

**GENERAL  
UN/CEFACT/1999/CRP.16  
UN/CEFACT/TMWG/N089/Rev.1  
18 March 1999**

**Techniques & Methodologies Working Group ( TMWG )**

**TMWG POSITION ON XML  
RECOMMENDATIONS TO UN/CEFACT**

**SOURCE: TMWG/XML TASK GROUP  
STATUS: APPROVED  
ACTION: FOR UN/CEFACT REVIEW**

## **Abstract (provided by UN/CEFACT Steering Group)**

UN/CEFACT has been carefully monitoring internet technology developments and in particular the evolution of XML. In order to put this in context with other developments in Electronic Business, the UN/CEFACT Steering Group (CSG) requested the TMWG to review the XML developments and offer guidance. A summary of that guidance is:

- XML, the eXtensible Markup Language, is an official recommendation by the World-wide Web Consortium (W3C) and is seen to offer opportunities in electronic commerce.
- A mature XML supported by appropriate standardisation and development tools has an importance place in electronic commerce.
- XML technology by itself does not solve all of the electronic data interchange related problems, but it can potentially help within the context of Internet forms related processing.
- Today, electronic forms or Web pages used for electronic commerce do not operate in a consistent manner.
- UN/CEFACT believes that the starting point for consistency for commercial business forms on the WEB should be the use of the UN Layout Key (UNLK).
- XML is best described as a somewhat loose standard; being 'extensible' means that anyone can devise their own data-identifying tags. Many different 'language' agreements can spring up between any two or more consenting parties!
- Standardisation is absolutely required within the context of Business to Business electronic commerce.
- There are issues involving the mass, uncontrolled proliferation of Document Type Definitions (DTD) specific to identical business processes, and the random selection of tag names without guaranteeing uniqueness. This uncontrolled proliferation leads to lack of interoperability when it comes to data integration within an application.
- Without a standardisation process, there will be a proliferation of tag names and DTDs for the same processes, creating severe interoperability issues.
- One aspect of XML that has attracted a lot of interest is the ability to attach semantic tags to structured data, giving it a basic capability to represent

structured business documents. So XML can perform much the same function as UN/EDIFACT syntax but data integration into applications is still required.

- More work is needed to evaluate the long term impact of XML since the technology (XML 1.0 specification and the associated tools) itself is still evolving.
- We believe the primary focus of UN/CEFACT in XML related work should be centred around the use of the UNLK.
- Any future edi development by EWG should be focussed on harmonisation and identification of the core set of data to support EDI transactions. Message Implementation Guidelines should be the source to create XML DTDs, not the EDI Directories.
- UN/CEFACT, in order to influence this process to the benefit of international trade simplification, needs to be proactively involved in the effort to creation a single global XML repository. There is a vast depth of knowledgeable resources within UN/CEFACT that could be encouraged to work on this effort.

## **Introduction**

This document is divided into two main parts;

- Overview of XML – Provides a high level view of XML's strengths, weaknesses, opportunities, and threats from a UN/CEFACT perspective
- Technical Annexe – Provides a list of XML related terms, a high-level purpose statement, TMWG position on the current state of the technology, and reference including some URLs to use for further research.

## **Historical Positioning of XML**

A markup language identifies individual pieces of information by embedding tag names in the text. XML, the eXtensible Markup Language, is receiving a lot of publicity by technical publications, especially as it comes from the printing/presentation background. XML is an official recommendation by the Worldwide Web Consortium ( W3C ). The same publications are now just beginning to realise the issues that have been raised by TMWG, and are classifying its deployment as risky due to the immaturity of the technology and its proliferation without a standardisation process.

XML came into existence because HTML, the HyperText Markup Language used for 'publishing' Internet Web pages, was too limited in its capabilities. Both XML and HTML are derivatives of SGML, the Standard Generalized Markup Language, which is an ISO standard. Said to be over-complicated and too feature-rich, SGML's main usage is in markup for information retrieval from major military specifications, and for printing/publishing layout descriptions. SGML dates back to GML created in 1969 by Charles Goldfarb in the IBM laboratories.

XML is best described as a somewhat loose standard; being 'extensible' means that anyone can devise their own data-identifying tags. Many different 'language' agreements can spring up between any two or more consenting parties!

