SUSPICIOUS ACTIVITY REPORTING

— From a Portfolio of NIEM Success Stories —

The great Grand Coulee Dam, eastern Washington State. Massive, spectacular, iconic, isolated—vital. And in the summer, swarmed with tourists.

Families crowd their RVs into choice spots opposite the Dam's towering concrete curves, waiting for the laser projections that cover the Dam's wall nightly. Tourists from eastern nations with strong engineering cultures take tours down the elevators into the guts of the Dam. Cyclists cruise the broad concrete walkway along the Dam's summit. Everyone seems to have binoculars, a camera or a camcorder, and everyone seems to be looking, snapping or filming non-stop.

Because it is a dam and hydroelectric, many ask technical questions. About the scale, the concrete, the flow, the power, how it was built, how it is maintained, what aspect of this endless structure does what? These are the usual innocuous avenues of inquiry.

In the midst of this flow of jovial tourists, always observant, are the local police, Dam security, and Federal agents. For them there is one primary, ongoing question: how do we keep this superstructure safe from vandals, criminals and terrorists? Further: how to distinguish between the tourist seeking souvenir images, and the terrorist engaged in pre-operational planning—with an eye to destroying the Dam and all that stands in its path below?

In our era of heightened awareness—when "If You See Something, Say Something" is the by-word—every cop on the beat is a sensor with eyes and ears alert. As is every citizen.

When something doesn't seem quite right, cops are expected to make a note of it—even if the observed conduct is not a crime. Perhaps a tourist tells a police officer that someone with a 'strange accent' is asking "too many" questions of a tour guide...what can that mean—and in an America rich in regional dialects and immigrants old and new, what is a "strange accent," anyway?

In the past, officers may have noted the mention and filed it away or shared it with the next duty tour using stickies, notepads or the backs of envelopes. Recognizing the potential value of many of these observations, many departments have in recent years formalized the procedure, adopting a "Suspicious Activity Report," or SAR, as protocol.

Today, then, officers receiving information or observing it themselves may make formal note—and fill out a SAR. While not all departments have adopted the formality of a SAR, many have. But how best to make use of this data—captured, as it is by many departments, in non-standard formats, with differing definitions, and maintained by independent agencies with neither the will, authority, finances nor process to share and collaborate?

What if, for example, the next week, a SAR is lodged about a similar person asking similar questions nine hundred miles away at Hoover Dam outside Las Vegas? How will these agencies, or even members of the same agencies at disconnected places, *connect the dots*? What will enable them to see a pattern in seemingly unrelated events—if there is one? Events that should raise not just eyebrows, but serious concerns, and trigger effective follow-ups?

It would be easy for officials to secure the Grand Coulee Dam, or for that matter any other infrastructure, from such pre-operational terror exploits: simply close it to the public and secure its perimeter.

But in an open society, even in an age of terror, officials charged with the Dam's security must maintain safety even while they assure that Americans and visitors from other countries are free to enjoy the benefits of leisure time visits. Winston Churchill once famously rebuffed a senior aide's recommendation to close London's museums and theaters during the Blitz. "Dammit man," the PM said. "We're fighting to keep them open!"

An open society also guarantees liberties—meaning, for example, that citizens should ordinarily be free to take photographs of dams without fear of interrogation by police officers; they should ordinarily be free to ask questions of tour guides without becoming the subject of law enforcement reports identifying them as potential terrorists. And their names should not ordinarily reside in law enforcement databases simply because they visited the Grand Coulee Dam one week and the Hoover Dam the next.

And yet...somewhere amongst the innocent tide of visitors has been, and one day will likely again be, men and women, and perhaps boys and girls, who are engaged not in innocent past-times but in pre-operational planning for terror strikes. Men, for example, training to pilot planes who show no interest in learning how to land them...

Until very recently those involved in pre-operational planning for terror in the United States had little to worry about. Police departments **defined** suspicious activity differently. They **recorded** suspicious activity differently, if at all. State, local and Federal systems were not built to interoperate and could not easily exchange data. Disparate laws prevented many State and local agencies from sharing information with Federal enforcement organizations. *What would become of it? Where would it be stored? Who could access, see and use it?*

Enter NIEM

Separate and apart from matters of terror and terrorism, even before 9/11, a collaboration of State, local and Federal law enforcement officials had made progress in establishing new capabilities for the sharing of information about crimes, court cases, and related matters. These capabilities rested on agreements for the "naming" of like things called by different names in their computer systems; the process for arriving at such naming agreements; and governance of the relationship between parties entering into these agreements.

Because over the years many and disparate computer systems had sprung up on the American law enforcement landscape, all with their own names for common things, a lack of *interoperability* among justice-related systems at the State and local levels was common. But such technical obstacles to information sharing created risk, inefficiency, and affected performance—often with dire consequences.

