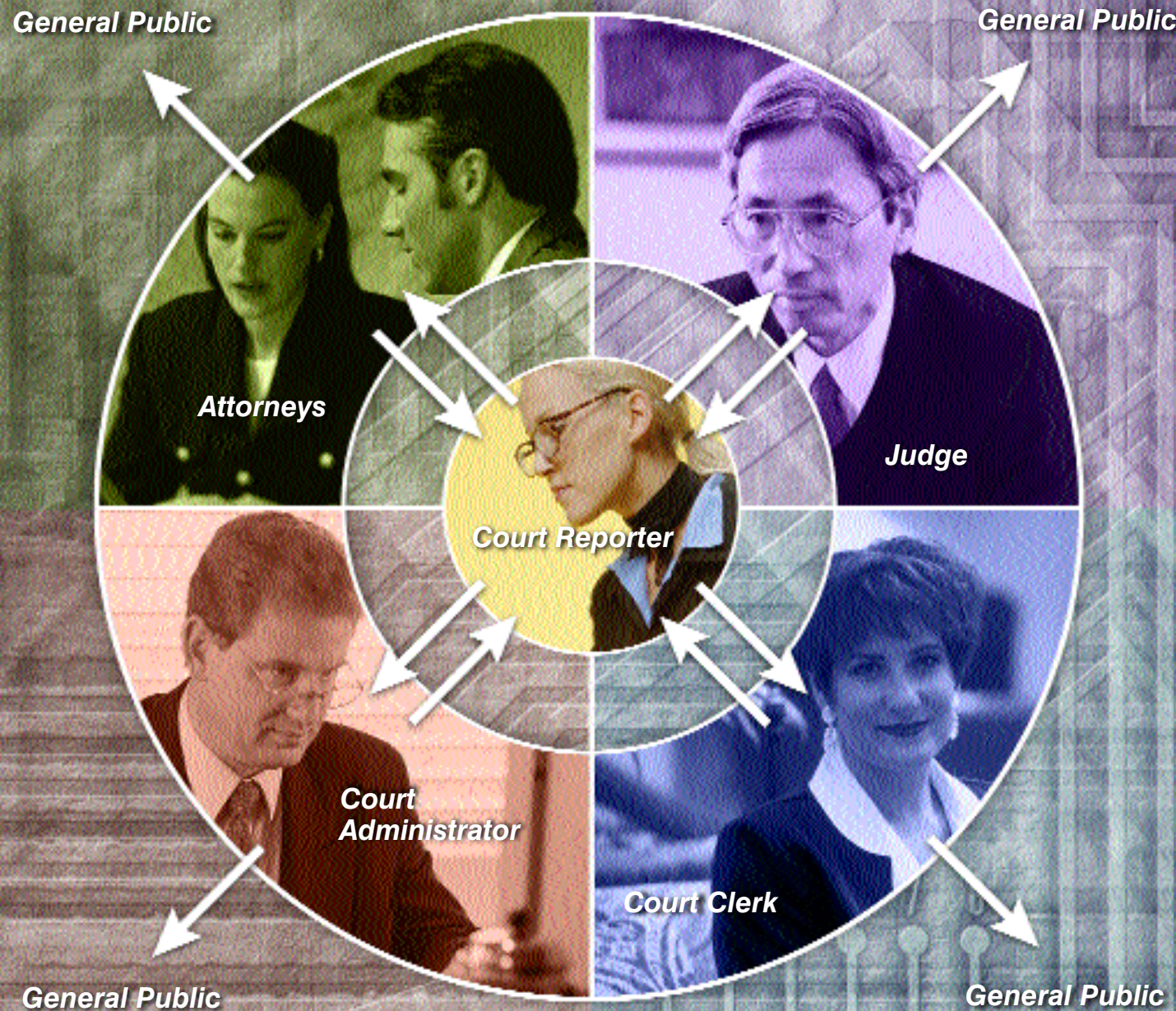


Legal XML: New Standards for Legal Documents

BY EDDIE O'BRIEN



Legal XML expands the amount of information that can be taken from the reporter's digital transcript and offers greater flexibility in how that transcript is searched and used.

Court reporters are often described as being the catalysts for bringing new technology into the legal environment. That's been the case since the 1970s, when reporters first introduced attorneys, judges and court administrators to computer-aided transcription. The progression then continued with the development of realtime and a host of other technologies, such as video-text integration, litigation-support software, Reporter Electronic Data Interchange and transcript delivery over the Internet.

