Communications of the Association for Information Systems

Volume 14 Paper

December 2004

E-Document Management Based on Web services and XML

Ou Liu
City University of Hong Kong, iuou@cityu.edu.hk

Jian Ma
City University of Hong Kong, isjian@cityu.edu.hk

Ji-Ye Mao
Renmin University of China, jymao@ruc.edu.cn

Ron C.W. Kwok
City University of Hong Kong, isron@cityu.edu.hk

Dong Li
National Natural Science Foundation of China

Follow this and additional works at: https://aisel.aisnet.org/cais

Recommended Citation
Liu, O., Ma, J., Mao, J., Kwok, R., & Li, D. (2004). E-Document Management Based on Web services and XML. Communications of the Association for Information Systems, 14, pp-pp. https://doi.org/10.17705/1CAIS.01433

This material is brought to you by the AIS Journals at AIS Electronic Library (AISeL). It has been accepted for inclusion in Communications of the Association for Information Systems by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Communications of the Association for Information Systems

E-DOCUMENT MANAGEMENT BASED ON WEB SERVICES AND XML

Ou Liu,
Jian Ma
Information System Department
City University of Hong Kong
isjian@cityu.edu.hk

Ji_Ye Mao
School of Business
Renmin University of China

Ron C.W. Kwok
Information System Department
City University of Hong Kong

Dong Li
Information Center
National Natural Science Foundation of China

ABSTRACT

Document management plays an important role in R&D project management for government funding agencies, universities, and research institutions. The advent of Web services and XML presents new opportunities for e-document management. This paper describes a novel solution for processing large quantities of electronic documents in multiple formats within a short timeframe. The solution is based on Web services for integrating two-tiered distributed processing. It also involves a document extraction process for handling multiple document formats, with XML as the intermediate for information exchange. The application of the solution at the National Natural Science Foundation of China (NSFC) proved successful, and the general approach may be applied to a broad range of e-document management settings.

Keywords: e-document management, Web services, XML, e-government solution

I. INTRODUCTION

Document management plays an important role in R&D project management [Eloranta et al., 2001] in government funding agencies, universities, and research institutions. At various stages of project management, such as proposal submission, project selection, and project evaluation [Bonett and Deckro, 1993, Jiang and Klein, 1999, Marble, 2003, Schniederjans and Wilson, 1991, Tian et al., 2004], many kinds of documents or forms are exchanged among a diverse range of stakeholders such as applicants, research administrators of fund recipient organizations, external

reviewers, and program directors and management. Prior research studied document management problems either in general or from a particular viewpoint such as document retrieval or document manipulation [Chin, 2001, Foo and Lim, 1997, Jones and Morrison, 1993, Lambrix and Padgham, 2000, Zantout and Marir, 1999]. This research focuses on e-document exchange and information extraction, which is an important topic of document management [Chang et al., 2003, Hao et al., 1996, Iocchi, 1999].

This paper describes an application developed for the National Natural Science Foundation of China (NSFC) to process research grant applications. The research problem arises from the need to develop a practical solution to deal with some challenges faced by NSFC. NSFC is the largest and most prestigious government funding agency for basic research in China, with an annual budget of over RMB 2 billions (over 250 million US dollars) in 2003. Each year, NFSC receives more than 40,000 grant applications from over 1,400 universities and research institutions in China. Since 2003, all applications have been submitted and processed electronically using the solution described in this paper.

The management of research grant applications is a document-intensive task for NSFC. The traditional approach is primarily paper-based. Proposals are prepared by applicants using various word-processors and then printed out for submission. Fund recipient organizations then review and endorse the proposals, and send multiple hardcopies to NSFC via special delivery services or personally delivered bulky packages. Much of the proposal information is re-entered into management information systems at NSFC for evaluation and administrative purposes. The project selection process is coordinated by senior management of NSFC and administered by the program directors from seven departments organized corresponding to research disciplines. It needs to be simple and transparent to ensure effective communication among stakeholders.

The widespread use of the Internet offers an opportunity to enhance the proposal submission process. For example, the National Science Foundation (NSF) of the United States developed a system, named FastLane (https://www.fastlane.nsf.gov), which allows online submission of electronic proposals. In 2000, NSFC started to experiment with the use of the Internet to streamline its project management processes. After a successful pilot run in 2002, NSFC rolled out electronic submission nation-wide in January 2003.

