

XML Markup of Court Documents for Data Capture and Re-Use

The OASIS LegalXML Member
Section's Electronic Court Filing
Technical Committee

Court Filing Documents Subcommittee

August 3rd, 2007

The Commitment

- From the Charter of the LegalXML Electronic Court Filing Technical Committee (TC):
 - ***“The TC will develop techniques and principles for marking up legal documents and forms for the purpose of data capture and re-use.”***

Subcommittee Products

- Written ASSUMPTIONS about how court documents can pick up their tags so their filed documents can be processed with a maximum of automation, along with other assumptions about how this sort of system would operate.
- Report on the result of researches into how such pattern form based documents can be assembled, using available technology, assuming that most filers would not be able or willing to insert tags on their own.
 - What is available to help produce court filings that are tagged correctly so data processing and information exchanging can be automated?
 - How does this also work to produce human-readable documents?

Assumptions

1. Most court filings are submitted in a format that could be treated as a “pattern form.”
2. Much of “Clerk Review” entails capturing basic data elements, most often by a clerk re-typing the data into a target system (e.g., a document index, calendaring program).
3. Minimum data elements for processing most court filings include:
 - Identification of the court
 - Reference number to the case
 - “Caption” for the case (e.g., Smith v. Jones)
 - Title for the filed document
 - Filer(s) of the filed document
 - Type of document filed (e.g., Pleading, Petition, Order)
 - Date and time for key events related to the filing

Assumptions

4. Conventions for inserting XML tags for data elements should be compliant with NIEM, GJXDM.
5. XML tagging in court filed documents is to be done in addition to XML tagging that is due to the use of LegalXML standards for Electronic Court Filing.
6. Users/filers/document creators are not able or willing to insert XML tags into documents as they create them.
 - User-created tagging would be error-prone.

Assumptions

7. Method 1: XML tags for data elements would be placed in specific corresponding locations in a **“pattern form”**.
 - Tagging would have to be designed for the document type.
 - Tagging will be “under the hood,” not observed by the author...provided a current “pattern form” is used (obtained afresh each time it is to be used).
 - Forms administration, including updating and access management are ever more important challenges for the court and clerk to administer.

Assumptions

8. Method 2: A “smart form” would present the form user with a series of questions, modifying the end product based on responses (**“wizard” driven**) *and* it would assign the correct data element tags in the process.
 - Designing “smart forms” correctly will be a challenge.
 - Maintaining, updating, and controlling access to such forms will be an ongoing challenge (and expense) for the court/clerk.

Assumptions

9. Forms management is always challenging because private entities routinely capture and re-sell court forms.
10. Documents created to capture key data elements for automation of clerk review must also produce a human-readable version of the document.
11. Electronic court filing systems are capable of delivering such XML-mark-up documents as court filings and to preserve them in standard digital form indefinitely.
 - XML tagging can be useful for future queries and searches, i.e., not solely for processing incoming court filings.

Assumptions

12. Development of XML court filing documents affords opportunities to:
- Establish rules/methods to ensure that certain tagged “sensitive” data will be automatically hidden (dropped out, obliterated, **redacted**) unless displayed to owner of an appropriate security privilege level.
 - Establish rules/methods whereby forms vendors, law firm managers, and others can add XML tagging for their own purposes (e.g., client billing).

Data elements in the King County *Judgment and Sentence* form are in filled-in blanks, check boxes, and interlineated notes by the judge. To use the data, an authenticated copy must be used, key data elements located to be manually keyed in to other document(s) or system(s).

XML TAG NAMES

SUPERIOR COURT OF WASHINGTON FOR KING COUNTY

STATE OF WASHINGTON,

Plaintiff, No. 01-1-10261-0 KNT

Vs. ANGELO M. [REDACTED]

Defendant.

JUDGMENT AND SENTENCE

30

I. HEARING

1.1 The defendant, the defendant's lawyer, SARA [REDACTED] EN, and the deputy prosecuting attorney were present at the sentencing hearing conducted today. Others present were: _____

II. FINDINGS

There being no reason why judgment should not be pronounced, the court finds:

2.1 **CURRENT OFFENSE(S):** The defendant was found guilty on 04/16/2002 by the court:

Count No.: 1	Crime: THEFT IN THE FIRST DEGREE	Crime Code: 0A514
RCW: 9A, 56.030 (1)-(A)		Incident No.:
Date of Crime: 05/31/2001		
Count No.:	Crime:	Crime Code:
RCW:		Incident No.:
Date of Crime:		

county

courtName

caseNumber

caseTitle

count

offenseDate

plainLanguageDescription

chargeIdentifier

**Standard data elements
found throughout the
document**

The Paper Document

Human-readable level

(display as words/data on a surface for human consumption)

Elements of trust

(assumptions, inked signatures, media, procedures, laws)

Optional: backup, security, other copy

(protection against loss of information, microfilm, etc.)

