AN XML METADATA APPROACH TO SEAMLESS DOCUMENT EXCHANGE

JASON UNDERWOOD
Construct IT, University of Salford, Salford, M7 9NU, UK
ALASTAIR WATSON
CAE group, School of Civil Engineering, University of Leeds, LS2 9JT, UK

Abstract

A three year ESPRIT project - ProCure – is ultimately aiming to take a significant but achievable step forward in the application of available ICT to the Large Scale Engineering (LSE) construction industry. The ProCure consortium consists of 5 industrial partners supported by 4 associated research and expert partners. The project combines leading expertise from 3 member states to support ICT deployment by 3 industrial collaborative groups, i.e. UK, Germany and Finland. The basis of the project is in the partners' belief that sufficient ICT is now available to achieve deployment, with care, in real projects, with an acceptable risk of failure. This paper presents work undertaken within the project to investigate the various Metadata standards that exist in order to define a minimum Metadata set based on these standards for the implementation of two demonstrators for XML based automated document exchange between a simulation of a Corporate Document Management System and a simulation of a collaborative construction project web site.

Keywords: ICT, Collaborative Construction Project Web Site, Metadata, XML Schema.

INTRODUCTION

The ProCure Esprit project is a three-year project that commenced at the beginning of 1999 (URL-1). The goal of the project is to take a significant but achievable step forward in the application of available ICT to the Large Scale Engineering (LSE) construction industry. Partners in ProCure consist of 5 industrial partners representing 3 member states, i.e. UK, Germany and Finland, comprising of a client, three types of engineering contractor, and a major product (steel) supplier. The basis for such a project is the partners' belief that sufficient ICT now exists to achieve deployment in real projects, all be it with care and an acceptable risk of failure.

Changes are occurring within the industry and larger clients and engineering contractors are establishing new patterns of working within the current industry structure, which will evolve considerably in the future. Such strategies acknowledge the vital role of ICT in supporting information flow across the many interfaces and information sharing to eliminate interfaces. However, the technology appears complex and its application difficult to envisage for LSE practitioners.

The approach being adopted by ProCure is to define scenarios and specify the ICT for deployment within the first year in addition to agreeing the real projects for the three member state pilots. The three pilots will each be carried out over a 15 month period with staggered starting dates. The three pilots are each different in nature due to the particular businesses sectors they are supporting. However, technological commonality does exist between them in that they are all adopting a Collaborative Construction Project Web Site for the management

and control of documents throughout the life-cycle of a project. This paper presents the work being undertaken within the UK Pilot to implement two demonstrators for XML based automated document exchange between a simulation of a Corporate Document Management System (CDMS) and a simulation of a collaborative construction project web site (CCWS). The work described in this paper has entailed:

- investigating the various Metadata standards that exist in order to define a minimum Metadata set based on these standards,
- extending these standards to form a more comprehensive set, i.e. to support the identified message transactions between the CCWS,
- translation of the IFC 2x Metadata subset from EXPRESS into modular XML Schema subsets based on identified simple modular XML messages.
- implementation of two demonstrators to show 1) an XML transaction by which a CDMS can publish a document with its associated metadata directly to a CCWS, and 2) a complementary XML transaction by which a CDMS can obtain the metadata relating to a particular user from a CCWS.

COLLABORATIVE CONSTRUCTION PROJECT WEB SITES/ HOSTING SERVICES

The construction industry, which currently accounts for some $3.4 trillion in business worldwide, suffers from inefficiency and a high degree of fragmentation. Internet companies have recognised that the internet holds the potential of untangling the issues of collaboration among companies involved in the life-cycle of a building. Consequently, there has been a recent spate of new Internet companies arrive on the AEC scene offering a wide array of project hosting services (URL-2; URL-3; URL-4; URL-5; URL-6; URL-7). These companies aim to provide a solution to collaborative project working in the form of internet-based communication, collaboration and management tools. For a nominal fee, companies are now able to publish and download project information, i.e. documents/drawings, on the Web, allocate project teams, facilitate procurement and bidding, in addition to a host of other services, using standard technology and without the need for specialist intervention. These systems provide such benefits as:

- Managed document workflow.
- Fast searching and easy to find information.
- Fast publishing and downloading.
- Audit trails/logs.
- Automated revision control.
- Automatic generation of notification emails.
- Multi-format storage.
- Standardised document issuing procedures.