A major challenge faced by NSFC is that the due date for a great majority of the proposals is March 31. Proposals must be processed within one week so that review forms can be distributed to the external reviewers on time. For example, in a short interval prior to March 31, 2003, NSFC received 34,820 electronic applications from over 1,400 universities/research institutions nation-wide, and all of them were processed by the solution described in this paper. About 150,000 evaluation forms were filled in by external reviewers and processed by the system during April and June. The sheer volume and peak load present a unique challenge for NSFC. Other major funding agencies around the world either distribute proposal submission dates over the whole year (e.g., the U.S National Science Foundation) or they simply do not have the same volume.

Another major challenge facing NSFC is the diversity of document formats and Internet accessibility. Throughout the grant application and review processes, application forms and evaluation forms are submitted in different document formats, reviewed and exchanged by various stakeholders. Therefore, a well-accepted intermediate is needed for the interchange of multi-formatted documents across different application platforms. Internet access is not free of charge to many applicants in China, nor is it reliable or accessible to those in certain remote areas. Therefore, it is infeasible for most applicants to prepare their proposals online or use the Web to fill in a major portion of the application data. Alternative means for document submission are needed in addition to the Internet.

Distributed architecture is a common approach to solving workload problems [Casavant et al., 2001, Jain and Purao, 1991]. A two-tiered distributed solution was developed for NSFC to handle the challenges and to support the existing management process.

In brief, as a result of our work, the following procedure is now used: At the local level, applicants submit their proposals electronically to their respective universities or research institutions first for

review and endorsement. Each organization extracts required data from the proposals, and then forwards the proposals along with extracted data in batch to NSFC. The e-document management solution consists of two parts: (1) an Internet-based Research Information System (IRIS) used at the local level and (2) The Internet-based Science Information System (ISIS, http://isis.nsfc.gov.cn) used at NSFC. The distributed processing gives rise to the use of Web services as the means for integrating the two levels, and to the use of XML to facilitate information exchange across different application platforms.

II. THE WORKFLOW AND SYSTEM ARCHITECTURE

In the NSFC solution, submission and processing of application proposals involve the following steps:

1. A registered applicant downloads the proposal form in Microsoft Word format, or Kingsoft Word Processing System (WPS) format, which is the domestic rival of Word in China. The applicant fills in the form offline, and then submits it.
2. The local IRIS then extracts data from the submitted form, and generates PDF document for review and record keeping.
3. After gathering all submitted application forms, the organization manager of a fund recipient organization submits the processed application files (i.e., a data file containing extracted data, and the application documents) from IRIS to ISIS through Web services.
4. Once ISIS receives the submitted files from registered IRISs, it merges the data files into its central database, and saves application documents into designated directories of a document repository.

A detailed workflow, to solve the peek workload problem corresponding to these four steps, is shown in Figure 1. It involves three kinds of stakeholders:

- registered applicants,
- registered organizations, and
- NSFC.

The data extraction and process are divided into three parts: (1) proposal checking, protection, and preprocess by embedded macros before submission; (2) proposal data extraction and PDF file conversion by IRIS; and (3) proposal data integration and document storage by ISIS.

A distributed system architecture with Web services is shown in Figure 2. IRIS is used by fund recipient organizations to accept online submission of proposals (arrow 1), and to extract required information for local storage and management. A stand-alone subset of IRIS, MiniIRIS, is used for this purpose (arrow 2) by organizations that cannot use IRIS because, for example, fast and reliable Internet access is not available. Prior to the submission deadline, the extracted data and proposals are sent to ISIS via Web services (arrow 3). For MiniIRIS users, the means for file transmission is FTP, compact disks or even hard disks. In addition, applications from organizations with less than 5 applications per year can be submitted to ISIS directly (arrow 4).

To integrate the two tiers and allow necessary interaction between NSFC and the fund recipient organizations, both ISIS and IRIS provide services for each other in the form of Web services. XML is used as the intermediate format for information exchange, e.g., transmission of extracted proposal data. Together these two techniques allow the use of various computer platforms (e.g., Unix, and MS Windows) and electronic document formats (e.g., MS Word, WPS, and databases). Detailed implementations are presented in the following sections.