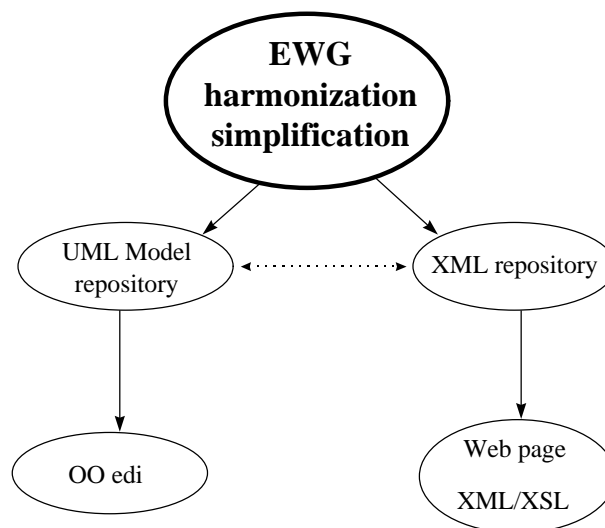
XML is seen as POP - People Oriented Presentation (i.e. people-internet-computer), while classical EDI is MOM Message Oriented Middleware (i.e. computer-computer any-which-way).

It must be understood that a mature XML supported by appropriate standardisation and development tools has an importance place in electronic commerce. TMWG is not dismissing XML technology, but we see that XML is really a “family” of complementary technologies ( the attached Annexe shows some of these complementary and often overlapping technologies ). TMWG views that due to this “hype”, the XML gurus have a case of classic “scope creep”. By trying to solve *too many business problems* using XML, they have created *too much technology*. And now it appears that W3C has *too much work* on their plates, which will take *too much time* to progress to the recommendation status and physically implement.

### ***Impact of XML on UN/CEFACT***

We are currently seeing the emergence of XML as a potential replacement for HTML. One aspect of XML that has attracted a lot of interest is the ability to attach semantic tags to structured data, giving it a *basic* capability to represent structured business documents. So XML can perform much the same function as UN/EDIFACT or ANSI ASC X12 *syntax*. This is being termed as XML/EDI. However XML/EDI *as such* is no closer to being a complete data interchange standard than UN/EDIFACT or ANSI ASC X12 *syntax*. This fact tends to be overlooked in some of the more optimistic statements on the future of XML/EDI. However, we are now observing several initiatives that suggest using a global repository of XML semantic tags and document type definitions (DTDs). This is a credible step towards standardisation of XML/EDI interchanges.

TMWG believes the *primary* focus of future edi development by EWG should be harmonisation and the identification of the core set of data to support EDI transactions. The following diagram shows how the core data would then feed into both UML modelling and an XML repository.



The process illustrated above could reduce unnecessary incompatibilities between OO-edi and Simpl-EDI's use of XML for web based EDI, providing a helpful symbiosis between the two lines of development. However there is a danger that the momentum of XML could attract an inappropriate amount of EWG resources, whereas they could be working with BPAWG on modelling the business process. EWG involvement in the XML/EDI development should be kept within well-defined limits.

### ***TMWG Position on eXtensible Markup Language ( XML )***

This document discusses some of the issues and TMWG recommendations for using XML within an Electronic Commerce environment and the value added contribution that UN/CEFACT should make to this work effort:

- 1) The creation of a standard look and feel for UN Layout Key web pages,
- 2) The growing acceptance of XML as an interface format by commercial applications,

- 3) The need for standard tag names and DTDs to allow interoperability,
- 4) The desire to create algorithms to auto-generate XML DTDs directly from the EDI Directories,
- 5) The development of a simple XML repository, which could have a bi-directional feed to and from the UML repository,
- 6) The use of XML to format object requests and responses within an OO-edi environment to improve the interoperability between distributed object computing environments ( DCOM, CORBA, Java RMI, etc. ).

It is believed that XML has some fundamental opportunities in electronic commerce, which will be described below. However, we believe more work is needed to evaluate the long term impact since the technology ( XML 1.0 specification and the associated tools ) itself is immature. Since data integration into applications is still required, XML technology by itself does not solve all of the electronic data interchange related problems, but it can potentially help within the context of Internet forms related processing.

There are issues involving the mass, uncontrolled proliferation of Document Type Definitions ( DTD ) specific to identical business processes, and the random selection of tag names without guaranteeing uniqueness. This uncontrolled proliferation leads to lack of interoperability when it comes to data integration within an application.