METADATA AND STANDARDS

Metadata is defined as "data about data" or alternatively "information about information". In practice, Metadata is the administrative data used in both manual and computerized systems for the organization and management of documents. The most common types of Metadata is information about *documents*, e.g. Title, Name, Author, Editor(s), Type/Classification, Revision information, Size, Storage Location, Format, Status, etc., *people*, e.g. First Name, Family Name, Title, Organisation, Address, Location, Telephone Number(s), Facimile Number(s), Email Address, etc., and *organisations*, e.g. Name, Address, Telephone Number(s), Facimile Number(s), Web Address, etc.

There are a number of Metadata standards that currently exist. For example, data exchange standards such as the IAI's IFCs (URL-8); for sharing data throughout the project life-cycle and across disciplines and technical applications within the domain of Architectural, Engineering, Construction and Facilities Management, CIS (URL-9; URL-10); for the digital exchange and sharing of the engineering information concerning steel-framed buildings or similar structures, and STEP-CDS (Csavajda, et al, 1998); for 2D CAD data exchange, incorporate Metadata as a sub-set of the model. These three standards each support document, person, and organisation Metadata. Several ISO standards have also been published relating to Metadata. These standards range from simply providing a prerequisite for Metadata, to providing guidance for the implementation of records management strategies, procedures and practices, to the definition of a formal model/standard (ISO 11442, 1993; ISO 15489, 1993; ISO 19033, 2000; ISO 82045, 2000).

XML

XML (eXtensible Markup Language) (URL-11) is being hailed as the technology that will take the Internet to a new level and make it truly suited to business. XML proposes to simplify communication between systems and to speed up the retrieval of data - contained in a web page, other structured data or programming code - from different systems.

XML is a subset of the Standard Generalized Markup Language (SGML) intended to make it easy to interchange structured documents over the Internet. XML differs from SGML primarily in simplifying the sometimes intimidating formalisms of SGML, in order to ensure that an XML parser is simple enough to embed in even lightweight software, including Web browsers. It differs from HTML primarily in allowing the user to specify new tags, marking types of elements not foreseen in the HTML specification, and making it possible for common off-the-shelf browsers and other software to handle such user-defined element types usefully.

The XML 1.0 standard first edition was approved and published by the World Wide Web Consortium (W3C) in February 1998. Since then, XML technology has quickly gained favor as a universal data interchange format for networked systems, e.g. Microsoft, Netscape, Oracle, SAP, Sun, etc. The second edition of the XML standard was approved and published in October 2000. Among the practical benefits are:

Structure: to model data to any level of complexity

Extensibility: to define new tags as needed

Validation: to check data for structural correctness

Media independence: to publish content in multiple formats

Vendor and platform independence: to process any conforming document using standard commercial software or even simple text tools.

XML is being seen as the ideal way to facilitate business-to-business communications – Internet EDI (URL-12). In contrast to the method of business communications prior to the emergence of the Internet (EDI), it is suggested XML is simple to implement and at a relatively low cost. EDI has been seen as difficult to set up, involving major investment with dedicated software and expert staff required. XML will enable anyone to create documents that have active elements embedded in them, and send then via the Internet to anyone else. In addition, XML documents feature meta-data that makes it more flexible than conventional EDI, which requires both sender and receiver agreeing in advance on rigid formatting.

Already W3C is recasting the existing HTML 4.0 specification in XML, to create XHTML, suggesting that XML will soon become the base specification in which other Web formats will be defined (URL-13). More recently (September 2000) W3C published a two part working draft, which specifies the XML 1.0 Schema definition language (URL-14). The schema language, which is itself represented in XML 1.0, provides a superset of the capabilities found in XML 1.0 document type definitions (DTDs). Part 1 of the working draft, XML Schema: Structures, specifies the definition language which offers facilities for describing the structure and constraining the content of XML 1.0 documents. Part 1 is currently a public working draft that has been issued by the XML Schema Working Group for review by the public and by members and working groups of the World Wide Web Consortium. The Working Group intends to submit this specification for publication as a Candidate Recommendation very soon. Part 2, XML Schema: Datatypes, proposes facilities for defining datatypes to be used in XML Schemas in addition to other XML specifications. Part 2 is currently an internal working draft for review by members of the Working Group and is awaiting publication as a public working draft.