Court reporters may have another opportunity to prove their worth to

About the Author

Eddie O'Brien is president and CEO of Ringtail Solutions Inc. and co-chairman of the Legal XML Transcripts Workgroup. You can reach Eddie at eddie@ringtail.com.au or www.ringtail.com.au.

the legal system with the development of a new application. The World Wide Web — and the people who use it — are approaching new levels of sophistication with the advent of XML.

XML is short for Extensible Markup Language. It is based on the same basic principles as Hypertext Markup Language, the lingua franca of the Web. HTML is powerful but simplistic and imprecise. In contrast, as the Web moves toward its third generation, many are convinced that XML offers precise descriptions and standards that will open up new worlds of possibility.

Why Should We Care About XML?

Most of the major information technology leaders such as Microsoft, Oracle, IBM and Sun are all focusing on application and information solutions that will offer native XML support. In fact it's hard to find a major player who hasn't declared its intentions to be part of the front line to support the XML standard.

In the same way as the Internet has permeated our businesses and homes, so too will the effects of XML. There are a series of initiatives underway with respect to XML, and perhaps most notable of these is the establishment of Legal XML (www.legalxml.org). Founded in November 1998, Legal XML is a nonprofit group comprised of volunteer members from private industry, nonprofit organizations, government and academia. The mission of Legal XML is to develop open, nonproprietary technical standards for legal documents and related applications.

What Does it Mean to the Legal Industry?

The basic content and page-oriented format of a legal transcript will probably always be preserved, but page breaks are to a large degree an arbitrary definition based on the need to review information from paper.

Of course there is value in the consistency and referential nature of page breaks, but the opportunity of online

Imagine having a single legal transcript that a lawyer or judge can view online and with a simple click review any document in various presentations, such as the traditional page, paragraph, question and answer, cross examination, or simply allow a view of the witness list and exhibit list.

review, analysis and annotation of transcripts has become widespread. Tools such as LiveNote, Summation, e-transcript and Ringtail's Web-based CourtBook and FullText products have changed the way many lawyers and judges review textual information.

The next generation of litigation-support tools will probably be browser based and use XML standards to allow transcripts, court proceedings, contracts and other legal materials to be reviewed and analyzed in powerful new ways.

Imagine having a single legal transcript that a lawyer or judge can view online and with a simple click review any document in various presentations, such as the traditional page, paragraph, question and answer, cross examination, or simply allow a view of the witness list and exhibit list.

In other words, by separating structure and content from presentation, the same XML source document can be written once and then displayed in a variety of ways on a computer monitor.

An Example of XML Transcript Tags

Here is an example of what a transcript standard could look like in XML. Those words contained within the brackets <> are tags that would be inserted into the transcript by the court reporter or input automatically by the reporter's CAT software. The tags would then function as search tools, in that judges, attorneys and court administrators could use the tags to find the information they need. By helping to set a standard in legal XML, court reporters would be able to use an agreed-upon set of tags so the information in their transcripts would be searchable across a variety of litigation-support platforms.

At the document level:

<Transcript Provider> including name, address, phone, etc.
<Court>
<Jurisdiction> e.g., Civil Law Division
<Location> e.g., Washington, D.C.
<Date>
<Day of Hearing>
<Date continued from>
<Judicial Officer> There can be more than one.
<Case No.>
<Event> e.g., hearing, voir dire, etc.
<Type of Document> e.g., transcript of hearing, sentence, etc.
<Party Name> including title, surname, given name, etc. (one for each party)
<Party Type> e.g., plaintiff, respondent, etc. (one for each party)
<Legal Representative> including name, address, contact details (one for each party)

<Representative Type> one for each party
<Short Case Name>
<Start Time>
<Reporter>
<End Time>

At the section level:

<Witness No. >
<Witness Name> including title, surname, given name, etc.
<Witness Status> e.g., sworn, re-affirmed, etc.
<Mode Status> e.g., examination-in-chief, cross-examination, etc.
<Mode Start Time>
<Advocate> name of legal practitioner or advocate
<Continuing> flag if relevant
<Mode End> e.g., the witness withdrew
<Mode End Time>
<Question> tag for start of question
<Answer> tag for end of a question

At the page level:

<Page No. >
<Header>
<Footer>

At the line level:

<Line No. >
<Exhibit No. >
<Exhibit Description>
<MFI No. > MFI = Marked for Identification
<MFI Description>
<Citation>

What About the Court Reporter's Role?

In a sense, legal XML takes Reporter Electronic Data Interchange a step further. REDI uses the information captured by the court reporter to extract specific data for administrative purposes. This enhances the process of case flow and streamlines the inputting of data currently used in court management without additional hardware costs. The technology provides a cost savings by



allowing information to be efficiently accessed in greater detail.