### **1) Standardized Web Pages for UN Layout Key Implementation**

Today, electronic forms or “Web pages” used for electronic commerce do not operate in a consistent manner. In addition, most companies intend to add advertisements and flashy graphics to make the WWW experience pleasant. This is not required in a trade facilitation process, and hence various parts of the web page can be standardised using XML and its complimentary technology called eXensible Stylesheet Language ( XSL ).

We believe creating standard “form objects”<sup>1</sup> is a specific value added service that UN/CEFACT should provide. The starting point should be the UN Layout Key as previously recommended to the CSG, and the SIMAC WWW work should continue and take into account standard form objects. This would involve reuse of the UN Layout Key form objects based on the context of the transaction; e.g., whether it is the “purchase order” or “export cargo shipping instruction”. This affects any XML repository development effort, specifically if an ENTITY from one electronic form “feeds” another ENTITY with a different name in another electronic form with identical data. This is exactly the type of harmonisation effort that needs to continue.

*TMWG Recommendation:-* The harmonisation and simplification work effort must continue to identify the core data.

*TMWG Recommendation:-* Standard Web form objects need to be developed for UN Layout Key web pages.

---

<sup>1</sup> . A form object consists of XSL text and an embedded scripting language ( Javascript, Vbscript, ECMAScript, or similar ) to read the XML data stream and add presentation specific information such as font size and location ( x-y coordinates ) on the form .

## **2) Software Providers' announcements of XML-aware Applications**

Several software companies such as Oracle, SAP, and Intuit have announced their support of XML as one type of interchange file format. This means that these applications can export and import XML files. However, the SME, and purchasers of this software in general, should be forewarned that most of these interchange files have XML *tag names* that are identical to the application's *field names* and are not in accordance to any type of standard. Therefore, transmitted XML documents that do not conform would need to be preprocessed by mapping the transmitted tag names to the tag names needed by the application.

Some vendors envision the ability to import DTDs into the application and have the application provide a mapping facility to map the XML tag names to the application field names. This is a very unlikely solution for the SME, since it has been stated that many SMEs do not like to do ANY type of mapping. Furthermore, if there are a large number of XML DTDs, mapping will be a time consuming and cost prohibitive task for the SME.

Other companies have "data integration" tools that ease the integration into commercial applications, however, these tools are more expensive and not attractive to most SMEs. This type of tool is different than an XML-aware application, and can integrate XML and almost any type of format including comma-separated-value files.

If the SME does not have an XML aware application, there is a need to transform the XML stream into an application ready format such as a comma separated value file or use data access standards such as ODBC.

*TMWG Recommendation:-* UN/CEFACT should encourage W3C to develop a simple XML Repository as soon as possible to stop the proliferation of XML tag names and DTDs.

## **3) Standard tag names and DTDs ensure interoperability**

The most significant benefit of XML is that the Document Type Definition provides the specification for the content of the XML Document. Since the XML Document references the DTD specification within the file itself, more sophisticated "validating" XML parsers can read the XML file, and check to see if the XML document conforms to the specification.

Another benefit is that the tag names in the DTD can "expose the semantics" of the data in the XML document. Therefore it is not necessary to have an implementation guideline or EDI directory available to understand the content of the XML document.

The downside is that without a standardisation process, there will be a proliferation of tag names and DTDs for the same processes, creating severe interoperability issues. This is already evident in the procurement arena, where providers such as Ariba and

CommerceOne are putting out competing XML DTDs. See the following article: <http://www.internetwk.com/news0299/news020999-6.htm>. Standardisation is absolutely required within the context of *Business to Business* electronic commerce.

*TMWG Recommendation:-* UN/CEFACT should encourage CommerceNet, the XML/EDI group, W3C, ANSI ASC X12C/TG3 and member software companies to adopt standardised tag names and DTDs for XML interchange file formats to help the SME integrate data.

#### **4) Recasting existing EDI Directories into XML**

There have been many recommendations from several organisations to develop algorithms to auto-generate XML DTDs directly from EDI Directories. TMWG believes that this should not occur, but UN/CEFACT resources should focus on the simplification and harmonisation effort instead. Message Implementation Guidelines should be the source to create XML DTDs, not the EDI Directories. The fundamental reason is that tag names would be created that would never be used in a real life scenario, basically a superset of tag names. The analysis of industry and company MIGs would provide the core data ( or subset ) that is really used.

