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, swarming with tourists.

Families crowd their RVs into choice spots opposite the dam's towering concrete curves, waiting to see the nightly laser projections that cover the dam's wall. Tourists from other nations with strong engineering cultures also take tours, going down in the elevators to see the inner workings 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 nonstop.

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

In the midst of this flow of cheerful tourists, always observant, are the local police, dam security personnel, and federal agents. For them, there is one primary, constant question: How do we keep this superstructure safe from vandals, criminals, and terrorists? And how can we distinguish between tourists innocently seeking souvenir images, and terrorists engaged in preoperational planning with the intent to destroy the dam and all that stands in its path?

In an era of heightened awareness—when "If You See Something, Say Something" is the byword—every cop on the beat is a human

sensor with eyes and ears ever alert. As is every citizen.

When something doesn't seem quite right, citizens are encouraged to notify authorities, and police officers are expected to make a note of the observed conduct—even if the conduct is not a crime.

Suppose a tourist tells a police officer that someone with "a strange accent" is asking "too many" questions of a tour guide. What might this mean? And in an America rich in regional dialects, an America whose citizens include immigrants old and new from all over the world, how do we define a "strange accent," anyway?

In the past, officers may have noted the mention and filed it away, or shared the information with the next officer on duty, using stickies, notepads, or the backs of envelopes to record the information. But recognizing the potential value of many of these observations, many police departments have in recent years formalized the procedure, adopting a standardized Suspicious Activity Report, or SAR, as protocol.

Today in many departments, officers receiving information about suspicious behavior, or observing it themselves, may make a formal note—and fill out a SAR. (Not all departments have adopted the use of SARs, but many have.)

One question facing departments is how best to make use of this data—captured, as it is, by individuals in many different departments, in nonstandard formats, with differing definitions, and maintained by independent agencies that have neither the will, the authority, the finances, nor the process to share such information and collaborate in analyzing it.

What if, for example, the week after the suspicious activity is noticed at the Grand Coulee Dam, a SAR is recorded about a similar person asking similar questions 900 miles away at Hoover Dam outside Las Vegas? How will staff at separate agencies, or even members of the same agencies in different places, *connect the dots*? What will enable them to see a pattern—if there is one—in two seemingly unrelated events? Events that should raise not just eyebrows, but serious concerns, and trigger effective follow-up measures?

It would be easy enough for officials to secure the Grand Coulee Dam, or for that matter any other infrastructure, from preoperational terrorist exploits: simply close them to the public and secure their perimeters. But, as Winston Churchill once famously said in response to a senior aide's recommendation to close London's museums and theaters during the Blitz, "Damn it, man, we're fighting to keep them open!"

In an open society, even in an age of terrorism, officials charged with maintaining the security of locations such as the Grand Coulee Dam must do so, while also ensuring that Americans and visitors from other countries are free to enjoy the benefits of visiting them.

An open society also guarantees civil 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 amidst the tide of innocent visitors have been, and one day will likely be again, men and women, and perhaps even boys and girls, who are engaged not in innocent sightseeing, but in preoperational planning for terrorist strikes. Men, for example, who are training to pilot planes but who show no interest in learning how to land them.

Until very recently, those involved in preoperational planning for terrorist activity 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 with each other. Laws prevented many state and local agencies from sharing information with federal enforcement organizations. *What would become of the information? Where would it be stored? Who could access, see, and use it?*

Enter the National Information Exchange Model (NIEM)

Even before the terrorist attacks on September 11, 2001, 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 creation of a common language to be used in their computer systems; the process for arriving at such agreements; and governance of the relationship between parties entering into these agreements.

Because over the years many separate computer systems had sprung up on the American law enforcement landscape, each with different names for the same things, a lack of *interoperability* among justice-related systems at the state and local levels was common. Such technical obstacles to information sharing among agencies and departments created risk and inefficiency, and negatively affected performance—often with dire consequences.