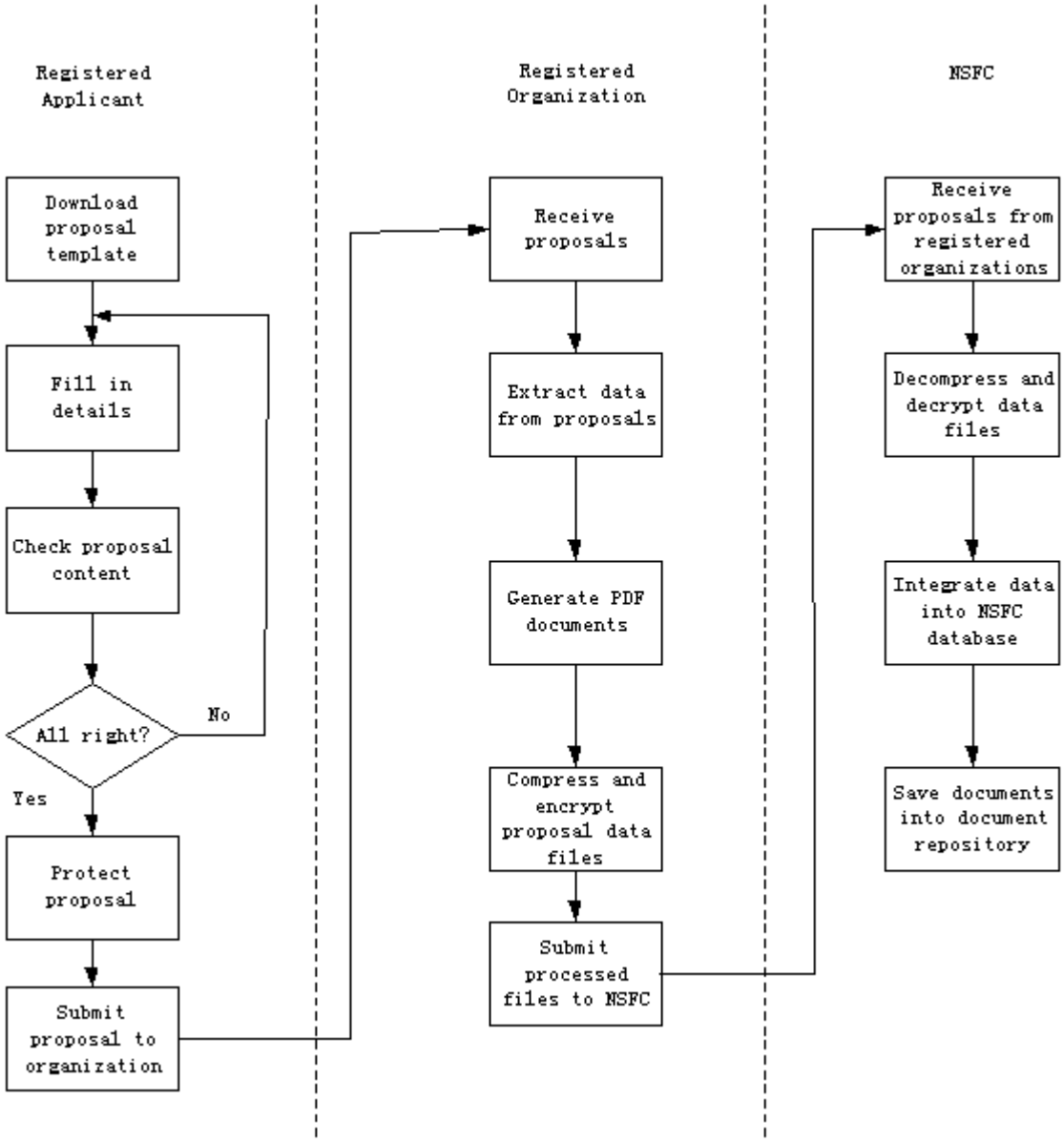

Figure 1. Workflow of Proposal Submission and Process

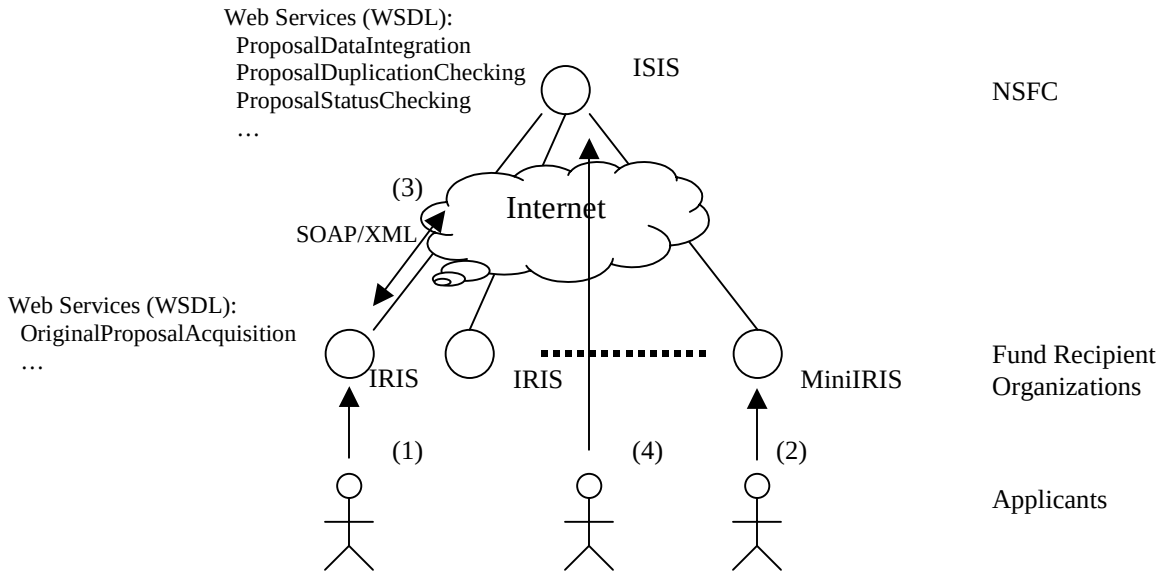

Figure 2. A Distributed System Architecture to Support the Two-Tiered Approach

III. WEB SERVICES ARCHITECTURE AND IMPLEMENTATION

In our proposed architecture, Web services are used to integrate the two tiers of information systems. The most important goal of system integration is to share and reuse the information/data and program/service in different information systems. The use of Web Services on the World Wide Web, which is analogized to the use of CORBA in client/server environments by some researchers, is expanding rapidly as the need for application-to-application communication and interoperability grows. Web services provide a standard means of communication among different software applications running on different platforms [Ferris and Farrell, 2003]. The platform- and language-independence feature of Web service technology makes it a suitable approach to information system integration. The power of Web services, apart from their great interoperability and extensibility thanks to the use of XML, is that they can be combined in order to support complex workflow, such as that in NSFC. In the NSFC case, C# is used to implement the services on Microsoft .NET platform.

In Figure 2, Each Web service includes some relevant functions/operations to fulfill a specific task. For example, ISIS provides the following services for those tasks that require validation at NSFC or cannot be fulfilled by IRIS:

- "ProposalDataIntegration" is responsible for the integration of the proposal data sent from IRIS. An IRIS packages the extracted data in a fund recipient organization and sends them in XML format to ISIS. The Web service receives and classifies the packages, and converts the data into the ISIS database.