In addition, as new organisations and individual companies continue the proliferation of “de facto standards”, any type of output from UN/CEFACT would only ADD to the confusion in the marketplace by providing only another XML solution for the SME to choose from. These companies creating these “standards” believe that 1) that the standards process is too slow, and 2) they can live without standards. It seems illogical to promote interoperability without the use of standards, especially as most of the W3C technologies are still at the “note” status, with limited resources to promote them to official “recommendations”.

*TMWG Recommendation:-* UN/CEFACT should NOT recast UN/EDIFACT messages into XML based on algorithms run against EDI Directories.

*TMWG Recommendation:-* UN/CEFACT should limit its resources’ use of XML to Web forms and distributed object messages.

#### **5) Repositories – the enabling technology for the future of EDI**

The XML proliferation is ongoing and cannot be controlled unless there is a competent agency to provide a registry of tag names and DTDs. It appears to TMWG that an organisation such as W3C, who has a large technical resource pool, would be the most acceptable to the industry as the custodian of an XML Repository. UN/CEFACT needs to be proactively involved in order to influence this process to the benefit of international trade simplification.

Information based on the harmonisation and simplification ( Simpl-EDI ) work effort should be the first to be loaded into this repository in order to preserve the investments to-date. We should not simply load everything into the repository, but ONLY the data that are actually being used. We view that the BSR could be used only if the existing



content is flushed and we start all over beginning with the harmonisation and simplification work effort.

There is a vast depth of knowledgeable resources within UN/CEFACT that could be encouraged to work on this effort. Additional harmonisation and simplification efforts would be directly beneficial to any UML modeling effort for OO-edi.

It is likely that DTD Editors will eventually output XML Metadata Interchange ( XMI ) files that will import the tag names and the nesting relationships into the UML model repository. The XML DTD could be transformed directly to a UML class diagram containing one or more classes. This content could be incorporated within a larger model that incorporates dynamic behaviour. Traditional SGML DTDs could be loaded into the repository as well. This would provide a relatively painless mechanism to incorporate short-term XML developments into the long term OO-edi standards development.

Conversely, the UML Repository will be able to export XMI files for loading information into other repositories. Therefore, since the XMI file contains information about the metadata, the XML Repository could import the tag names, which then could be used in Web forms.

It is viewed that these repositories could potentially converge into one, however, there may not be a need to do this if synchronisation between repositories is developed. For example, if a XML DTD is created and loaded into the XML Repository, it could kick off a process in the UML Repository to review this within the context of existing or new models. As the models are developed or refined, the XML Repository could be refreshed and versioned to reflect the incorporation in the UML models.

*TMWG Recommendation:-* UN/CEFACT should encourage W3C to develop a simple XML Repository as soon as possible to stop the proliferation of XML tag names and DTDs.

## **6) The Future - XML Formatted Object Messages**

TMWG also believes that the best and most proper use of XML is in the context of marshalling object requests and responses as XML formatted objects within a distributed object environment, which is a direct output from the UML repository. This involves formatting a message request as an XML document which is sent over the network to an object. The object then invokes the requested method, and returns back its response in XML format as well.

In addition, a client object could request the metadata about a server object, so that the client knows how to format the message requests. This opens the door to dynamically reading interfaces and learning how to interact with an object on the fly which is currently very difficult to implement in environments such as the CORBA Dynamic Interface Invocation.

W3C has a very good note entitled “WebBroker: Distributed Object Communication on the Web” ( <http://www.w3.org/TR/1998/NOTE-webbroker-19980511> ) which discusses exactly what TMWG has been promoting for a significant period of time. The idea is that in today environment object request brokers such as DCOM, CORBA and Enterprise Java Beans cannot communicate without expensive “bridges”, and XML formatted object messages would provide interoperability between these environments. This solution would utilise DOM to populate business objects, which drastically simplifies byte order processing in a distributed object environment.

*TMWG Recommendation:-* UN/CEFACT should encourage W3C to accelerate work on the W3C WebBroker Note and progress it to a W3C Recommendation as soon as possible.

*TMWG Recommendation:-* UN/CEFACT should limit its resources’ use of XML to Web forms and distributed object messages.