UK PILOT

Background

The concept for the UK Pilot is a test and demonstration, initially in a pre-pilot but ultimately in a real project, of the relationship between an Engineering Contractor: Taylor Woodrow Construction (TWC), a sub-contractor: a steel fabricator and a major Supplier: Corus Group (CG). For a selected project, TWC will cover the design co-ordination, project costing and construction management aspects including the management of the sub-contracted steelwork fabricator. CG will provide pre-contract design advisory services to the project through a Customer Design Service, will trace the steel from rolling mill to piece position in an erected frame, provide on-line enquiry and order tracking for customers and supply logistics. The UK Pilot is to be developed in two stages. The first stage is that of information gathering and testing of the technology to be used for the second stage, which is the “live” pilot that should utilise the selected technology to undertake an ongoing “real” construction project.

Requirements for the Storage, Transmission and Exchange of Project Information

It is necessary to have robust and secure systems for the storage and transfer of all the project information throughout the life-cycle of a project in an electronic form. Additionally such information should be held in data structures, which can be understood by all relevant parties.

The storage and physical transfer of project information can be achieved by using some form of web accessible neutral project information repository for the exchange and storage of all shared project data to fulfil the role of an Electronic Document Management System (EDMS). This may be achieved using a project specific Secure Web Site, project or company Extranet, or a more sophisticated EDMS such as Global Recall, ProjectWise, etc, which would allow access both over the WWW or by more direct means. Following strong concerns expressed about the contractual implications of using such an EDMS that is under the direct control of just one of the participants in a construction project, a decision was reached for the UK Pilot to adopt a CCWS that is established and managed independent of the possible commercial/contractual interest of a single participant. Furthermore, it is important that the system has the appropriate level of security to ensure that documents are available to all the intended recipients and to them alone. In addition to this secure arrangement, the capability to undertake a full audit of document issue and access is extremely important. All three ProCure Pilots are adopting the above requirements, however, due to the specific needs of the different projects, they are intending to use different approaches for their specific projects.

The proposed solution for the ProCure UK Pilot proposes that a CCWS holds ALL the "issued" project information, including sensitive contractual documents communicated between one or more project partners. Access to information on the CCWS must therefore be strictly controlled.

As previously mentioned, there has been a recent spate of new Internet companies arrive on the AEC scene offering a wide array of project hosting services. In order to select the most suitable CCWS for the UK Pilot a functional assessment exercise was undertaken. This task involved drawing up a more formal definition of the requirements for the CCWS required by the UK Pilot in order to appraise a small selection of candidate CCWS products. As the result of the functional assessment, *the-project* hosted by Sarcophagus Ltd. was selected to be adopted for the UK Pilot (URL-2).

It is envisaged that the CCWS is a collaborative environment for the management and exchange of documents between organisations and is therefore neutral of any individual organisation. The primary role of a CCWS is to give the participants in a project controlled rights to publish and accesses project documents. A key feature of current systems is their easy access via a web-based interface. This provides all project participants with a simple low cost means of access. However, the industrial participants in the UK pilot identified a limitation in this universal means of access. Companies are progressively moving towards Corporate Document Management Systems (CDMS) to manage the in-house production, storage and issue of documents across all projects. This being the case, a complex many-to-many relationship will be develop between multiple CCWS and CDMS, and the existing browser-based human-centric CCWS interface will become inappropriate. Clearly there is a need to establish a complementary universal machine-centric interface between the corporate and project systems.

SIMPLE MODULAR XML MESSAGES

The UK pilot has opted to adopt a modularized approach to the structure of message transactions between the CCWS and CDMS. By adopting such an approach the structure of each type of message is kept simple whilst eliminating the transaction of the whole metadata structure in each and every message transaction. Figure 1 shows the message transactions that have been identified as required between the CDMS and CCWS, together with the type of metadata associated with each message.

Upload a Document: Uploads to the CCWS a single document with its metadata, the identifier of the resulting document file is returned.

Download a Document: Sends the identifier of a document file to the CCWS, the document is download together with its metadata.

Download User Info: Sends the identifier of a user to the CCWS, that user's current metadata is returned.

Download Organisation Info: Sends the identifier of an organisation to the CCWS, that organisation's current metadata is returned.

Create New User: Sends minimal information about a new user to the CCWS, the resulting identifier of the (new) user is returned.

Edit User: Sends to the CCWS the identifier of an existing user with revised metadata for that user, conformation of the update is returned.

