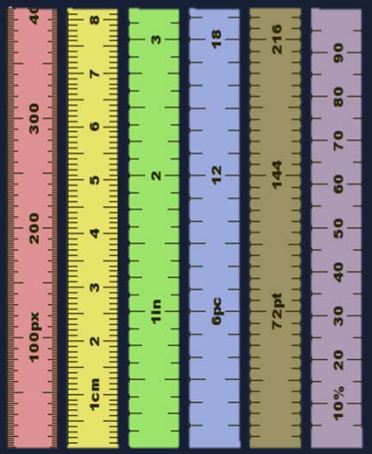
Proposed ANSI/NIST XML Short-Tag Format

Shahram Orandi
July 27, 2010



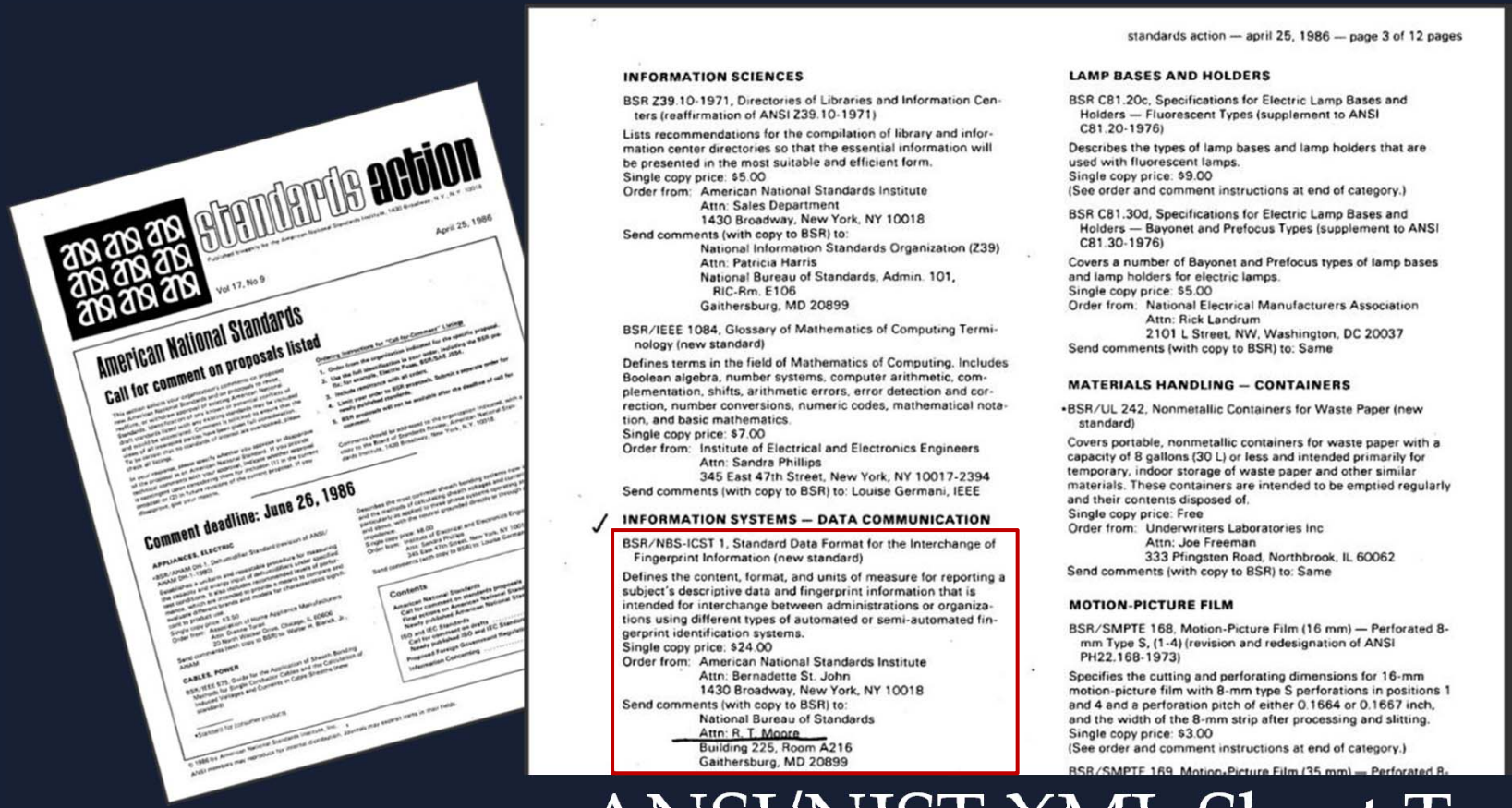
NIST's core competencies:

- 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?



standards action — april 25, 1986 — page 3 of 12 pages

INFORMATION SCIENCES

BSR Z39.10-1971, Directories of Libraries and Information Centers (reaffirmation of ANSI Z39.10-1971)

Lists recommendations for the compilation of library and information center directories so that the essential information will be presented in the most suitable and efficient form.
Single copy price: $5.00
Order from: American National Standards Institute
Attn: Sales Department
1430 Broadway, New York, NY 10018
Send comments (with copy to BSR) to:
National Information Standards Organization (Z39)
Attn: Patricia Harris
National Bureau of Standards, Admin. 101,
RIC-Rm. E106
Gaithersburg, MD 20899