- "ProposalDuplicationChecking" checks potential duplication according to various NSFC regulations, which could disqualify applications. For example, a proposal should not be identical or similar to existing or previous ones. Moreover, each applicant can only submit up to two proposals each year. A researcher with associate professor's title or below can only lead one project, whereas a senior researcher ranked at full professor or its equivalent can lead or participate in up to two projects at any given time. Conformance to these regulations can only be checked by ISIS at the national level. IRIS can only do it within a local organization due to limited data. Therefore, it needs to call ISIS for the service.
- Similarly, "ProposalStatusChecking" enables applicants from local IRISs to check the current approval status of a specific proposal submitted to NSFC.

In return, IRIS provides services to be requested by ISIS. These services allow NSFC to access additional information about applicants and their applications. For example, “OriginalProposalAcquisition” allows NSFC staff to get the original proposal documents from remote IRIS sites.

In Figure 2, SOAP (Simple Object Access Protocol) messages are sent between the service requestor (ISIS or IRIS) and the service provider for calling a Web service (arrow 3). SOAP is a standard messaging protocol across Web services in the Internet, and is represented in XML syntax and restricted by XML Schema. To learn how to interact with a Web service using SOAP/XML messages, the service requestor must retrieve the service description first. WSDL (Web Services Description Language) is used to describe functional interfaces of Web services.

A sample Web service description for “ProposalStatusChecking” is shown in Figure 3. It first defines the input and output messages in SOAP format, and then the operation interface of the Web service. Then, it makes a binding, and links to the Internet address of the real Web service implementation. With this definition, client applications can be connected correctly with the Web service by sending and receiving SOAP messages in the right formats.