Upload Transmittal: Uploads to the CCWS a document transmittal comprising the identifiers of users (the recipients), metadata and the identifiers of existing document files (the payload). A two-stage confirmation is returned: first that the content of the transmittal is valid, and second that the transmittal has been downloaded by all the recipients.

Download Transmittal: Pushes a transmittal (see above) from the CCWS to a CDMS recipient, confirmation of receipt is returned to the CCWS.

Group Uploaded Documents: Used to form a number of previously uploaded documents into a group. Sends metadata and the identifiers of existing document files to the CCWS, confirmation of the grouping is returned.

Structure Uploaded Documents: Used to form a number of previously uploaded documents into a group within which the files are (understood to be) held within a folder hierarchy. Sends metadata and the identifiers of existing document files to the CCWS, confirmation of the structuring is returned.

XML METADATA SCHEMA

IFC 2x Metadata Subset

The Metadata subset of the IFC 2x Final Release has been adopted as the basis for the Metadata for the exchange of information between the CCWS. The IFC 2x Final Release supports the Metadata associated with persons, organisations, and documents.

IfcActorResource

The ***IfcActorResource*** schema within the IFC model provides for the representation of information concerning a *person* or an *organisation* that will undertake work, hold responsibility, etc. Figures 2a & 2b show the EXPRESS-G diagram for the schema. ***IfcPerson*** and ***IfcOrganisation*** supports the identification of properties of persons and organizations, respectively, whose services may be used. A role performed by an *actor*, i.e. a person, an organization or a person related to an organization, is provided in ***IfcActorRoles***. *Address* information relating to a person or organisation is held using the ***IfcPostalAddress*** and ***IfcTelecomAddress***. ***IfcPostalAddress*** holds address information for the delivery of paper based mail. ***IfcTelecomAddress*** represents the address to which telephone, electronic mail and other forms of telecommunications should be addressed. In addition, the ***IfcActorResource*** schema supports relating persons to organizations via the ***IfcPersonAndOrganization*** entity and relationships between organizations using the ***IfcOrganizationRelationship*** entity, e.g. to form hierarchical organization structures.

IFC Documents Model

Document Metadata is supported within the IFC Documents Model of the ***IfcExternalReferenceResource*** schema. The IFC Documents Model enables reference to be made to documents stored externally to a populated IFC model or for information about documents to be stored within the model. The scope of the IFC Documents Model is to:

- manage information about documents
- manage reference to documents
- be equally applicable to documents that are paper based or stored electronically

Figure 3 shows the EXPRESS-G Diagram for the IFC Documents Model. Information about a document is held using the ***IfcDocumentInformation*** class. This identifies and names the document, names the document owner and, optionally, can include a description of the document. In addition, a revision identifier for the document may also be included. Document information can also include creation and revision times and the duration of its validity by reference to 'valid from' and 'valid to' attributes.

Documents frequently hold references to information held in other documents e.g. documents referencing standards that are also documents. A significant tree structure of document information referencing could be built up in this way. Such relationships between document information can be captured through the ***IfcDocumentInformationRelationship*** class which manages both relating and related document information and inversely captures the document information carrying the pointer and the document information to which pointers refer.

The IFC Documents Model supports information being referenced from external sources. Referencing of a document is by its location (address) to enable access through mechanisms such as the World Wide Web. This is done through the ***IfcDocumentReference*** class. This is a type of ***IfcExternalReference*** that has a label (which can be the reference address) and identifier. Additionally, a name attribute provides the document with a human readable extension or qualifier to the location. Optionally, as well as the document reference itself, information concerning the document can also be stored as an attribute of the document reference.

Where a document that is stored electronically, the ***IfcElectronicDocumentFormat*** class enables information about the format to be recorded. This includes the file extension used and information about the content using the MIME standard.

As depicted by the shaded entities in Figure 3, the IFC Documents Model schema has been extended in order to support the document exchange between *the-project* CCWS, i.e. Class, RevisionStage, Status, CollectionID, PaperScale, PaperSize, and Confidentiality.