### ***Summary of TMWG Recommendations for UN/CEFACT Resource Allocation***

- Harmonisation and simplification work effort must continue to identify the core data.
- UN/CEFACT should encourage W3C to develop a simple XML Repository as soon as possible to stop the proliferation of XML tag names and DTDs.
- UN/CEFACT should encourage CommerceNet, the XML/EDI group, W3C, ANSI ASC X12C/TG3 and member software companies to adopt standardised tag names and DTDs for XML interchange file formats to help the SME integrate data.
- Standard Web form objects need to be developed for UN Layout Key web pages.
- UN/CEFACT should limit its resources’ use of XML to Web forms and distributed object messages.
- UN/CEFACT should NOT recast UN/EDIFACT messages into XML based on algorithms run against EDI Directories.
- UN/CEFACT should encourage W3C to accelerate work on the W3C WebBroker Note and progress it to a W3C Recommendation as soon as possible.

### **Technical Annexe to TMWG Position Paper on XML**

This technical annexe is appended to provide readers of the Position Paper with additional detail should they wish to dip deeper into items mentioned in the main paper. Alternatively, they may wish to pass on the whole paper to colleagues with a deeper technical interest as a resource for their review and thought process.

XML Technology	Purpose	TMWG Position	Reference Materials
XML	<p>XML, eXtensible Markup Language, is similar to HTML, the HyperText Markup Language used for Internet Web pages. HTML and XML are derivatives of the generic markup language called SGML, used for for military specifications and publishing descriptions.</p> <p>(a) In common with SGML and HTML (also many sectoral derivatives such as OFC, OFX, Integrion Gold, BIPS et al) XML effectively tags every field (element of data) with a keyword or tag name according to a particular style. Such elements are also gathered together and the sets of 'grouped together' elements can also be given tag names.</p> <p>(b) The 'extensible' part of XML is emphasised as giving flexibility. Elements which are not in XML's vocabulary can readily be invented by the users, as it is possible to include the definition/specification of new elements within the message itself.</p> <p>Status: A W3C Recommendation, not a NOTE anymore.</p>	<p>(a) The essential difference with EDIFACT, ANSI X12 is the tagging of every field, whereas EDIFACT tags small 'record layouts' each of which contains several 'grouped together' elements. The nett effect is for the field-tagged markup languages to be somewhat more verbose, i.e. to use more characters to say the same thing!</p> <p>(b) It is not yet clear how this mechanism will actually function in a true application-to-application interchange. (One can easily envisage this in an application-to-presentation process such as printing or in offering services via a web-site. This betrays its concept origins in printing/publication) Nor is it clear how new invented elements get into the mainstream language for re-use by others and to prevent many new tags being invented for the same thing.</p>	<p>XML in general  <a href="http://www.w3.org">http://www.w3.org</a>  XML 1.0 Specification  <a href="http://www.w3.org/TR/REC-xml">http://www.w3.org/TR/REC-xml</a></p>
DTD	<p>The principle feature of a valid XML document is that it has, and conforms to, a DTD - document type definition. This feature, inherited from SGML, contains declarations that specify the overall structure of a document and the acceptable types of data content values.</p>	<p>There are some short-comings omitted, such as data typing, see XML-Schema DCD later.</p>	<p>ISBN 1-861001-52-5  "XML Applications"  XML in general  <a href="http://www.w3.org">http://www.w3.org</a></p>
XLL	<p>XLL was the one standard for XML linking, now separated into XLINK and XPOINTER. The term is still used collectively for both.</p>	<p>This has been superseded.</p>	<p>ISBN 1-861001-52-5  "XML Applications"</p>
XLINK	<p>XLINK is for linking between documents. A <i>link</i>, as the term is used here, is a relationship which is asserted to exist between two or more data objects or portions of</p>	<p>"...won't this be cool when they implement it?"  Unfortunately, as with much of XML,</p>	<p>ISBN 1-861001-52-5  "XML Applications"  <a href="http://www.w3.org/TR/WD-xlink">http://www.w3.org/TR/WD-xlink</a></p>

