

GOVERNMENT TECHNOLOGY**Government Technology****XML: Out of the Shadows**

By Jim McKay
May 26, 2005

The federal departments of Homeland Security and Justice recently agreed on a global data-sharing standard that could spur interoperability throughout the public safety community and beyond.

The move limits proliferation of incompatible XML data models -- which translate data into information that can be shared among multiple IT systems -- and opens the door to greater cooperation among law enforcement, firefighters, health organizations and others. The DHS and DOJ also tied use of the new standard -- known as the Global Justice XML Data Model (GJXDM) -- to federal grants for information exchange projects in fiscal 2005, so states and localities are challenged with getting up to speed on the promising data-sharing model.

Some have already. The DOJ's Office of Justice Programs (OJP) has tallied at least 50 fully implemented extensible markup language (XML) applications in the government justice and public safety realm thus far, including Amber Alert, possibly the most noteworthy example. That number could stretch to 200 or more when newer, developing implementations are considered.

That's exciting to those who see XML as a silver bullet for interoperability among disparate communications systems. The XML thrust received a big boost in February with the announcement that the DHS and the DOJ will use the GJXDM for interoperability.

A source from the OJP said the agency is "quite happy" about XML's progress, while acknowledging it's a long-term project that will take years to implement.

"The best news about this model is that there's no secret to duplicating its success," said Bureau of Justice Assistance (BJA) Director Domingo Herraiz. "We're receiving reports from numerous states on improved information sharing and the cost of efficiency of implementation." Herraiz said state and local governments should be excited about GJXDM's future. "The national cost savings alone would be in the billions, and the speed by which information could be shared is phenomenal."

Over the past three years, the BJA, working with the National Governors Association (NGA), allocated \$17 million to support state and local implementation efforts. Other funding sources that have contributed to state and local GJXDM development are the Edward Byrne Formula Grant, the Local Law Enforcement Block Grant program, the Justice Assistance Grant, the National Criminal History Improvement Program, the Office of Community Oriented Policing Services and the DHS.

Catherine Plummer, a justice information specialist at the National Consortium for Justice Information and Statistics (SEARCH), said adoption of the GJXDM by federal agencies and the requirement of the GJXDM conformance for receipt of grants will have a big impact on state and local governments.

"Whether states and locals have any clue what we're talking about is a different issue. But they are slowly becoming aware that this is a condition of

SIDEBAR**XML Makes a Mark in Maricopa**

In Maricopa County, Ariz., police don't have to spend all day entering data on arrestees into multiple systems, thanks to the county's commitment to the Global Justice XML Data Model (GJXDM).

In one of its several information exchange programs, the county deployed a booking and arrest system that automatically populates multiple systems with arrest data. Instead of having to be manually entered several times, data is entered once and automatically populates several systems, including the police agency system, the Arizona Department of Safety criminal records repository and the correctional system.

their grants if they're going to use XML," Plummer said. "If they are going to do any kind of information exchange development, those changes have to be GJXDM-conformant."

The new GJXDM standard is just what the doctor ordered in terms of the guidance states and locals have needed to move forward on interoperability.

"From a practitioner standpoint and from NLETS [National Law Enforcement Telecommunication System] as an organization, it's been a key to our success in promoting information sharing because there is this standard now that we can look to, use and implant," said Bonnie Locke, NLETS administration director.

Locke credited the OJP with bringing practitioners from state and local governments, and the private sector, together for meetings to work on the model. She described the meetings as ugly, long and drawn out. "I sat in on one of the early meetings in Atlanta a couple of years ago," she said. "Literally they sat in a room arguing about what would go into the model, what something would be called. It was amazing. They'd be arguing about [how to describe] eye and hair color."

Plummer remembers those meetings as well.

"The unique part of the development is that user requirements were built by a bunch of users looking at our real specifications, looking at real exchange data," Plummer said. "This is actually something built by state and local participants. That's amazing. The federal government has never let state and local define what they do."

She said research from SEARCH and NLETS, and participation from several states, helped yield the data that resulted in the GJXDM.

The partnership between the DOJ and DHS on the GJXDM could smooth the overlapping of XML from justice and law enforcement into the intelligence world and other agencies -- such as transportation, emergency services, and health and human services -- without fragmentation.

The Glue to Interoperability

XML is a standard or means of translation between communications systems. For example, the GJXDM version 3.0, recently released by the OJP and the Global Justice Information Sharing Initiative, consists of a data dictionary, a data model and a component reuse repository or database.