The proposed architecture can be easily extended to support more document management-related services, which can be requested through widely accepted Internet standards and protocols. The Web service-based architecture provides standards, convenience, reliability and security for system integration. Without the Web service mechanism, remote access to many services in ISIS such as “ProposalDuplicationChecking” would be very cumbersome.

```
<!-- The definition of “ProposalStatusChecking” service starts here -->

<message name="ProposalStatusSoupIn">                    <!-- input message -->
        <part name="parameters" element="nsfc:ProposalID" />
</message>

<message name="ProposalStatusSoupOut">                   <!-- output message -->
        <part name="parameters" element="nsfc:ProposalStatus" />
</message>

<portType name="ProposalStatusPortType">          <!-- definition of operations-->
        <operation name="GetProposalStatus">
                <input message="nsfc: PropStatusSoupIn" />
                <output message="nsfc: PropStatusSoupOut" />
        </operation>
</portType>

<binding name=" ProposalStatusSoup" type="nsfc: ProposalStatusPortType">
<!-- binding information of operations -->
        ...
        <operation name="GetProposalStatus">
                ...
        </operation>
        ...
</binding>

<service name="ProposalStatusChecking">           <!-- the Web service interface -->
        <port name="ProposalStatusPort" binding="nsfc: ProposalStatusSoup">
          <soap:address location="http://localhost/ ProposalStatusChecking.asmx" />
        </port>
</service>
```

Figure 3. WSDL Fragments of a Sample Web Service

IV. E-DOCUMENT EXTRACTION AND MANAGEMENT

XML is another key technology to the NSFC for e-document management. It is the intermediate representation format for data storage and transmission between different application platforms. As a global standard with wide acceptance, XML provides a simple, standard, self-describing way of storing and exchanging text and data. Moreover, XML also provides a means for Web interchangeable data representation.

The system architectures for document extraction and management in ISIS and IRIS are similar. They are illustrated in Figure 4 with IRIS as an example: (1) at the front-end, users, including applicants and external reviewers, download the proposal templates and submit the completed proposals in different formats; (2) at the back-end, corresponding program modules extract the incoming messages, then store the documents and extracted data into document repository and databases respectively.

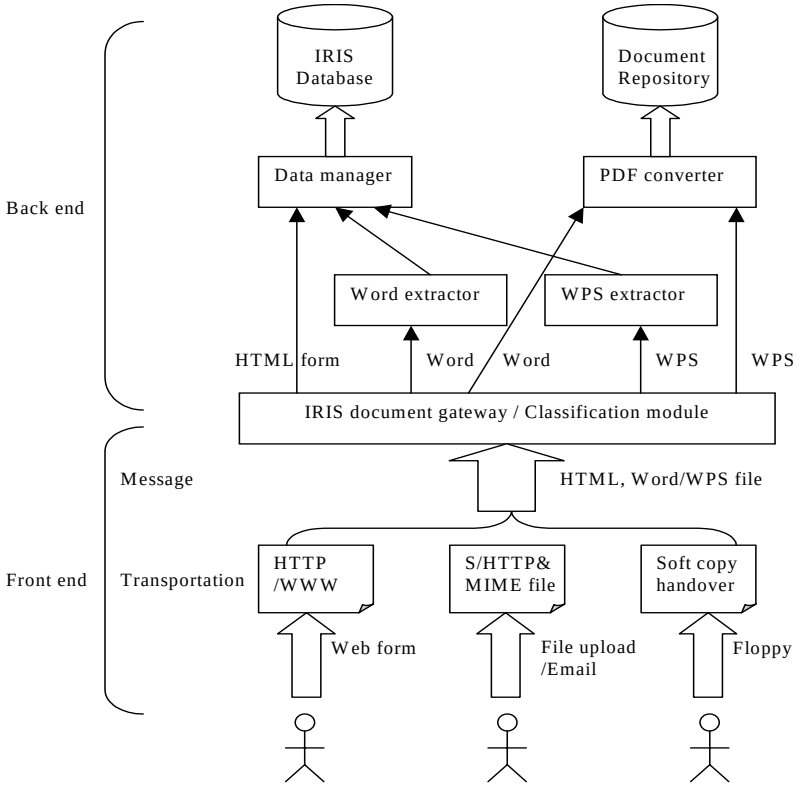

Figure 4. e-Document Extraction and Management

Front-end. Applicants interact with the systems using standard Web browsers, electronic mail applications, or even floppy disks for proposal submission. The multiple channels are necessary in light of the diversity of Internet accessibility in China. The front-end consists of two layers, the transportation layer for enveloping and carrying messages through the Internet, and the message layer for representation of the messages.