IFC 2x to XML Schema Translation

There are currently two XML Schema standards. The recently promoted XML Schema standard to "Candidate Recommendation" by W3C and Microsoft's non-W3C-complient implementation XML-Data Reduced (XDR). There are some differences in syntax, and even slightly structural differences between XDR and the XML Schema proposal. For example:

- XDR uses an XML file extension and XDR namespace whereas W3C uses an XSD file extension and namespace.
- Elements and attributes within XDR schemas are defined using *ElementType* and *AttributeType* elements. W3C schemas are defined using *complexTypes* elements within which a number of *element* and *attribute* elements are declared. In turn, *element* and *attribute* elements are either declared as a simple 'built-in' type or by a declared *simpleType* element.

Although XDR has been implemented in MSXML 2.0 and IE5 (and is the basis of Microsoft's BizTalk e-commerce initiative), once the W3C's proposal becomes a Final Recommendation, XDR will likely become obsolete (Hunter et al. 2000).

ProCure UK Pilot XML Translation

The translation conducted for the ProCure UK Pilot demonstrators has adopted the W3C "Candidate Recommendation" schema definition syntax. For the purpose of the ProCure UK Pilot, i.e. simple modular XML messages, the method of translation of the IFC 2x Metadata subset was conducted in the following manner (Figure 4):

- Entities are mapped to complexTypes in the XML model.
- Attributes are mapped to XML 'sub elements' (complexTypes/simpleTypes).
- Simple data types and enumerations are mapped to simpleTypes.
- As the ProCure UK Pilot is adopting simple modular XML messages for the exchange of information between the CCWS, references between entities/elements are not required to be encoded. Therefore, the XML schema subsets for the module messages are composed of simple element/sub element structures.

IAI BLIS XML Translation

Subsequent to the ProCure UK Pilot XML translation, the IAI BLIS group released the XML schema translation for IFC 2x (BLIS-XML) in December 2000. The BLIS-XML is a methodology for transforming EXPRESS data models into 'XML Schema', which is currently based on XDR. However, it is proposed that the newer schema definition syntax by W3C will be used in the future. The BLIS-XML translations are as follows (BLIS, 2000):

- Entities in the EXPRESS model map to elements in the XML model. As XDR does not support inheritance each element in the BLIS-XML schema that inherits attributes in the EXPRESS schema has a copy of the attributes it inherits.
- Attributes are transformed in one of two ways:
  1. Simple data types are mapped to Element attributes.
  2. LIST or SET attributes are mapped to sub elements with the following exceptions:
     - A LIST or SET of object references is transformed to idrefs.
     - SELECT types are modeled using the group construct.
- Enumerations map directly to the enumeration data type in XML.
- References are encoded using the id/idref/idrefs construct in XML. Each entity has a XMLID attribute of type ID, which can be any string value that is legal for a token of type ID.

Methods for IFC to XML Translation

There are currently a number of methods for the translation of EXPRESS to XML. These include:

1. ETEB (EXPRESS Typed Early Binding).
2. OSEB (Object Serialization Early Binding).
3. CEB (Containment Early Binding)
4. Pdnet Early Binding
5. eConstruct approach (EXPRESS to XML via UML).
6. BLIS approach.

Approaches 1. and 2. are part of the upcoming international standard, published by the International Standardization Organization – ISO 10303-28 (URL-15). The third and forth approaches are proposed extensions within the same organization. The fifth approach has been defined as part of the e-Construct ESPRIT project (URL-16) that is aiming to develop an XML vocabulary including a number of national flavours specific for the European BC industry - BCML (Building-Construction Mark-up Language).

The Model Support Group (MSG) of the IAI have recently been tasked by the International Council to come up with a strategy on how to provide XML for IFC 2x – the envisioned ifcXML – by establishing the XML task group. The MSG has started to collect and survey the different methods for translating an EXPRESS model into XML. In order to ensure the IAI adopt the correct translation method for ifcXML sufficient time is required to complete the survey and to propose recommendations.

The BLIS approach is currently the only available XML translation of IFC 2x. It is also a much simpler approach for the translation of EXPRESS to XML than that defined in ISOCD 10303-28. However, the BLIS approach has been defined such that data can be converted from Part21 files to BLIS-XML files and back to Part21 files without loss of information. Such conversion of data from Part21 to XML and visa versa is not a requirement for the

purpose of the UK Pilot but the simplicity of translation and schema structure is a prime concern.

XML METADATA EXCHANGE DEMONSTRATORS

Overview

The demonstrators implemented XML-based information transfer between a simulation of a CDMS and a simulation of a CCWS (specifically the-project). Two XML transactions where implemented to demonstrate and prove the concept, to show how the CCWS side might be implemented, and to develop (and prove against the-project) the required XML schemas - thus enabling a full set of specifications and schemas to subsequently be developed.