XML has been called the "glue" that could pave the way for future interoperability among agencies and jurisdictions. It's based on a common language that acts as a catalyst in defining common terms between disparate systems.

The data dictionary is the data model's underpinning structure. It is, in effect, a spreadsheet containing identification of data elements, and the meanings or definitions of those data elements, all of which are unique. The data model builds relationships between the data elements, and the result, in simple terms, is that disparate systems connect via the unique identifiers.

"Law enforcement agencies don't want to stand in line at our jail and do data entry," said John Doktor, technical director of the Integrated Criminal Justice Information System agency for Maricopa County. That's literally what they had to do in the past. Now the police can enter arrest data into a state booking system, and it updates the others immediately.

"By entering into one system, the data is XML-tagged, comes over to the county and goes into the pre-booking system," Doktor said. "When the officer shows up with the person at the jail, they're already registered, if you will, in our hotel."

Maricopa County staff got a head start on implementing these information exchange packages by attending National Consortium for Justice Information and Statistics (SEARCH) training sessions, according to Catherine Plummer, justice information specialist for SEARCH. "They learned so much in that process that they have completely automated their arrest and booking

A problem arises when the data is defined or represented differently, perhaps by different organizations and without a standard model, according to John Wandelt, research scientist for the Georgia Tech Research Institute and a key contributor to the reuse repository's development.

Absent a common model, the development of too many XML elements and schemas is possible, which could lead to fragmentation.

"The courts have their community, their standards," Wandelt said. "And there are different constituencies in law enforcement. Initially they had their own efforts, so then you have XML that doesn't interoperate with XML."

Much of the data shared between justice and public safety agencies, or even justice and transportation agencies, is similar, so each organization needed to define and represent those elements in the same manner for interoperability to take place, otherwise fragmentation can occur.

"The justice domain communicates with other domains -- homeland security, health, transportation, INS [U.S. Citizenship and Immigration Services], finance," Wandelt said. "Very often when those transactions go across those domains that overlap, that semantic overlap needs to be represented and semantically consistent between those domains. We can't have fragmentation, where all these individual groups define the same things but represent them in different ways."

A few years ago, things looked that way in the XML world, when multiple groups began developing XML-based exchange documents.

"There were dozens, if not hundreds, of different law enforcement specifications with different groups that were working those," Wandelt said. "The thought that came to mind is, 'Why are we building all these documents?' And when I say document, I mean an XML electronic transaction."

OJP officials had the same thoughts.

"We found there was XML work going on in the intelligence community," said an OJP source. "There was XML work going on in the state and local courts and criminal history area we commonly call the 'rap sheet world.'"

Brave New World

The idea was to bring all those "worlds" together. A task force acting on behalf of the Global Justice Information Network Advisory Committee and the OJP decided to include justice and public safety terminology -- not just law enforcement terminology -- in the data dictionary.

The task force combed the country for justice applications that ease interoperability among agencies and jurisdictions, collecting about 16,000 data elements used in approximately 35 of those implementations. Redundancies were removed and semantic differences were resolved. The groups eventually reduced the number of elements to less than 3,000.

Collecting those data elements was laborious, but developing common definitions for data exchanges that are reusable across multiple agency lines was even trickier.

segments with the county with a new jail system."

SIDEBAR

XML Links Alaska Justice Systems

In Alaska, an XML-based exchange engine paved the way for the state to more efficiently process a backlog of traffic citations. The Alaska Administrative Office of the Court found itself with a backlog of 17,000 traffic citations that needed to be manually entered in three different locations: the Anchorage Police Department, the court and the state Bureau of Justice criminal records repository.

Obviously the backlog meant the state wouldn't collect on the estimated \$1 million in fines until the citations were entered into the system. The backlog made it more difficult to collect in the long run because inevitably people changed locations during the lull. It also made moot the tax intercept program, where the state, in effect, keeps a portion of a tax refund if the individual has unpaid fines.

So the state implemented an XAware XML-based exchange engine to pass the citation from

"The most tedious part of building this [data model] was coming up with really good, nonambiguous definitions," Plummer said. "It's not perfect, but it's a good start, and it's based on real data, real research on exchange. It's not people sitting around a room making things up."

She said most agencies use core data, and only certain agencies use data outside that core. For example, department of transportation databases may include 50 data elements to describe snow.

"Most of us don't need that, but there are transportation databases that actually record that much detail about snow," Plummer said. "The nice thing is if we can structure them the same way in the core, they're all reusable. So if I come in and I have a justice exchange, and I need 20 different descriptions of snow, I can actually go into transportation and reuse those components because they're structured consistently. That's the goal."