The transportation layer makes use of different transportation protocols depending on the mode of submission. Simple Mail Transfer Protocol (SMTP), Secure Multi Purpose Internet Mail Extension (S/MIME) and Secure HyperText Transfer Protocol (S/HTTP) are the three major ones being used. Security measures have been applied to the usage of these protocols. The system requires users to provide their certified key when performing the transfer. The keys are released by NSFC during the user registration and authentication process.

In the message layer, the HyperText Markup Language (HTML) standard and third party document standards (e.g., Microsoft Office and Kingsoft WPS) are used as the representation vehicles. Applicants can submit proposals to IRIS in one of the following two ways, HTML form filled in online with attachments, or full proposal based on Word or WPS formats.

When choosing HTML form format, users need to fill in all required proposal data to be stored in the IRIS database. IRIS then generates a pre-filled document (with the submitted data already filled in) in Word or WPS format by option. The applicant downloads the form and fills in other required contents.

When choosing Word or WPS format, users can download a proposal template and fill in all required data (e.g., Figure 5). The templates were created with (1) embedded XML tags for the data fields to be extracted, (2) data objects such as list-boxes and pull-down-lists enabled with macros to allow ease of data entry and representation, and (3) macro programs for validity checking and data capture. Before submission, the completed form must be checked by the built-in macros for data completeness, validity, and conformance to business rules, and then protected with password against unauthorized changes. After checking and protection, all required data to be extracted are captured and stored as XML strings hidden in the proposal document. The data can be extracted efficiently by corresponding modules in the back-end systems for storage in databases. Extraction is a key measure in the e-document management solution to reduce the workload at the server side and increase system efficiency.

Back-end. Documents submitted to IRIS are sent to the corresponding extraction modules after being classified by the classification module. The various extractors (shown in Figure 4) extract required data from the submitted documents in different formats, and then generate data in XML