Where judicial, welfare, and health agencies all might have information about a child at risk of abuse, for example, each data system could use different naming conventions to refer to the child. A "youth" in one system was a "minor" in another and a "juvenile" somewhere else—even though they all referred to identical things in the real world. But so long as there was no way to translate one to the other, it would be impossible to exchange data meaningfully among them—or in time.

As a result, dots that should have been connected—dots which might point to real risk for a child, for example, or even perhaps less risk than thought—went unconnected. Authorities would sometimes discover *too late* to prevent harm, and sometimes, it might be said, even moved *too early* in such matters, breaking apart families.

However, with the advent of extensible mark-up languages—XML and its many subject matter-specific flavors—much changed.

Using XML-based metadata (data about data), State and local justice agencies and their Federal counterparts who wished to exchange information—where lawful and appropriate—could keep their own "legacy" system names for things, and agree instead to a *metadata* dictionary.

With the metadata agreed to in an *information exchange model*, everyone could "speak" their own language, leaving their huge legacy systems unchanged except for the tagging of information but send and understand messages to and from others. The XML-based exchange model enabled all to translate and share data between systems quickly.

"I call them 'automobiles.' You call them 'passenger vehicles.' Let's both agree to tag those things 'cars,'" for example—meaning two systems could exchange data about the same "automobile/passenger vehicle" provided they used the agreed-upon tag, "cars."

It was elementary but an important breakthrough, whether for efficiency, transparency, or improved performance. Analysts could run reports, for example. Statisticians could find patterns, and policy makers could better understand results, trends and options.

The data could get connected.

From a systems and budgetary perspective there was real benefit. If law and policy permitted, organizations could exchange data without having to rename everything in their databases to conform. That lowered costs and reduced obstacles to information-sharing significantly. New agencies could join the network easily and improve the total value of the network to all. Once "cars" was agreed to, for example, *anyone* who wanted to exchange information about "cars" could *reuse* the same tag. And system updates and changes would only mean adding or adjusting the metadata, not rewriting entire legacy code.

Global Justice XML, as it became known, emerged as a "win-win" for everyone, *transforming* the value of the information assets in disparate systems, which until now were isolated and of limited value, into a fused "common operating picture." And much was learned about the *process* of getting to those crucial agreements—lessons about governance, rule-making, and the step-wise method—which assured consistency in approach and results.

In the same way, a *national* information exchange model, based on the same principles of step-wise development, and utilizing XML, should make it possible for *any* system owner to exchange information with *any* other system owner—whether law enforcement, health, energy, transportation—provided they each made their systems conformant with a shared metadata dictionary.

NIEM's roots run deep to its sources not just within Global Justice XML at the State and local level, but across the Federal Government, and to the national level. Over the past decade these three strands have come together to establish NIEM as a significant new national resource for information sharing.

At the national level, a keen new awareness of vulnerability and response to 9/11 led to the creation of the Department of Homeland Security, passage of the Intelligence Reform and Terrorism Prevention Act of 2004, and establishment of the Program Manager for the Information Sharing Environment. It culminated in the decision by the Departments of Justice and Homeland Security, in 2005, to adapt the Global Justice XML body of work to a new national enterprise, the National Information Exchange Model, or NIEM.

A related initiative focused on streamlining information gathering and sharing across the Federal Government. It started with the Clinger-Cohen Act of 1996, continued with the E-Government Act of 2002, the establishment of the Federal Enterprise Architecture within OMB, and OMB's publication, in 2005 of the Data Reference Model. NIEM is today the leading implementation of that reference model.

Information Sharing in the Age of Terror

There is no single source for information related to terrorism. Awareness is gained by gathering, fusing, analyzing, and evaluating relevant information from a broad array of sources on a continual basis.

In an age of asymmetric warfare and terror, ordinary crime, industrial espionage, and commonplace financial transactions can all be vectors of support, planning and operations for terrorist strikes.

As a result, important data and information may be observed by cops on the beat, housing inspectors, bank tellers, fire marshals, or shipping companies—as well as gathered through the formal agencies of the law enforcement and intelligence communities.

Until the opening of fusion centers, that information often remained isolated in systems and organizations that could not, or would not share information. Fusion centers are an analytic resource that support the efforts of State and local law enforcement to prevent and investigate crime and terrorism in local communities. Fusion centers receive information from a variety of sources, including Federal, State, and local entities, and ensure timely and relevant information is provided to the right stakeholders within their geographic area of responsibility. Though fusion centers pre-date the September 11, 2001 terrorist attacks, the concept gained momentum and was promoted by State and local law enforcement and homeland security officials during post-9/11 discussions as a more effective way to protect their communities.