Grants Force the Issue

With federal grant money now tied to GJXDM implementation, state and local agencies are being forced to learn what XML has to offer, Plummer said.

In February, the NGA Center for Best Practices announced grants to six states for implementing the GJXDM. The NGA -- with funding from the DOJ, OJP and BJA -- awarded \$50,000 each to Colorado, Kansas, Kentucky, Nebraska, Pennsylvania and Wisconsin for 12-month pilots.

The pilots will encourage use of XML in sharing justice information among existing systems. Wisconsin, for example will develop its "justice gateway" to expedite information transfer between law enforcement, the courts and corrections quickly and without moving a lot of paper. Kentucky plans to use the GJXDM to electronically transmit data collected from the Automated Fingerprint Identification System to the prosecutor's case management system when a suspect is booked.

"We've been talking about this for years, but now people are faced with the fact that they have to use it," Plummer said. "I think most practitioners buy into the fact that they want to share information, they want information to be timely, they want to track information on offenders, they all agree on the business case of information sharing."

She said most at the state and local levels understand that XML is a kind of translator but are confused about the details. "How do I do this?" and "Do I have to change my system?" are commonly asked questions, Plummer said. "Well, no you don't. That's the point. It's really not quite the obstacle most people think it is."

At press time, Plummer was developing a users' guide to XML and GJXDM conformance that will be available through SEARCH.

"The point is that exchanges among these systems are what need to be conformant," she said. "You don't need to change your system. You can keep your legacy system. But what you pump out of that system needs to be transformed into something conformant."

That means data "pumped out" of one system needs to be tagged or annotated so it can be manipulated and interpreted by another, different system.

That's all the user needs to know about XML, Wandelt said.

"The end-user, the person on the street, he doesn't see XML. He just sees a Web interface or whatever

one legacy system to another, eliminating the need for multiple entries of the same information.

At press time, the 17,000-citation backlog was nearly history. The 12,000 or so citations that had been updated had saved about 400 man-hours because of the ability to enter the data once and update three different systems at the same time.

interface he has. He just knows he can get to more things than before," he said. "It's like when you look at your Web browser and see HTML. You don't need to read that, you just look at the Web page. It's the same with XML."

Still More Obstacles

Plummer remembers working on XML schemas while an employee of New Mexico.

"We made up our own XML because there wasn't a standard," she recalled. "Now with the data model, we have something much more consistent."

Conformance to the GJXDM will enhance XML-based information sharing, but there are still fears the model could grow too large. "With the Department of Homeland Security on board, there is now a concept of a governmentwide information model" Plummer said. "The latest term I've heard is the National Information Exchange Model (NIEM)."

Plummer said the larger model would be based on the GJXDM but encompass many more agencies. The NIEM might include domains such as the EPA and the departments of Education and Labor. The worry from Plummer's point of view is one of resources. She said there are not enough people available who understand how to develop the XML structure, and expanding too soon could stretch resources too thin.

"You have a core set of people who have been building this and are stretched to the max. It's kind of hard to think of going into all these other areas."

An OJP representative also expressed concern that the model could become too large and cumbersome.

"Where it gets complicated is the data model because when you're trying to implement it in schemas as a single model, it's too big," the source said, adding that Georgia Tech is developing a "subschema generator," a technical tool to help break down the model into subschemas instead of one large model. It would select certain subsets or hierarchical relationships, using only the necessary portions of the model.

Privacy is another obstacle to sharing information, Plummer said.

"There are different rules and regulations on individuals' privacy in every state," she said, describing a hypothetical situation in which Florida puts information on the Internet and sells it to consumers, but New Mexico doesn't. "If New Mexico shares something through an exchange, and [that information] somehow gets consumed in Florida and is on the Internet, that could violate New Mexico's laws. That's an issue."

She said technology can secure the data but the policy-level decisions about the data must be carefully thought out. SEARCH offers a free tool, the Justice Information Exchange Model, designed to help policy-makers record their decisions about ownership and access to key data.

"It provides a place for policy-makers and operational people to really go through their exchanges and create a road map for what they're going to share and the business rules for that exchange."

Plummer said limited resources for training and lack of education are obstacles that must be overcome. She said a good education and outreach strategy must be deployed to inform state and local government officials.

"We need to come up with really good ways to get policy-makers onboard," she said, adding that though it's a big help that grant money is being tied to GJXDM conformance, it's not enough. "Education and outreach are our challenges. That's where we really need help to get the word out and get people onboard to believe they need to use the GJXDM and to look for the help to use it appropriately."

[Jim McKay](#)