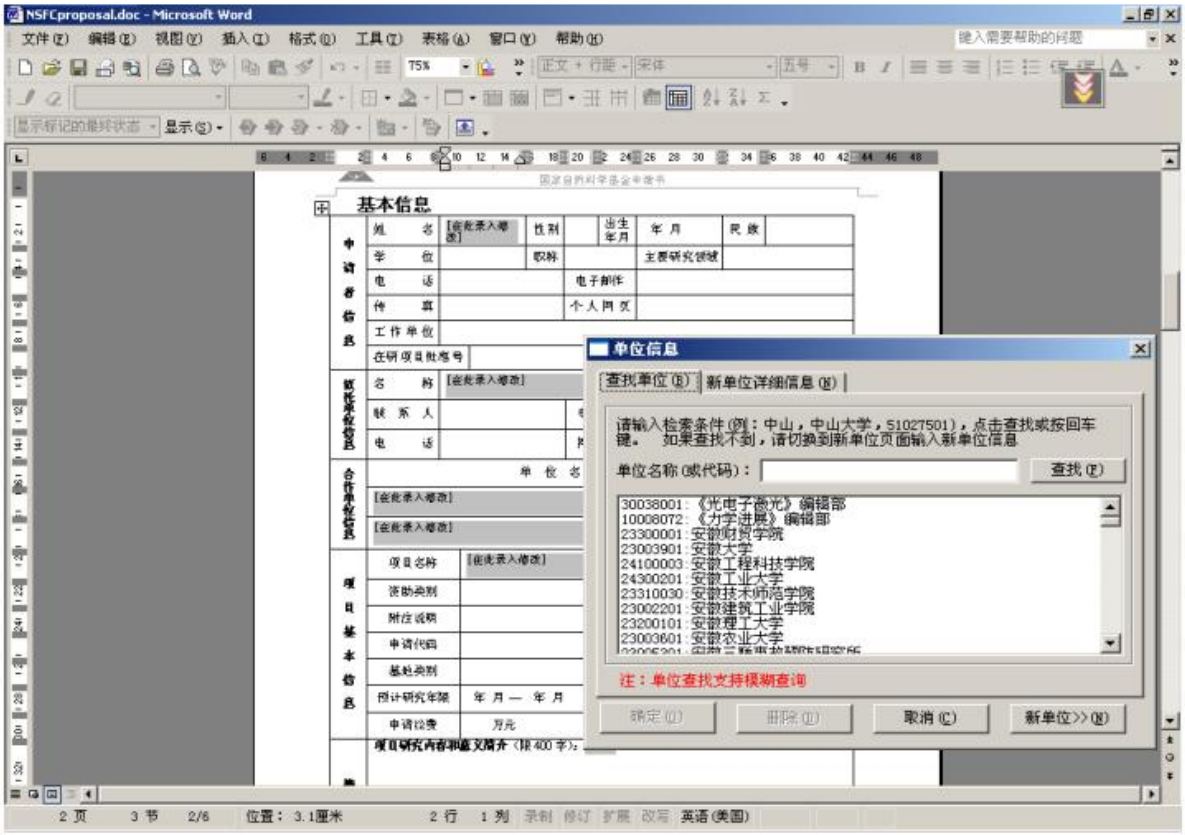

Figure 5. Word-based Proposal Document to be Prepared Offline

format for a module called “Data manager,” which parses the XML data and exports them into IRIS database. This approach makes the solution less dependent on the implementation platform. Even when some components, such as the DBMS, are to be changed, only minor modification is needed for the system. Finally, a module called “PDF converter” converts the whole proposal documents into PDF format and keeps a copy in the document repository.

The implementation of the extractors varies. For a HTML-based document, a simple web server can extract the data from the POST/GET data packet. For Word or WPS documents, data extraction involves picking up the XML strings embedded in the documents. Since XML is used as the intermediate data format, the workload of those extractors is reduced greatly. The data extraction process can be extended to support other third party document formats because the use of XML format as intermediate is generalizable.

V. IMPLEMENTATION

FUNCITONAL MODULES

This subsection describes the major functional modules in Figure 4.

Word extractor. Data extraction functions are performed by programming a relevant Microsoft Word API (Application Programming Interface). Prior to proposal submission, embedded macro programs call Microsoft Word components to validate the contents, capture required data, for project management, and hide them in XML strings (Section IV). At the back end, the Word extractor obtains the required data from the Word file at the server side. The ActiveX standard is used as the basic protocol to communicate with Word components.

WPS extractor. Kingsoft WPS file format is another commonly used document standards in China, released as part of Kingsoft’s WPS Office. Like Microsoft Word format, WPS file format is also selected as a document standards in ISIS. WPS Office series also provide an API for the manipulation of document content. Similar to data extraction in a Word file, embedded macros extract data, which are retrieved at the server side by calling WPS Office supporting components.

Data Manager. Results in XML format from the above two extractors are sent to the data manager to populate the database. The data manager includes two sub modules, a database access module and a data integration module. The database access module provides the basic support for database communication, including connection creation, transaction coordination, and query execution. The data integration module receives the extracted proposal data, parses them by translating the serialized XML text and integrates them into existing data set. The data integration module is wrapped as a Web service so that it can be called by IRIS from remote sites in a standard manner for data exchange.

PDF converter. Adobe Acrobat Portable Document Format (PDF) is selected as the standard format of document repository in ISIS because it is one of the most widely used document standards for e-document exchange and management. Besides portability and platform-independence, it provides support to the Document Object Model (DOM). The PDF converter is implemented by the programming interface of Adobe Acrobat. It converts Word or WPS files to PDF files by calling the corresponding components.

SECURITY CONSIDERATIONS

ISIS/IRIS systems ensure data security in multiple ways at both document and transportation levels. At document level, once a proposal passes the built-in validity check, the applicant is prompted to provide a secret key to prevent modifications by anyone other than the author. During the proposal review process, the system sends encrypted proposals to reviewers so that only the reviewers selected can view the proposal using their assigned key. The keys are issued by NSFC separately during the user registration and authentication process.

At the transportation level, Secure Socket Layer (SSL) technology and Public Key Infrastructure (PKI) are applied during data transfer between the server and clients. The system requires users

to provide their certification keys when performing the transfer. When using MiniIRIS to submit a package of proposals to ISIS, a specially designed algorithm is used to compress the data package, which can only be decompressed by a corresponding algorithm; a sequence number is also assigned to each university/institute to authenticate its identity while submitting the package.

Strict access control is also implemented in the system architecture to ensure security. The Intranet servers and Internet servers are separated with different authorizations to different user groups. Furthermore, different users can only access their pre-specified Web pages related to the authorized functions and information.