The *National Commission on Terrorist Attacks Upon the United States* (the "9/11 Commission") identified a breakdown in information-sharing as a key factor contributing to the failure to prevent the September 11, 2001 attacks. Its critiques spurred policy that led the Federal Government to support the establishment and sustainment of a national integrated network of State and major urban area fusion centers, and designate fusion centers as the primary focal points within the State and local environment for the receipt and sharing of terrorism and other homeland security-related information and intelligence. Fusion centers provide the Federal Government with critical State and local information and subject-matter expertise that it did not receive in the past—enabling the effective communication of locally generated terrorism-related information.

Yet until recently true *fusion* of data across multiple disciplines and its meaningful analysis was mostly out of reach. At best, the fusion centers provided a *place* where many agencies established co-located terminals, and analysts could exchange views. That itself was a significant gain. At least the data products were going *somewhere* and analysts from different agencies were *talking*. But with data streaming in and no real way to share except by word of mouth, the fusion centers could easily become simply big *new* places where otherwise meaningful information went to die.

The data needed to be melded together in ways that did not rely entirely on humans. While humans would always remain in the "loop," they could not do it all. Machine-to-machine exchange was critical for bringing large volumes of data meaningfully to analysts' eyes for evaluation, and to leaders for decision.

There is no single source for information related to terrorism. Awareness is gained by gathering, fusing, analyzing, and evaluating relevant information from a broad array of sources on a continual basis.

Surely a building block of any successful data fusion could be the lowly but foundational Suspicious Activity Report. With any luck such reports would soon be streaming in, pawns in the great game of chess being played in the war on terror. How to manage, make sense of, and take advantage of this potential treasure trove of data? For somewhere in there, one day, would surely be the dots, again: crucial information about pre-strike planning activities of terrorists on domestic soil.

The NIEM IEPD

In 2007, building on their successes in developing early justice system applications, State, local, and Federal officials and private sector partners came together to explore how to apply XML capabilities and lessons learned to standardizing suspicious activity reporting around the nation.

The Los Angeles Police Department (LAPD) in particular had been in the forefront of such efforts, pioneering Suspicious Activity Reporting and formalizing its management through its own Counterterrorism and Criminal Intelligence Bureau. How could the LAPD's and other pioneers' efforts be leveraged nationally to establish a SAR capability nationwide?

Established as the "Information Sharing Environment Suspicious Activity Report ("ISE-SAR") Functional Standard Development Team" the group confronted a wide disparity of approaches, capabilities, and procedures across the nation's many reporting jurisdictions.

Even as a matter of *definition*, for example, there existed no agreement as to what *constituted* reportable suspicious activity. What Alabama counted as suspicious and reportable, Illinois might take for granted and not report.

With practice disparate city to city, State to State, some saw a risk to Americans' privacy and civil liberties from proposals to "fuse" such data. The American Civil Liberties Union, for example, raised its voice loudly to denounce fusion centers as threats to the Republic and the Constitution. For individual States, laws clearly constrained the sharing of information with Federal agencies, and would require careful re-work with legislatures to authorize. They in turn would be looking for lawful approaches that were mindful of the privacy and civil liberties of citizens.

How would all of these issues get ironed out—so that there was uniformity in the information being gathered and reported, consistency in its process and treatment—and the notion of "suspicious" activity, let alone its handling, was left neither to the avid imaginations nor jaded eyeballs of, potentially, thousands of individual reporters?

Taking Up the Challenge

The ISE-SAR Functional Standard development team—some 35 people with diverse backgrounds including law enforcement, homeland security, and intelligence subject-matter experts and technology experts—met early in 2007 for two and a half pivotal days.

A team leader explained, "We told them we need to figure out a standard way to start sharing information." And that meant standards—standards for what data was collected, how it was collected, and how it would be shared.

At its January meeting, the development team defined what would become elements of a SAR *Information Exchange Package Documentation*, or "IEPD."

"We told them we need to figure out a standard way to start sharing information." And that meant standards—standards for what data was collected, how it was collected, and how it would be shared.

The IEPD would be the document that defined the terms that would comprise a Suspicious Activity Report anywhere a SAR was used or generated by participating agencies. From a technical perspective, it comprised the *data elements* of agency reporting, and as such specified the terms to be shared across jurisdictions, and their metadata tags. For this purpose, the IEPD would draw upon the metadata dictionaries *already* contained in NIEM, to every extent possible *reusing* terms, both those that were based on Global Justice XML, and new entries from other domains.

Using the NIEM construct had another benefit: it provided a framework for discovery and agreement of key policies and business processes. This process eventually led to the development of a SAR Process that includes: multi-level training, a tiered vetting process, a privacy and civil liberties framework, and the ability to share data technically through the SAR IEPD standard. In fact, the NIEM process facilitated a constructive dialogue with privacy and civil liberties advocates—moving the debate from general characterizations on the dangers of collecting SAR data to discussing specific data elements that should be afforded certain privacy protections.