BSR/IEEE 1084, Glossary of Mathematics of Computing Terminology (new standard)

Defines terms in the field of Mathematics of Computing. Includes Boolean algebra, number systems, computer arithmetic, complementation, shifts, arithmetic errors, error detection and correction, number conversions, numeric codes, mathematical notation, and basic mathematics.
Single copy price: $7.00
Order from: Institute of Electrical and Electronics Engineers
Attn: Sandra Phillips
345 East 47th Street, New York, NY 10017-2394
Send comments (with copy to BSR) to: Louise Germani, IEEE

✓ **INFORMATION SYSTEMS — DATA COMMUNICATION**

BSR/NBS-ICST 1, Standard Data Format for the Interchange of Fingerprint Information (new standard)

Defines the content, format, and units of measure for reporting a subject's descriptive data and fingerprint information that is intended for interchange between administrations or organizations using different types of automated or semi-automated fingerprint identification systems.
Single copy price: $24.00
Order from: American National Standards Institute
Attn: Bernadette St. John
1430 Broadway, New York, NY 10018
Send comments (with copy to BSR) to:
National Bureau of Standards
Attn: R. T. Moore
Building 225, Room A216
Gaithersburg, MD 20899

LAMP BASES AND HOLDERS

BSR C81.20c, Specifications for Electric Lamp Bases and Holders — Fluorescent Types (supplement to ANSI C81.20-1976)

Describes the types of lamp bases and lamp holders that are used with fluorescent lamps.
Single copy price: $9.00
(See order and comment instructions at end of category.)

BSR C81.30d, Specifications for Electric Lamp Bases and Holders — Bayonet and Prefocus Types (supplement to ANSI C81.30-1976)

Covers a number of Bayonet and Prefocus types of lamp bases and lamp holders for electric lamps.
Single copy price: $5.00
Order from: National Electrical Manufacturers Association
Attn: Rick Landrum
2101 L Street, NW, Washington, DC 20037
Send comments (with copy to BSR) to: Same

MATERIALS HANDLING — CONTAINERS

•BSR/UL 242, Nonmetallic Containers for Waste Paper (new standard)

Covers portable, nonmetallic containers for waste paper with a capacity of 8 gallons (30 L) or less and intended primarily for temporary, indoor storage of waste paper and other similar materials. These containers are intended to be emptied regularly and their contents disposed of.
Single copy price: Free
Order from: Underwriters Laboratories Inc
Attn: Joe Freeman
333 Pfingsten Road, Northbrook, IL 60062
Send comments (with copy to BSR) to: Same

MOTION-PICTURE FILM

BSR/SMPTE 168, Motion-Picture Film (16 mm) — Perforated 8-mm Type S, (1-4) (revision and redesignation of ANSI PH22.168-1973)

Specifies the cutting and perforating dimensions for 16-mm motion-picture film with 8-mm type S perforations in positions 1 and 4 and a perforation pitch of either 0.1664 or 0.1667 inch, and the width of the 8-mm strip after processing and slitting.
Single copy price: $3.00
(See order and comment instructions at end of category.)

BSR/SMPTE 169, Motion-Picture Film (35 mm) — Perforated 8-

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.

ANSI/NIST XML Short Tag

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

ANSI/NIST XML Short Tag

Here is a short tag fragment:

```
<?xml version="1.0  encoding="UTF-8"?>
<ANSINISTShort>
    <Record>
                <RecordType>01</RecordType>
                <DAT>20090921</DAT>
                <DAI>DAI000000</DAI>
                <ORI>MDNCANIST</ORI>
                <TCN>FBI_JABS0001</TCN>
                <DOM>NORAM</DOM>
                <NSR>19.69</NSR>
                <NTR>19.69</NTR>
                <VER>0500</VER>
                <TOT>NFXX</TOT>
                ...
```

And the same data, in long-tag format

```
<?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/niem-core/2.0
 xmlns:incits="http://biometrics.nist.gov/standard/2-2008-AnnexG"
 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>DAI000000</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></nc:IdentificationID>
            </ansi-nist:DomainVersionNumberIdentification>
            <ansi-nist:OrganizationName>NORAM</ansi-nist:OrganizationName>
         </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:TransactionCategoryCode>NFXX</ansi-nist:TransactionCategoryCode>
         ...
```

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:

Short Tag	Short Tag w/Attribute Markers
<ANSINISTShort> <Record> <RecordType>01</RecordType> <DAT>20090921</DAT> <DAI>DAI000000</DAI> <ORI>MDNCANIST</ORI> <TCN>FBI_JABS0001</TCN> …	<ANSINISTShort> <Record> <RecordType>01</RecordType> <DAT altSys="تاريخ">20090921</DAT> <DAI altSys="目的地">DAI000000</DAI> <ORI altSys="походження">MDNCANIST</ORI> <TCN altSys="NIEM_TCN">FBI_JABS0001</TCN> …

ANSI/NIST XML Short Tag

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