For both demonstrators the CCWS client is simulated by a human interacting with Internet Explorer 5.0 (IE5.0), and the CCWS server is simulated using active server pages (asp) running on Information Server Service (ISS). The CCWS client for each demonstrator is a web page (asp) comprising two HTML forms and a piece of client-side JavaScript code. The first of the HTML forms enables the user to enter/select the necessary metadata. Once the necessary metadata has been entered/selected the user submits the metadata to the CCWS, triggering the client-side JavaScript. The client-side JavaScript uses the Document Object Model (DOM) of IE5.0 to generate the XML metadata based on the information entered/selected by the user, in conformance with the respective modular XML message schema, before copying the XML metadata (ASCII text) to a text input form element in the second of the forms and sending the it to the CCWS server via the HTTP POST protocol. Upon receipt of the XML modular message, the CCWS uses the DOM of IIS to parse the XML message, in order to retrieve the data contained within the message.

The DOM is the mechanism defined by the W3C for structural access and modification of XML documents. It allows XML documents to be viewed as a tree of objects, rather than serialized text. There are several advantages with using the XML DOM over manipulating the text directly:

- The XML DOM guarantees that any document created will be well formed as the tree is created as a tree of nodes as opposed to serialized text. Problems such as forgetting to write the end tag for an element to a stream cannot occur.
- The XML DOM allows documents being read to be parsed for well formedness and validity.
- The XML DOM allows documents to be created ad-hoc, with the serialization to a stream only happening after the document has been constructed as a tree.
- The XML DOM allows the search of XML documents to retrieve nodes quickly.

Upload Document Demonstrator

This demonstrator shows an XML transaction by which a CDMS can publish a document with its associated metadata directly to a CCWS. Clearly this is one of the key functionality's that would be required of a XML message-based interface between a CDMS and a CCWS. In designing the document upload message, advantage was taken of the machine-to-machine

nature of the communication by using a short form of the required metadata. Thus, the user who is uploading a document is simply identified by their internal ID on the CCWS (implying that the CDMS will need to track these). This directness has several advantages, particularly as the same message must be capable of unambiguous communication with different CCWSs.

In the demonstrator the human representing the CDMS enters/selects the metadata, relevant to the document required to be uploaded to the CCWS, into a simple form on the CCWS client interface such as Title, Description, Class, Revision Stage, Revision, and Collection ID (Figure 5). Once the relevant metadata has been entered/selected, the user then submits the metadata to the CCWS, i.e. the CCWS interface generates the XML metadata message and posts this to the CCWS. When the CCWS receives the XML metadata message it parses the message to retrieve the necessary document metadata. With this document metadata the CCWS is able to update its system appropriately based on the upload of the document from the CDMS e.g. audit trails/logs, etc. Figure 6 shows the XML upload document metadata message the CCWS receives.

Download User Info Demonstrator

This demonstrator shows a complementary XML transaction by which a CDMS can obtain from a CCWS the metadata relating to a particular user (who may have just uploaded a document). In this case the request message is very simple, the user being identified by their internal ID on the CCWS. It is the reply message that is more interesting allowing the CCWS to return what information it has about that user to the CDMS.

In the demonstrator the human representing the CDMS selects an available user ID from a CCWS client interface and submits this to the CCWS (Figure 7). By submitting the user ID, the CCWS interface is triggered to generate a simple XML message and posts this to the CCWS. The CCWS receives the XML metadata message and parses the message in order to retrieve the user ID. Based on the selected user ID, the CCWS retrieves the relevant user information from its underlying rational database and generates the user info XML metadata message before returning this to the CDMS (Figure 8). With this user information metadata the CDMS is able to update its system as appropriate.

CONCLUSIONS

This paper has presented part of the work that has been undertaken within the UK Pilot of a EU funded project - ProCure. The ProCure project is aiming to take a significant but achievable step forward in the application of available ICT to the Large Scale Engineering (LSE) construction industry. The project consists of three member state pilots from UK, Germany, and Finland, which although are different in nature due to the particular businesses sectors they support, do share technological commonality in that they are all adopting a Collaborative Construction Project Web Site for the management and control of documents throughout the life-cycle of a project.

