Proposed ANSI/NIST XML Short-Tag Format

Shahram Orandi
July 27, 2010
NIST's core competencies:
- Measurement science
- Rigorous traceability
- Development and use of standards

One of these standards is the ANSI/NIST standard.

ANSI/NIST XML Short Tag
Where did the ANSI/NIST standard begin?

ANSI/NIST XML Short Tag
What’s good about ANSI/NIST-Traditional:

-Almost a quarter century old: well penetrated in the field, stable, proven and tested.

-The format is the basis of, or supported by, nearly all law enforcement applications on the globe.

-Structurally ”Monolithic”

-Simple binary encoding format (it was designed for an 8-bit world with slow computers and even slower networks/modems)

ANSI/NIST XML Short Tag
What’s lacking for ANSI/NIST-Traditional:

- Requires niche tools for data manipulation.

- General purpose systems can’t do much to the data unless made specifically aware of the format (i.e., COTS data orchestration & crypto components need to be customized or specifically coded to handle it).

- It’s built for an 8-bit world, and it has some of its limitations built in.

ANSI/NIST XML Short Tag
The ANSI/NIST NIEM-Conformant XML was designed to be:

- A modern format, compatible with modern information technology constructs such as web services, SOAP, etc.
- Have greater availability of tools via large selection of XML tools, libraries and appliances.
- General purpose XML orchestration tools can parse and manipulate most of this data format readily.
Some lessons learned from ANSI/NIST NIEM-Conformant XML:

• Complexity: translates to time and money
• Dependencies in flux (3 updates to NIEM in 3 years)
• Radical change from Legacy tag format, very US-English-centric
• Faced a mountain of legacy data and systems that may not be ready for a very large change.

ANSI/NIST XML Short Tag
What is short-tag?

An XML data “serialization”* approach for biometric (and biographic) data, using existing ANSI/NIST field 3 or 4 character mnemonic references. E.g., IMP vs. FingerprintImageImpressionCaptureCategoryCode

*Serialization in layman's terms: Taking a complex data structure and bulldozing it flat for transfer, storage, etc.

ANSI/NIST XML Short Tag
And the same data, in long-tag format

```xml
<?xml version="1.0" encoding="UTF-8"?>
<itl:NISTBiometricInformationExchangePackage
    xmlns:ansi-nist="http://niem.gov/niem/ansi-nist/2.0"
    xmlns:itl="http://biometrics.nist.gov/standard/2-2008"
    xmlns:nc="http://niem.gov/niem/ncore/2.0"
    xmlns:ebts="http://cjis.fbi.gov/fbi_ebts/beta_1.0.2">
    <itl:PackageInformationRecord>
        <ansi-nist:RecordCategoryCode>01</ansi-nist:RecordCategoryCode>
        <ansi-nist:Transaction>
            <ansi-nist:TransactionDate>
                <nc:Date>2009-09-21</nc:Date>
            </ansi-nist:TransactionDate>
            <ansi-nist:TransactionDestinationOrganization>
                <nc:OrganizationIdentification>
                    <nc:IdentificationID>MDNCANIST</nc:IdentificationID>
                </nc:OrganizationIdentification>
            </ansi-nist:TransactionDestinationOrganization>
            <ansi-nist:TransactionOriginatingOrganization>
                <nc:OrganizationIdentification>
                    <nc:IdentificationID>MDNCANIST</nc:IdentificationID>
                </nc:OrganizationIdentification>
            </ansi-nist:TransactionOriginatingOrganization>
            <ansi-nist:TransactionUTCDate>
                <nc:DateTime>2009-09-21T15:27:43Z</nc:DateTime>
            </ansi-nist:TransactionUTCDate>
            <ansi-nist:TransactionControlIdentification>
                <nc:IdentificationID>FBI_JABS0001</nc:IdentificationID>
            </ansi-nist:TransactionControlIdentification>
            <ansi-nist:TransactionDomain>
                <ansi-nist:DomainVersionNumberIdentification>
                    <nc:IdentificationID>NORAM</nc:IdentificationID>
                </ansi-nist:DomainVersionNumberIdentification>
            </ansi-nist:TransactionDomain>
            <ansi-nist:TransactionImageResolutionDetails>
                <ansi-nist:NativeScanningResolutionValue>19.69</ansi-nist:NativeScanningResolutionValue>
                <ansi-nist:NominalTransmittingResolutionValue>19.69</ansi-nist:NominalTransmittingResolutionValue>
            </ansi-nist:TransactionImageResolutionDetails>
            <ansi-nist:TransactionMajorVersionValue>05</ansi-nist:TransactionMajorVersionValue>
            <ansi-nist:TransactionMinorVersionValue>00</ansi-nist:TransactionMinorVersionValue>
        </ansi-nist:Transaction>
    </itl:PackageInformationRecord>
</itl:NISTBiometricInformationExchangePackage>
```