XML Technology	Purpose	TMWG Position	Reference Materials
	<p>data objects. XLINK allows you to:-</p> <ul style="list-style-type: none"> <li>• create your own link elements</li> <li>• use any element as a linking element</li> <li>• create bi-directional links with one-to-many and many-to-one relationships</li> <li>• specify traversal behaviour - i.e. how users get between links</li> <li>• create link databases to specify and manage links outside of the documents to which they apply (for ease of updating)</li> <li>• aggregate links</li> </ul>	<p>there is no implementation of the W3C standard as yet.</p> <p>Used with XPointer for Repository lookup, even though this is not fully implementable at this time.</p>	
XPOINTER	<p>XPointers are used in conjunction with URIs to specify part of a document. They give the means for pointing into documents</p> <p>XPointer consists of a series of locations separated by a period, each of which specifies a location in the document, usually in terms of the previous location. There are several forms of absolute location in a document, the root, the origin, and the most useful an element id.</p>	<p>"As we do not even have an XML compliant browser, we certainly don't have an XLink or an Xpointer compliant one"</p> <p>Used as a reference to Repository entries.</p>	<p>ISBN 1-861001-52-5 "XML Applications" <a href="http://www.w3.org/TR/WD-xptr">http://www.w3.org/TR/WD-xptr</a></p>
XSL	<p>XSL allows the user to specify how parts of a document should look (e.g. font, size, colour) and is extensible in that it allows the user to create new formatting tags and properties.</p> <p>XSL enables formatting information to be associated with elements in the source document to produce formatted output. The formatted output is created by formatting a tree of <i>flow objects</i>. A flow object has a class, which represents a kind of formatting task. A flow object also has a set of named characteristics, which further specify the formatting. A core set of flow object classes is outlined later in this document.</p> <p>The association of elements in the source tree to flow objects is through <i>construction rules</i>. The construction rules contain a <i>pattern</i> to identify specific elements in the source tree, and an <i>action</i> to specify a resulting sub-</p>	<p>TMWG see this as the necessary tool for XML-based Web pages, even though tools for Rapid Application Development(RAD) are not available.</p> <p>August '98 XML guru comment "Look at this neat stuff. Won't it be great when we have a tool?"</p> <p>Megginson "...surprised to see Microsoft promise an implementation for XSL, which is still relatively immature..."</p> <p>Bray "...worries that Microsoft could create forward-compatibility problems if the implementations in IE5 differs</p>	<p>ISBN 1-861001-52-5 "XML Applications" Named quotes from Internet World Oct 19 1998 <a href="http://www.w3.org/TR/NOTE-XSL.html">http://www.w3.org/TR/NOTE-XSL.html</a> <a href="http://www.w3.org/Style/XSL">http://www.w3.org/Style/XSL</a></p>

XML Technology	Purpose	TMWG Position	Reference Materials
	<p>tree of flow objects. The stylesheet processor recursively processes source elements to produce a complete flow object tree.</p> <p>In addition to construction rules XSL also supports <i>style rules</i> which allow the merging of characteristics. While only one construction rule can be invoked for a particular source element, <i>all</i> applicable style rules are invoked, allowing per-characteristic merging.</p> <p>The capabilities provided by XSL allow:</p> <ul style="list-style-type: none"> <li>• formatting of source elements based on ancestry/descendency, position, and uniqueness</li> <li>• the creation of formatting constructs including generated text and graphics</li> <li>• the definition of reusable formatting macros</li> <li>• writing-direction independent stylesheets</li> <li>• extensible set of formatting objects</li> </ul>	<p>from eventual XSL specifications..." and "...I'm kind of uneasy with this massive charge in the direction of XSL..."</p>	
XMI	<p>Proposal to use XML for the transfer syntax and interchange format, including:-</p> <ul style="list-style-type: none"> <li>- DTDs to enable transfer and verification of :- <ul style="list-style-type: none"> <li>• UML based models</li> <li>• MOF based metamodels</li> </ul> </li> <li>- Specification of a precise MOF to XML mapping to enable:- <ul style="list-style-type: none"> <li>• Use of OCL to specify stream production rules</li> <li>• Interchange of any MOF-based metamodel</li> <li>• Automatic generation of DTDs</li> </ul> </li> </ul> <p>Potential XMI usage scenarios are:-</p> <ul style="list-style-type: none"> <li>- Interchange of UML and other MOF compliant models between:- <ul style="list-style-type: none"> <li>• modeling and design tools, generators</li> <li>• tools and repositories</li> <li>• repositories</li> </ul> </li> </ul>	<p>This is the essential file format required by BOTH the UML and XML repositories to allow synchronisation between the tools.</p> <ul style="list-style-type: none"> <li>• Older CASE tools need this support</li> <li>• Additional DTDs need to be developed for IDEF0, IDEF-1X, and other legacy modeling techniques.</li> </ul>	<p>Presentation: Proposal to OA &amp; DTF RFP - 3 OMG TC Meeting June 10 '98</p>
DOM	<p>The Level 1 DOM (Document Object Model) specifies how to access and manipulate an HTML or XML document, and compliant software is beginning to emerge. Level 2 is being worked on at the moment (Oct</p>	<p>One of the big questions is how to actually access, manipulate and use the information that is stored in XML. DOM is potentially one of the most</p>	<p>ISBN 1-861001-52-5 "XML Applications" ISBN 0-13-616822-1 "Designing XML Internet Applications"</p>