“This just wasn’t dreamed up,” one participant said. “We flowed out a typical transaction and said, ‘Okay, let’s start with that guy who’s taken a picture of the dam. How did it go through the process? Who gets involved and what system supports it?’ We mapped out the process. What’s the precipitating event? What triggers an exchange? What is applicable and what is not?”

“We picked one or two exchanges, and talked about what data elements should be in there,” another participant recalled. “You need a name, and ‘Oh wait a minute, there’s a whole bunch of different names. There’s a person who reported this, there’s the guard that was there, there’s another witness, and there is the suspect and then there’s maybe even a target because they were looking at their binoculars at another building or another person.’ We started modeling the data. We built a data model or domain model around that exchange.”

Perhaps most importantly the ISE-SAR Functional Standard development team arrived at a good first start of a standard for “suspicious activity,” putting some rigor to the term and its use. “Suspicious behavior,” it said, would be defined as “Observed behavior reasonably indicative of pre-operational planning related to terrorism or other criminal activity.” It would include surveillance, photography of facilities, site breach or physical intrusion, cyber attacks, and testing of security, for example.

“Suspicious behavior,” it said, would be defined as “Observed behavior reasonably indicative of pre-operational planning related to terrorism or other criminal activity.”

In January 2008, the Office of the Program Manager for the Information Sharing Environment issued the ISE-SAR Functional Standard codifying the SAR IEPD, the SAR business process and information flow, and the standard’s governance. By the end of 2009, the Nationwide SAR Initiative (“NSI”) had been launched for evaluation purposes in three States and nine cities. It was soon embraced and endorsed by multiple police organizations, and linked to the Department of Homeland Security, to the Department of Defense’s Northern Command, and to the FBI’s eGuardian system.

Further pilot projects and operational developments followed, including within California’s Administrative Office of th e Courts, Florida’s Law Enforcement eXchange, New York’s Division of Criminal Justice Services eJustice Portal, Pennsylvania’s JNET, and the Texas “Path to NIEM.” Federal adoptions also proliferated within FBI, Homeland Security, and Department of Justice systems.

Looking Back, Looking Forward

“There is now for suspicious activity reports,” a program manager stated, *“a standard way*

to express and share information between agencies. You have a standardized set of data. When you look at it from an aggregate level, you start making sense of it. You can start to see patterns or similarities and anomalies."

The development of the SAR IEPD showed that the IEPD is a data dictionary, but much more. Its construct is a formal *process* by which agencies develop, test and prove the exchange of data in reports or queries. It formalizes not just content, but a development path. Those who step down the IEPD road for the development of an exchange model have a well-defined path.

"There is now for suspicious activity reports," a program manager stated, "a standard way to express and share information between agencies."

Moreover, the finished IEPD becomes what is called an "*artifact.*" It is a document in standardized format that anyone can see and quickly understand, and which *persists* even if system developers move on to new positions or leave agency service altogether. This is important as agencies do from time to time reorganize; new individuals come on to the work force, and veteran employees retire.

Once finished, the IEPD can provide a *reusable* basis for any new system to join in the same exchange—meaning it is *scalable* and *extensible*. An IEPD thus permits dynamic network growth. When a new agency wishes to share information with agencies already conformant with the IEPD, they find that the metadictionary is already built, meaning all they have to do is find the right metadata tag for their term. This saves them work, and gives them wide benefits quickly from participation.

Ultimately, the more users on the network the better—for with more users, "network effects" are enhanced for all users, meaning improved efficiency, better information sharing across organizations, and overall gains to performance. Dots can get connected better, faster, and cheaper.

LAPD Commander Joan T. McNamara assessed the operational impact of SAR this way, "While the number of investigations and arrests are important, they are almost secondary to our new-found ability to connect events that in the past would have appeared unrelated. This paints an amazing picture in real time."

□ □ □ □ □

Recently, the ACLU noted that these "strong Federal guidelines" are a "welcome improvement" and called for legislative watchfulness.

The ISE-SAR Functional Standard is moving toward broad adoption, supporting the introduction of two new White House endorsed Program Management Offices—the Nationwide SAR Initiative Program Management Office, and the National Fusion Center Program Management Office.

Globally, Canada has implemented the standard, and Sweden is using the SAR IEPD to enable improved information sharing with their public safety operations.

A Portfolio of NIEM Success Stories is sponsored by the
Office of the Program Manager, Information Sharing Environment (www.ise.gov)

For more information on the National Information Exchange Model visit www.niem.gov

For further information on the Nationwide Suspicious Activity Reporting Initiative
please visit www.ncirc.gov/sar