The work undertaken within the UK Pilot has resulted in the implementation of two demonstrators for XML based automated document exchange between a simulation of a CDMS and a simulation of a CCWS to show:

1) an XML transaction by which a CDMS can publish a document with its associated metadata directly to a CCWS, and
2) a complementary XML transaction by which a CDMS can obtain the metadata relating to a particular user from a CCWS.

From an examination of the various Metadata standards that currently exist, the UK Pilot has adopted the Metadata subset of the IFC 2x Final Release as the basis for the Metadata for the exchange of information between the CCWS. The subset Metadata of the IFC 2x Final Release has been extended to form a more comprehensive set, i.e. to support the identified message transactions between the CCWS. The UK pilot has opted to adopt a modularized approach to the structure of message transactions between the CCWS and CDMS, thereby keeping the structure of each type of message simple and eliminating the transaction of the whole metadata structure in each and every message transaction. The extended IFC 2x Metadata subset has been translated from EXPRESS into modular XML Schema subsets based on the identified simple modular XML messages and the W3C schema definition syntax. The implementation of both demonstrators has adopted web based technology – active server pages together with client-side JavaScript - built on top of ISS and using the IE5.0 and IIS DOM's to create and parse the XML metadata transaction messages. For the two demonstrators a human interacting with a specific web page (asp) via IE5.0 simulates the CCWS client, and the CCWS server is simulated using asp running on ISS.

The implementation of these two XML messages transaction demonstrators has:

- demonstrated and proved the concept of an XML Metadata approach to seamless document exchange between a CDMS and a CCWS,
- showed how the CCWS side might be implemented,
- developed and proved against *the-project* the required XML schemas, which in turn will enable a full set of specifications and schemas to subsequently be developed.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the contribution of the ProCure project partners and in particular the UK Pilot team. The work presented within this paper is part of the ProCure ESPRIT project 29948 funded by the European Commission.

REFERENCES

URL-1, PROCURE: ICT at work in the LSE Procurement Chain, http://cic.vtt.fi/projects/procure/.
URL-2, Sarcophagus Ltd: Information Systems Consultants – THE PROJECT Repository, http://www.sarcophagus.co.uk/project_repository.htm.
URL-3, Buzzsaw.com, http://www.buzzsaw.com/.
URL-4, Cephren, http://www.cephren.com/.
URL-5, Integration, http://integration.arup.com/.
URL-6, Sarcophagus Ltd: Information Systems Consultants, http://www.sarcophagus.co.uk/.
URL-7, 4projects, http://www.4projects.com/.
URL-8, International Alliance for Interoperability, http://www.interoperability.com/.
URL-9, Welcome to the official web site for the second release of the CIMsteel Integration Standards, http://www.cis2.org/.

URL-10, CIS Homepage, http://www.cae.civil.leeds.ac.uk/past/cis/cis.htm.
Csavajda, P., Haas, W., Kalpakidis, G. (1998) STEP and its implementation in factory design – the STEP-CDS initiative, Proceedings of ECPPM '98.
ISO 11442 (1993) ISO 11442 (all parts) Technical Product Documentation: Handling of Computer-based Technical Information, International Organisation for Standardisation.
ISO 15489 (1993) ISO 15489 (all parts) Records Management, International Organisation for Standardisation.
ISO 19033 (2000) PD ISO/TR 19033:2000 Technical Product Documentation – Metadata for Construction Documentation, British Standards Institute.
ISO 82045 (2000) IEC/ISO 82045 (all parts) Management Data (Metadata) for Technical Documents, International Organisation for Standardisation.
URL-11, Extensible Markup Language (XML), http://www.w3.org/XML/.
URL-12, XML/EDI: "the e-Business framework", http://www.xmledi.com/.
URL-13, XHTML™ 1.0: The Extensible HyperText Markup Language, http://www.w3.org/TR/xhtml1/.
URL-14, XML Schema, http://www.w3.org/XML/Schema.
Hunter, D., Cagle, K., Gibbons, D. (2000) Beginning XML Programming, Wrox Press Ltd, 2000.
BLIS (2000) BLIS-XML Methodology for Transforming EXPRESS data Models to XDR, BLIS Project Companies, November 2000.
URL-15, XML representation of EXPRESS schemas and data, http://www.pdtsolutions.co.uk/standard/xml-express/.
URL-16, Welcome to eConstruct!, http://www.econstruct.org/.