For example, where judicial, welfare, and health agencies might all have information about a child at risk of abuse, each data system might use different words 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 the same person in the real world. As long as there was no way to translate these terms from one system to another, it was impossible to exchange data meaningfully among the systems—or in at least some cases, to do so in time.

As a result, dots that should have been connected—dots which might point to a child at risk—went unconnected. Authorities would sometimes discover the information *too late* to prevent harm; at other times they might have moved *too quickly*, breaking apart families unnecessarily.

With the advent of extensible mark-up languages (XML) and their many subject matter-specific terms, much changed.

Using XML-based metadata (data about data), state and local justice agencies and their federal counterparts that wished to exchange information—where lawful and appropriate—could keep their own "legacy" system names for things and agree to use a *metadata* dictionary to facilitate interagency or interdepartmental communication.

With the metadata agreed to in an *information exchange model*, each department or agency could continue to "speak" its own language, leaving the huge legacy systems unchanged except for the tagging of information. But each could now also send and understand messages to and from other agencies and departments. The XML-based exchange model enabled all participating entities to quickly translate and share data between their systems.

For example, one department might use the term "automobile." Another might say "passenger vehicle." They would agree to use the data element "car." This would allow their

computer systems to efficiently exchange data without having to change their internal terminology.

This was an elementary but important breakthrough, allowing for greater efficiency, transparency, and improved performance in information sharing. Analysts could run reports, statisticians could find patterns, and policymakers could better understand the results, trends, and options showing up in their data.

The dots could get connected.

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

Global Justice XML, as this became known, emerged as a "win-win" tool for everyone, *transforming* the value of information assets in disparate systems, which previously had been 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 ensured consistency in approach and results.

In the same way, a National Information Exchange Model (NIEM), 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 from law enforcement, health, energy, or transportation—provided they each made their systems conform to 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. Over the past decade, these three strands of government 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, the passage of the Intelligence Reform and Terrorism Prevention Act of 2004, and the 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 previous initiative focused on streamlining information gathering and sharing across the federal government started with the Clinger-Cohen Act of 1996 and 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. Today, NIEM is the leading implementation of that reference model.

Information Sharing in the Age of Terrorism

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.

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.

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

Fusion Centers

Until the initial openings of fusion centers in 1996, information gathered by all these sources often remained isolated within systems and organizations that could not, or would not, share information with each other.

Fusion centers receive information from a variety of sources, including federal, state, and local entities. They then ensure that timely and relevant information is provided to the right stakeholders within their geographic areas of responsibility. The fusion centers are an analytic resource that supports the efforts of state and local law enforcement to prevent and investigate crime and terrorism in local communities. Though fusion centers predated 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. The critiques of the Commission 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 to designate fusion centers as the primary focal points within the state and local environment for the receipt and sharing of information about 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 nationwide communication of locally generated terrorism-related information.

Yet until recently, true *fusion* of data across multiple disciplines and its meaningful analysis was still mostly out of reach. At best, the fusion centers provided a *place* where many agencies established colocated terminals, and analysts could exchange views with each other in real time. That in itself was a significant gain. At least the data products were going *to centralized locations*, and analysts from different agencies were *talking to each other about the information they were reviewing*. But with data streaming in and no real way to share it except by word of mouth, the fusion centers could easily have become more big *new* places where otherwise meaningful information went to die.

In order to make the best use of the data received, it 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 all the analysis needed alone. Machine-to-machine exchange was critical for bringing large volumes of data meaningfully to analysts' eyes for evaluation, and to leaders for decision-making.

Surely one building block of any successful data fusion could be the simple but foundational Suspicious Activity Report (SAR). With any luck, such reports would soon be streaming in, pawns in the great game of chess being played in the war on terrorism. A key question

was how to manage, make sense of, and take advantage of this potential treasure trove of data. For somewhere in there would surely be a set of dots in need of connection again, one day: crucial information about the prestrike planning activities of terrorists on domestic soil.

The SAR Information Exchange Package Documentation (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 pioneering efforts of the LAPD and others be leveraged nationally to establish a SAR capability nationwide?