ANSI/NIST XML Short Tag
Short-tag’s goals:

- Be an evolution of Legacy (Part-1), not something radically different. If you’re familiar with Legacy, you’ll understand most of short tag out of the box.

- Guarantee 1:1 equivalency with legacy format. Legacy can be translated to Short-tag, and back without pieces breaking off permanently.

- Short-tag will be lock-stepped with Legacy. No out of sync updates to XML/Legacy. When one evolves, so will the other.

- Will use legacy tag structure that is well accepted (i.e., TOT, ORI, etc.) rather than English words.

- Simple: Means cheap, easy to implement.

- No external dependencies.

- 1:1 allows for legacy users to maintain legacy systems and jump to XML anywhere in their Legacy pipeline that they’re comfortable with.

- Lets stakeholders get footing with XML on their terms. Either a little... Or a lot... Or none-at-all (caveat: if someone sends you short-tag, you’ll need to recode it to legacy at your front door).

ANSI/NIST XML Short Tag
Some positive side effects of Short-tag:
- While XML is on average significantly larger than comparable binary format, short-tag is relatively small for XML.

- XML’s attributes allow for flexibility and the possibility of inserting data transform markers or other data markers inline. Here is a hypothetical example:

<table>
<thead>
<tr>
<th>Short Tag</th>
<th>Short Tag w/Attribute Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;ANSINISTShort&gt;</td>
<td>&lt;ANSINISTShort&gt;</td>
</tr>
<tr>
<td>&lt;Record&gt;</td>
<td>&lt;Record&gt;</td>
</tr>
<tr>
<td>&lt;RecordType&gt;01&lt;/RecordType&gt;</td>
<td>&lt;RecordType&gt;01&lt;/RecordType&gt;</td>
</tr>
<tr>
<td>&lt;DAT&gt;20090921&lt;/DAT&gt;</td>
<td>&lt;DAT altSys=&quot;تاريخ&quot;&gt;20090921&lt;/DAT&gt;</td>
</tr>
<tr>
<td>&lt;DAI&gt;DAI000000&lt;/DAI&gt;</td>
<td>&lt;DAI altSys=&quot;目的地&quot;&gt;DAI000000&lt;/DAI&gt;</td>
</tr>
<tr>
<td>&lt;ORI&gt;MDCANIST&lt;/ORI&gt;</td>
<td>&lt;ORI altSys=&quot;походження&quot;&gt;MDCANIST&lt;/ORI&gt;</td>
</tr>
<tr>
<td>&lt;TCN&gt;FBI_JABS0001&lt;/TCN&gt;</td>
<td>&lt;TCN altSys=&quot;NIEM_TCN&quot;&gt;FBI_JABS0001&lt;/TCN&gt;</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Summary of Short-tag proposal:

- Be an evolution of Legacy (Part-1)
- Guarantee 1:1 equivalency with legacy format at all times.
- Short-tag will be locked into the standard as a whole.
- Will use legacy mnemonics that is already well accepted (TOT, etc.)
- Simple.
- No external dependencies.
- Smaller than some other XML serialization approaches (bonus, easier on resources)

ANSI/NIST XML Short Tag
Q & A?

Contact:
Shahram Orandi
sorandi@nist.gov
301-975-3261

ANSI/NIST XML Short Tag