XML Technology	Purpose	TMWG Position	Reference Materials
	<p>'98) and will include manipulation of the document's style information.</p> <p>The W3C wg is developing a platform- and language-neutral program interface that will allow programs and scripts to access every element in a document and update the content and structure of the documents in a standard way.</p> <p>The DOM Level One (Core) specification defines a set of object definitions that are sufficient to represent a document instance (the objects that occur within the document itself). This specification extends the DOM Level One (Core) specification such that document type definitions, entities, CDATA marked sections, and conditional sections can also be represented.</p> <p>The DOM objects and interfaces are designed to be:</p> <ul style="list-style-type: none"> <li>• sufficient for representing the content of parsed HTML and XML documents without loss of [significant] information. The supported HTML version is 4.0; the supported XML version is 1.0.</li> <li>• sufficient to construct an entirely new document instance programmatically that is identical to the parsed form of a given HTML or XML document. This means that it has sufficient constructive power to build any useful document object hierarchy, and that an implementation could be written such that the external document parser merely calls the methods specified in the level one specification to build the object hierarchy.</li> <li>• the foundation for the rest of the document object model levels, which means it must be simple, flexible, and extensible.</li> <li>• thread-safe: The operations supported by the DOM will not corrupt the document object or return corrupted state (as far as this API is concerned). Higher level consistency support</li> </ul>	<p>important standards in XML because it gives implementers a common vocabulary to use in manipulating an XML document.</p> <p>But it's still evolving.</p> <p>Bray "...the usefulness of XML support in IE5 will depend on the depth of Microsoft's DOM support..." and "...waiting for Microsoft to clarify whether IE5's DOM interfaces will allow programmatic access to actual untranslated XML code..."</p>	<p>XML'98 Conference, Chicago</p> <p>Quotes from Internet World Oct 19 1998</p> <p><a href="http://www.w3.org/TR/WD-DOM">http://www.w3.org/TR/WD-DOM</a></p> <p><a href="http://www.w3.org/TR/REC-DOM-Level-1">http://www.w3.org/TR/REC-DOM-Level-1</a></p>

XML Technology	Purpose	TMWG Position	Reference Materials
	mechanisms such as explicit locks or transactions are outside of the scope of the level one specification. For level one of the DOM, the assumption is that only one thread operates on the document at a time.		
XML SCHEMA	<p>Schemas define the characteristics of classes of objects. In XML a schema is simply a definition of the way in which a document is marked up. The two XML schema proposals that exist are still at the Note stage:-</p> <ul style="list-style-type: none"> <li>• XML-Data</li> <li>• Document Content Description.</li> </ul> <p>Another schema being proposed is:-</p> <ul style="list-style-type: none"> <li>• Schema for Object-Oriented XML (see SOX)</li> </ul>	<p>XML Schemas have arisen because of weaknesses in the traditional DTDs.</p> <p>The transformation between XML Schema to DTD does suggest that information in the XML Schema is lost.</p>	<p>ISBN 1-861001-52-5 "XML Applications"</p> <p>XML in general <a href="http://www.w3.org">http://www.w3.org</a></p>
DCD	<p>DCD is a Document Content Description. DCD semantics are a superset of those provided by XML DTDs.</p> <p>DCD takes the approach of describing constraints that apply to the structure and contents of XML documents. Apart from declaring what element types, attributes and values can be used in a document instance, a DCD constrains the structure and contents of a document that uses it. The DCD specification permits attributes and elements that are not from the same DCD; others can be referenced by namespaces</p>		<p>ISBN 1-861001-52-5 "XML Applications" <a href="http://www.w3.org/TR/NOTE-dcd">http://www.w3.org/TR/NOTE-dcd</a></p>
XML-Data	<p>DTDs do not explain data well; there is no data typing, all data is a string.</p> <p>XML-Data provides a greater range of popular datatypes than the traditional DTD does.</p>		<p>ISBN 1-861001-52-5 "XML Applications" <a href="http://www.w3.org/TR/1998/NOTE-XML-data/">http://www.w3.org/TR/1998/NOTE-XML-data/</a></p>



