CEFACT

**Proposal for a UN Repository for XML Tags
Based on UN/EDIFACT**

SOURCE: Dick Raman, Head of Delegation, EEMA
STATUS: Working document
ACTION: For discussion/approval, fourth session of CEFACT,
14-16 September 1998, agenda item 11

Introduction

This paper proposes that CEFACT establish a Global Repository for the translation of UN/EDIFACT in XML tags. The European Electronic Messaging Association (EEMA) EDI Working Group is prepared to assist the UN/ECE Secretariat in the set up and operation of such a repository, which could be crucial in the advancement of the use of EDI over the Internet.

Background

XML stands for Extensible Markup Language and as such it is a subset of SGML (Standardized General Markup Language), which is also the mother of HTML. In terms of complexity, XML can be placed between SGML and HTML. It uses the SGML syntax to transport components across the network. XML is a flexible, powerful language, geared towards transmitting documents in a structured format across the World Wide Web. The combination of XML and EDI semantic foundations, called XML/EDI, will provide a complete framework where a set of different technologies work together to create a format that is usable by applications as well as humans. These technologies are XML, EDI, templates, agents and repository. These components all work together to create an XML framework for business use:

- The XML tags can actually replace existing EDI segments or data-element identifiers, which produces a somewhat bigger file than an EDI file, but in which all the labels of the data-elements (in other words the descriptions or explanations) are included as well.

- The templates are essentially rules that determine how the XML files should be interpreted. It can define the layout of the file and is supplemented by DTDs (Document Type Definitions) that enable transaction operability.

- The agents can interpret the templates to perform whatever job needs to be done, but they can also interact with the transaction and help the user to create new templates for each specific task.

- The repository is a location where shared Internet directories are stored and where users can manually or automatically look up the meaning and definition of XML/EDI Tags. The repository is in fact the semantic foundation for the business transactions.

What makes XML/EDI different from previous initiatives is that it can use the know-how of business processes, captured in EDI messages, but tries to put it in a Web environment, whereby the same file can be viewed by a user or can be processed by an application.

Why is XML/EDI so useful?

One of the weaknesses of EDI was always that it was quite invisible and it actually only worked well when it was application-to-application. For a long time we have created dummy applications that would allow the user to view or create the contents of an EDI message. This would then be translated and sent out as an EDI message, so that the (usually bigger) partner would receive EDI. With XML/EDI this becomes a lot easier. One of the benefits of XML/EDI is that it can be embedded in the workflow of Electronic Commerce applications.

One can send an EDI message in an XML format to a supplier who can then pick it up and present it through a browser to a person who is in charge of approving incoming orders. That person can add some data to the incoming order, which actually signals his approval - this could for instance be a digital signature. The approved order can then be sent into the supplier's application as a plain EDI file. In this way, as a human or an application creates a message, it can travel through an organization, and can be sent to another organization and switch between humans and applications as well. Every time the XML file will become larger, containing new updated information. As far as the general idea behind XML/EDI is concerned, this is quite exciting.

What is missing?

The objections to wrapping an EDI message in an XML format is that a lot of overhead is created in the message. At this point there is no consensus as to what the tags in XML/EDI should be like. Also, today many initiatives are deployed on XML/EDI. There is strong support from the industry for XML. Companies like Microsoft, IBM, Sun and many others have already created tools for handling XML files. However, all initiatives to apply XML/EDI have one thing in common: they lack a **consistent, generic foundation**. Although XML offers the possibility to create any kind of file, using any kind of tag, there is no defined way to base an XML/EDI message on a unique, global and agreed repository like UN/EDIFACT.

The global repository raises the main issue of consensus building and management of the process as well as the technical aspects of managing codes and terminology. In that context, the importance of UN/EDIFACT as the consensus and business agreement process between business partners should be highlighted. The UN/EDIFACT directories are based on a clear, open and global consensus building process as well as a well-defined maintenance process. Consensus building secures that all parties will play according to the published rules of the UN/EDIFACT directories and therefore does not require business partner discussions in order to initiate transaction globally.

Rather than having HTML for a user and EDIFACT for an application, with XML/UNEDIFACT, we could have an EDI message that can be an EDIFACT orders message written in an XML format; therefore it is presentable to the application just like the EDI file. The same file uses the embedded templates and rules that explain how it should be displayed to a user and can be viewed through a browser.

XML/UNEDIFACT technical feasibility

In the view of the SIMAC Group, there should be a sequential numbering of the tags for XML/EDI; this would be an identifying code only, starting with a letter. This would provide a neutral starting point and would be most efficient in size. Every (qualified) data-element in the EDIFACT directories would have its equivalent in the repository as an XML tag. The format, name and description of this tag is derived from the EDIFACT equivalent. As such the Repository can be created automatically from the UN/EDIFACT directory. The EDIFACT UNSMs will always prescribe what tags (=data-elements) would be legal in a message. In that way, the know-how of business processes and transactions that is captured in these UNSMs does not go to waste; on the contrary it is now exploited in a much more open and versatile way.

Why is it urgent?

Setting up a Repository for XML tags based on UN/EDIFACT should be done now, to avoid a situation in which different communities create their own tags. We could then find ourselves in a situation where these initiatives cannot be merged into a global approach and the work we have put in creating UN/EDIFACT would be lost.

Proposal to CEFACT

If a repository of XML, tags would be created based on UN/EDIFACT, this would provide a solid basis to set up messages in an XML format. UN/CEFACT is in the unique position to create such a Global Repository for XML tags that can be used in any community as a foundation for XML/EDI.

The UN/ECE secretariat could be commissioned to set up this Repository on their Web site so that it could be available to anyone - free of charge. The initial repository could be set up in such a way that only the most used data-elements are included, while a flexible mechanism can be created to provide anyone with the possibility to add tags for those elements that are already defined in the UN/EDIFACT directories. Any request for new data-elements should go through the normal DMR process as not to undermine the existing structure.

XML/EDI might not be the ultimate solution, but a global repository would be extremely useful for conventional EDI and a powerful tool in support of proposed solutions for SME's and other forms of EC.