Legal XML expands the amount of information that can be taken from the reporter's digital transcript and offers greater flexibility in how that transcript is searched and used. Legal XML is based on tags that

would be manually input by the court reporter or automatically included by the reporter's CAT software (some

possible tags are listed in the sidebar on this page). Attorneys, judges and court administrators would then use litigation-support software to search through the transcript with these XML tags. For example, a judge may use the tags to study all the citations mentioned in a trial, the attorney to examine all the testimony of a witness, and a court administrator or court clerk to pull out the case number, the day of the proceeding and whether the activity for that day involved voir dire or a hearing. By establishing a standard for legal XML, court reporters would then be using the same tags, so

no matter the jurisdiction or location, anyone could use the tags to search the transcript or other legal documents.

As the co-chairman of the Legal XML Transcripts Standards Workgroup, I recently met with NCRA Executive Director Mark Golden regarding the Association's involvement in creating legal XML standards. One of the key objectives for the Transcripts Workgroup will be to encourage all stakeholders to contribute to the new standard, and the first stage for me was to introduce Legal XML and the mission of the organization to NCRA and also to encourage NCRA's leaders and members to become involved in the development of legal XML standards for transcripts.

Obviously, one of the key contributors to any transcript standard should be the court reporter. And, with NCRA's recent decision to aid in the development of legal XML standards, we'll be in a much better position to develop requirements that meet the needs of court reporters and others who will make use of XML.

To join the Legal XML discussion or development group, visit www.legal.xml.org. You are free to join both lists. However, discussions on the developer's list may be too technical for someone who simply wants to stay abreast of Legal XML events.

Finally, a few important points about the Legal XML standards: They are not mandatory or obligatory. They are not fixed, but rather the intention is to create a flexible standard that may evolve with time. They should not add unnecessarily to the task of producing transcripts. They should be for everyone's benefit. They should not affect the ability to produce, review and print transcripts as we use them today. ■

An Introduction to XML

1. XML is a method for putting structured data in a text file.

For "structured data" in the legal environment think of transcripts, court proceedings, contracts, electronic filings and other legal documents. Programs that produce such data often also store it on disk, for which they can use either a binary or text format. The latter allows you, if necessary, to look at the data without the program that produced it. XML is a set of rules, guidelines, conventions — whatever you want to call them — for designing text formats for such data in a way that produces files that are easy to generate and read (by a computer), that are unambiguous, and that avoid common pitfalls, such as lack of extensibility, lack of support for internationalization/localization and platform dependency.

2. XML looks a bit like HTML but isn't HTML.

Like HTML (HyperText Markup Language), XML makes use of tags (words bracketed by < and >) and attributes (of the form name = "value"). But while HTML specifies what each tag and attribute means (and often how the text between them will look in a Web browser), XML uses the tags only to delimit pieces of data and leaves the interpretation of the data completely to the application that reads it. In other words, if you see "<p>" in an XML file, don't assume it is a paragraph. Depending on the context, it may be a price, a parameter, a person or perhaps something else.

3. XML is text, but isn't meant to be read.

XML files are text files but they are not meant to be read. They are text files because that allows experts such as court reporters to more easily present textual information with rich formatting extensions for communication and presentation.

4. XML is a family of technologies.

There is XML 1.0, the specification that defines what tags and attributes are, but around XML 1.0 there is a growing set of optional modules that provide sets of tags and attributes or guidelines for specific tasks.

5. XML is verbose, but that is not a problem.

Since XML is a text format, and it uses tags to delimit the data, XML files are nearly always larger than comparable binary formats. The advantages of a text format are evident (see 3 above), and the disadvantages can easily be solved at a different level. Disk space isn't as expensive as it used to be, and zip programs can compress files.

6. XML is new, but not that new.

Development of XML started in 1996. But in fact the technology isn't very new. Before XML there was SGML, which was developed in the early '80s and is widely used for large documentation projects. And, of course, HTML, whose development started in 1990. The designers of XML simply took the best parts of SGML, guided by the experience with HTML, and produced something that is no less powerful than SGML, but vastly more regular and simpler to use.

7. XML is license free, platform independent and well supported.

By choosing XML as the basis for some project, you buy into a large and growing community of tools and engineers experienced in the technology. And since XML is license free, you can build your own software around it without paying anybody anything. The large and growing support means that you are also not tied to a single vendor. XML isn't always the best solution, but it is always worth considering.