the GJXDM technology is contingent upon the possibility of future integration with other formats and ease of use by a wide variety of organizations.

On the level of the technologies used for storage and transmission of data, GJXDM has a wide support base (with XML and Web Services being increasingly common). However, its effective use will depend on the level of interoperability that can be achieved at the level of the data model – e.g., how flexible and easy it is to extend. Several case studies, especially ones going beyond existing applications of Justice XML, will be required to answer these questions and provide the guidance for adoption.

One of such case studies currently proposed by the Center for Advanced

Defense Studies and the George Washington University involves using GJXDM schema to create a format for exchange of car crash data. This project would use the data generated at the National Crash Analysis Center, located at The George Washington University Northern Virginia campus [6].

5 Cooperation at the national and international level

The primary goal of the GJXDM to date has been to facilitate the justice information exchange among the jurisdictions within the USA. However, with increased attention to the terrorist activity worldwide and a growing need for international cooperation in this area, it is evident that the future implementations of the GJXDM will require adoption at a wider variety of government agencies, with a higher level of internationalization, and guaranteed compatibility with similar efforts developed elsewhere in the world. Special attention should be paid to supporting operational tasks, such as data extraction and GJXDM object relationship tracking.

This direction of GJXDM research involves both technical and political tasks. The technical tasks would be to investigate how well the current version of the GJXDM schema and tools support languages other than English, and what is required to ensure this support in the future. Exploring the existing similar developments in other countries and developing an integration case study would support this cause and provide valuable data to the current maintainers of GJXDM and future implementers. The Center for Advanced Defense Studies is currently evaluating the possibility of carrying out such a case

study in collaboration with the Law School of the Jean Moulin University, Lyon 3 (France).

The policy aspects of using the GJXDM schemas should also stay in focus. Interoperation between justice systems in different jurisdictions, even within the same country, is a complicated procedure, which needs to conform to a variety of policies, such as data storage and privacy regulations. The Cyber Security Policy and Research Institute has a long track record of staying on top of government IT developments, and will provide wide expert support in the implementation policy issues.

6 Conclusion

GJXDM is a rapidly developing data exchange format for justice information sharing, which represents multiple benefits to law enforcement and Homeland Security. To ensure its wide adoption and effective use in the future, a number of research and development efforts need to be carried out. In particular, they include performance testing and optimization, security framework development, and interoperability case studies. The Center for Advanced Defense Studies and its partners propose feasible solutions to meet this goal.

References

- [1] GJXDD Performance Testing Project Test Plan, Draft version, June 22, 2004
- [2] SEARCH, The National Consortium for Justice Information and Statistics, <http://www.search.org/default.asp>
- [3] Global Justice XML Data Model, U.S. Department of Justice, Office of Justice Programs <http://www.search.org/default.asp>
- [4] DOSSIER: Documents Security System for Internet Electronic Resources, project proposal, IJIS Institute, Spring 2004
- [5] SAML, OASIS Security Services Technical Committee http://www.oasis-open.org/committees/tc_home.php?wg_abrev=security
- [6] National Crash Analysis Center, <http://www.ncac.gwu.edu/>
- [7] DOT starts building crash center, September 10, 2003, FCW Media Group, <http://www.fcw.com/fcw/articles/2003/09/08/web-dot-09-10-03.asp>
- [8] Global Justice Information Sharing Initiative, <http://it.ojp.gov/>
- [9] Georgia Tech Helps Provide Foundation for New Justice Information Sharing Initiative, Science News, <http://www.newswise.com/p/articles/view/505703/>
- [10] Mark Kindl, Global Justice XML Data Model (GJXDM): A Technical Walk-Through, Georgia Tech Research Institute (GTRI), 2004
- [11] Structure and Design Issues for Developing, Implementing, and Maintaining a Justice XML Data Dictionary, report of the Justice XML Structure Task Force (XSTF), 13 April 2004, http://justicexml.gtri.gatech.edu/JusticeXMLStructureTaskForceReport_a.pdf
- [12] Interoperability Strategy Concepts, Challenges, and Recommendations, *Industry Advisory Council (IAC) Enterprise Architecture SIG*, <http://it.ojp.gov/documents/Interoperability.pdf>
- [13] Technology considerations in the development of integrated justice data exchange standards, *IJIS Institute*, <http://it.ojp.gov/technology/files/IJIS-Standards.pdf>
- [14] Investigating Injection of Malicious Code into Arbitrary Executable Code, Newton Howard, Sergey Kanareykin, technical report to ARL, 2003
- [15] Integrated Justice: The Foundation of Homeland Security, Karen K. Diamond, Unisys Corp., 2004, http://www.unisys.com/public_sector/insights/articles/articles.htm?insightsID=21503