XML Technology	Purpose	TMWG Position	Reference Materials
SOX	<p>SOX is a Schema for Object Oriented XML. It is a proposed facility for defining the structure, content and semantics of XML documents, to enable XML validation and higher levels of automated content checking. It provides an alternative to XML DTDs , giving basic intrinsic datatypes, an extensible datatyping mechanism, content models and attribute interface inheritance, a namespace mechanism, and embedded documentation.</p> <p>SOX should:</p> <ul style="list-style-type: none"> <li>• enable software mapping from SOX documents into data structures in relational databases, common programming languages and interface definition languages(e.g. Java, IDL, COM, C and C++)</li> <li>• enable re-use at the document design and application programming levels</li> <li>• be able to express domain abstractions and common relationships among them directly and explicitly (e.g. subtype/supertype etc)</li> <li>• support the generation of common application components directly from SOX documents</li> </ul>	<p>Quote from the book, "...we will not be looking at that, although there is already partial implementation of it in some of the XML freeware that is available..."</p>	<p>ISBN 1-861001-52-5 "XML Applications" <a href="http://www.w3.org/TR/NOTE-SOX">http://www.w3.org/TR/NOTE-SOX</a></p>
WIDL	<p>The Web Interface Definition Language (WIDL) is an XML vocabulary designed to enable application-to-application communications. WIDL enables an application to express, in an abstract sense, the services it offers, thus enabling it to communicate with other applications regardless of differences in the underlying programming language, the XML DTD, or even the ability to send receive XML data.</p> <p>It was developed for the WebMethods web automation toolkit and has been proposed as a standard to W3C.</p> <p>WIDL describes resources of WWW in order to enable automation of all interactions with HTML/XML</p>	<p>WebMethods B2B Integration Server is the first deployed and proven application integration server based on XML technology. It was designed to enable organisations with legacy applications and data structures to compete in the electronic landscape. To accomplish this, WIDL and B2B developer were developed. (see link in reference material column)</p> <ul style="list-style-type: none"> <li>• This takes after OMG's IDL</li> </ul>	<p>ISBN 1-57521-334-6 "Presenting XML" ISBN 0-13-616822-1 "Designing XML Internet Applications" <a href="http://www.software.ibm.com/webserver/appserv/doc/aberdeen/12981072.htm">http://www.software.ibm.com/webserver/appserv/doc/aberdeen/12981072.htm</a></p>

XML Technology	Purpose	TMWG Position	Reference Materials
	documents and forms. It describes the location of services, the input parameters to be submitted to each service and the output parameters to be returned by each service. This last requirement implies a capability to reliably extract specific data elements from web documents and map them to output parameters. WIDL supports object-oriented interfaces.	<p>syntax.</p> <ul style="list-style-type: none"> <li>The WIDL becomes an 'object wrapper' to existing or new applications.</li> </ul>	
WDDX	WDDX stands for Web Distributed Data eXchange, which uses XML to define how different applications can exchange data structures. It is a non-RPC (remote procedure call) based way to move complex data structures over HTTP between application servers. It can be thought of as a very high-level API built on top of the DOM.	<p>From Allaire, makers of the Cold Fusion application server.</p> <p>This competes directly with WebBroker</p>	
WebBroker	This is a set of DTDs that allow the exchange of object messages between HTTP servers. It promises to be the "united distributed computing model for the Web, encompassing both document publishing and distributed software object communication".	<p>There is a very good mapping between COM+CORBA, Java etc data typing.</p> <p>This would eliminate the need for an Object Request Broker, hence the name "WebBroker" for OO-edi.</p> <p>This currently is independent of DOM, and could be independent of HTTP in the future (swap out with different transport)</p>	<a href="http://www.w3.org/TR/1998/NOTE-webbroker">http://www.w3.org/TR/1998/NOTE-webbroker</a>
XQL	XML Query Language. Many people want a standard way to query XML: one proposed solution is XQL, a query language designed specifically for XML documents. This would provide a simple language for specifying nodes, paths and conditions based on a model for the structure of XML documents.	<p>What is needed is the ability to index tag names across multiple documents to speed up searches.</p> <p>The syntax of queries is unlikely to have a concept similar to JOIN.</p>	
		<p>Watch this space...</p> <p>...it's growing...</p>	