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

Even *defining* suspicious activity proved challenging: for there was no agreement as to what constituted reportable suspicious activity. What officers in Alabama considered suspicious and reportable, their cohorts in Illinois might not consider suspicious, and therefore might not report.

With disparate practices from city to city and state to state, some saw a risk to Americans' privacy and civil liberties in proposals to "fuse" such data. The American Civil Liberties Union (ACLU), for example, raised its voice loudly to denounce fusion centers as threats to the Republic and to constitutional protections. In some states, laws clearly constrained the sharing of information with federal agencies, and careful work with legislatures was needed to authorize such sharing. Legislators, in turn, looked for lawful approaches that were mindful of the importance of protecting the privacy rights 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 the way it was processed and treated? And that the definition of "suspicious" activity, and procedures for how it was handled, were subject neither to overly avid imaginations nor to the jaded or careless eyeballs of potentially thousands of individual reporters?

Taking Up the Challenge

The ISE-SAR Functional Standard Development Team—35 experts with diverse backgrounds in law enforcement, homeland security, intelligence, and technology—met early in 2007 for two and a half pivotal days to create guidelines for defining and reporting suspicious activity.

One team leader explained, "We told the group we needed to figure out a standard way to start sharing information." That meant developing standards—standards for what kind of data was to be collected, how it would be collected, and how it would be shared.

At this meeting, the Development Team defined what would become the elements of a SAR Information Exchange Package Documentation, or IEPD.

The IEPD is 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 across agencies and departments. This process eventually led to the development of a SAR process that included multilevel training, a tiered vetting process, a privacy and civil liberties framework, and the ability to share data technically through the SAR IEPD standard.

The NIEM process also facilitated a constructive dialogue with privacy and civil liberties advocates—moving the debate from general discussions about the dangers of collecting SAR data to identifying specific data elements that should be afforded certain privacy protections.

"This wasn't just 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 the information 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 you realize, '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 suspicious person and then there's maybe even a target, because they were looking through their binoculars at another building or another person.' We started modeling the data. And 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 basis for establishing a standard for defining "suspicious activity," putting some rigor to the term and its use. Suspicious activity, it said, would be defined as "observed behavior reasonably indicative of preoperational planning related to terrorism or other criminal activity." This would include such things as surveillance or photography of facilities, breach of or physical intrusion into a site, cyberattacks, and the testing of security.

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, the Department of Defense's Northern Command, and the FBI's eGuardian system.

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."

> **"There is now for suspicious activity reports a standard way to express and share information between agencies."**

The development of the SAR IEPD showed that the IEPD is an effective data dictionary, but it is also much more. Its construct offered a formal *process* by which agencies developed, tested, and proved the validity of data exchanged in reports or queries. It formalized not just content, but a development path. Those who use the IEPD for the development of an exchange model have a well-defined path to follow.

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

The IEPD provided 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 users in a new agency wish to share information with agencies already conformant with the IEPD, they find that the metadictionary is already built, so all they have to do is identify the right metadata tags for their system's terminology. This saves them work and gives them wide benefits quickly as they join the network.

Ultimately, the more users in 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 performance gains. Dots can get connected better, faster, and at less cost.

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 newfound ability to connect events that in the past would have appeared unrelated. This paints an amazing picture in real time."

▫ ▫ ▫ ▫ ▫

The ISE-SAR Functional Standard is moving toward broad acceptance and adoption. The ACLU recently noted, for example, that SAR's "strong federal guidelines" are a "welcome improvement" and called for legislative watchfulness. New "fusion-center-in-a-box" solutions have entered the commercial marketplace. And the White House has introduced two new Program Management Offices—the Nationwide SAR Initiative Program Management Office and the National Fusion Center Program Management Office.

The standard has also been implemented in Canada, and Sweden is using the SAR IEPD to enable improved information sharing with its 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 NIEM, visit www.NIEM.gov.

07052011