Optional: managed record

(adding value to documents and their information)

Optional: machine-readable version

(word-processed, imaged, PDF, etc.)

The **XML** Document

Human-readable level

(display on a surface for human consumption / "style sheets")

Data & information "marked-up"

(lets applications process based on standard mark-up)

Standard vocabulary, thesauri, etc.

(uniform semantic meanings, "under the hood")

Security, authenticity, reliability

(establish technical trust in the electronic document)

Technical architecture, messages, exchanges

(establish trust in exchanges of documents, data)

Using XML...

- Once key-stroked (or otherwise captured), *data should never be key-stroked again!*
- Hold on to the meta-data—not just the human-readable (on paper)
- Repurpose/reuse data, but ensure validation, trust, etc., are achieved
- Support standards; they grow in importance with the scope of our business interactions

Why Do We Make Electronic Records Act Like Paper?

- For human use, a document has to be displayed as “words on paper” (or “on screen”)
- We think of records in terms of paper containers – shelves, boxes, folders, envelopes, staples, clips, books, binders...
- With paper, data is, we believe, fixed to the medium on which it appears
 - Hard to imagine data freed from its medium
 - We mistrust data when separated from the medium
- We need to re-define “Paper”—to tell it like it is...

~~PAPER~~ **Definition** ~~PAPER~~ .

highly portable, non-
refreshable screen on
which we display
electronically stored
information.

—Roger Winters (1999)

Overview of Proposal

- Create the NEW Documents Specification through:
 - Defining the steps of work on the new Documents Specification following well-known software engineering practices
 - And by creating an implementation of the new Documents Specification
- LegalXML already published “Court Document 1.1,” but it did not include automation of data capture during clerk review

How This Fits in with ECF 3.0.1

- ECF Specification doesn't distinguish among types of "payloads" delivered by a standardized "envelope."
 - Allows for PDF's, Images, Text, XML, MS Word, and other formats (support based on court choice) wrapped in a common level enveloping structure "on the wire."
 - The NEW Documents Specification will allow for all variations of a "document" as the payload of an ECF 3.0.1 message.

How This Fits in with ECF 3.0.1

- An example: WARRANT
 - A warrant document in one county, however customized it may be, can be “wrapped” into the common level defined by ECF 3.0.1.
 - No matter how customized and varied the inner XML document is, the outer layer will have universal, common, justice-specific related elements that pertain to a warrant.
 - Using the new Documents Specification, the construction of the specific document will be achieved by using a common building framework.

Creating the Implementation

- An implementation, the evolution of the specification, and compliance with ECF 3.0.1 -- all take place in parallel.
- Implementing a demonstration system:
 - Will serve as a feedback source to obtain information, criticism, and ideas for improving the Specification.
 - Will also generate feedback on ECF 3.01, since it must also be an implementation of ECF 3.0.1.

Next Steps

- Obtain criticism and other feedback from the TC:
 - About the proposal for building the Specification
 - About the Assumptions
 - About an initial Implementation
- Adjust work plan based on TC feedback.
- Drafting of proposal for building a new Court Filing Document Specification (the charter for work).

Next Steps

- Reconstituting the Court Filing Documents Subcommittee.
 - Expertise on Clerk Review and other document processes.
 - Expertise on implementing the current ECF specification.
 - Involvement of creators, users of court case documents.
 - Expertise on NIEM and GJXDM
 - Current Subcommittee members may leave, become “Observers,” remain involved for this work, etc.
- Convene Documents Subcommittee meetings:
 - Identify tasks, take assignments to draft Specification.

Next Steps

- Anticipated work plan:
 - Seek out, create, and refine Use Cases
 - Create needed System Architecture
 - Create needed System Design
 - Implement components of Design
 - Perform iterations of:
 - Revision of the Specification
 - Implementing new Court Filing Document Specification
 - Draft new Court Documents Specification
 - LegalXML ECFTC receives, reviews, revises, approves

Actions to Be Taken

- Obtain go-ahead from ECFTC (NOW)
- Co-chairs (McElrath & Winters) to contact those on current roster of Subcommittee:
 - Request each to change their status in the Subcommittee to reflect one's actual interest in this Specification:
 - Voting member (active participant)
 - Observer (able to read, review, not to participate)
 - Drop from Subcommittee (if unavailable; on request)

Contact the Subcommittee

- Co-Chair Roger Winters
 - King County Judicial Administration
 - (206) 296-7838
 - Roger.Winters@kingcounty.gov
- Co-Chair Rex McElrath
 - Georgia Courts
 - (404) 657-9218
 - mcelratr@gaaoc.us
- Michael Alexandrou
 - Georgia Courts
 - (404) 656-7788
 - alexandrm@gaaoc.us