PERFORMANCE CONSIDERATIONS

Multiple means are allowed for proposal submission in light of potential Internet access and connection speed constraints (Section II). Moreover, given the peak-load challenge, performance is an important issue to the success of the solution. In addition to the distributed process (Section II), workload balancing is also applied to databases, Web servers and application services:

1. Separate database and database servers are used for storing proposals, experts, review comments, and public queries based on their business process duration and access frequency.

2. Three Web servers and Microsoft's Web workload balancing technology are used to release the bottleneck during the peak time.

3. Application services, such as document extraction, PDF generation and email processing, can be executed in parallel on multiple machines controlled by a task manager. As a result, when one server encounters failure, its tasks can be taken over by other servers and its information can be recovered.

VI. EVALUATION OF THE SOLUTION

ADVANTAGES OF THE SOLUTION

The e-document solution developed for NSFC was proven successful by the pilot run in 2002 involving 16 universities and research institutes. In 2003, it was rolled out at the national level and sustained the annual peak load in March. The system architecture with Web services allowed efficient communication and coordination among distributed document management systems. It greatly improved data reliability and quality, timeliness, and consistency between NSFC and local universities/research institutes.

The e-document management solution shortened the cycle time, and saved processing expenses, by eliminating the need for multiple hardcopies of proposals, postage associated with external reviews, and data entry work. It also enabled a re-engineered business process. Proposal data were extracted and available for decision-making and subsequent management processes (e.g., assigning external experts for peer evaluation of proposals) soon after the submission deadline.. Furthermore, proposals can be sent to external reviewers as PDF documents as part of the online evaluation process. As a result, NSFC decided to waive the RMB500 (about US$60) processing fees paid by the applicant of each proposal. Whereas five hardcopies were still required by NSFC in 2003, only one copy with signatures and official stamps was required in 2004 for archival and legal considerations. This paperwork reduction will result in further efficiency and savings for all stakeholders.

E-document management also benefits fund recipient organizations. Compared to the previous paper-based approach, major benefits include reduced workload and handling of paperwork. Previously, administrators needed to review the proposals and check their validity and completeness manually, (e.g., to make sure NSFC regulations are followed and they contained no obvious omissions) to improve the application's probability of success. Much of this work is

now done by the e-document solution Organizations, however, still need to review and endorse the proposals. Since data extraction is done at the local level, much data entry/re-entry is eliminated. Furthermore, common managerial reports and statistics can be generated by IRIS.

A unique strength of the NSFC e-document is that applicants can use a popular word processor to prepare their entire proposals offline, and submit them to their local organizations. This step can be performed without Internet access, which is a necessary and useful alternative to a national e-document management solution for a developing country like China. Moreover, the Word/WPS-based full proposal affords several distinct advantages:

1. compared to Web-based form-filling, the Word/WPS-based proposal is easy for frequent access and modification, and easy to share among collaborating researchers;

2. consistency is maintained in the format and layout between the electronic copy and hardcopy for signatures and review; and

3. the process of proposal preparation is convenient and compatible with users' normal way of document preparation, allowing custom formatting and the use of advanced word processing functions.

SURVEY

A survey was conducted on a sample of 120 professors and research administrators from 16 universities in China after the pilot-run of electronic submission of proposals in 2002. The great majority of the respondents (95%) to our survey found the e-document solution easy to use. Such high usability score showed that the offline preparation of full proposal was easy to use, and much welcome by researchers. Similarly, 95% of the users either strongly agreed or agreed that e-document management was convenient. Most importantly, 87% either strongly agreed or agreed that it enhanced efficiency.

VII. DISCUSSION AND CONCLUSION

Although it was developed to address several unique challenges faced by NSFC, the two-tiered approach can be adapted and applied to a broad range of e-document management settings. For example, as far as we know, most of the universities in North America and Asia perform their local project administrative functions in collaboration with national funding agencies in a two-tiered manner. Whereas e-Government/business applications are used at both levels, there is no direct information sharing between their systems. Typically, individual researchers submit their proposal data and documents electronically to the national research agencies such as the NSF. Our two-tiered solution based on Web services and XML can facilitate greater efficiency of Internet-based applications and inter-organizational integration.

The use of XML as the intermediate for data extraction and interchange between different document formats and platforms is also novel and applicable to other e-document management tasks. As discussed in Section IV, with its global standard and wide acceptance, XML provides a simple, standard, self-describing way of storing and exchanging text and data. Therefore, we believe that our solution is generalizable.

ISIS currently provides user interfaces and data management in three language standards including Simplified Chinese characters, Traditional Chinese characters, and English. All display and prompts to users are kept in a data file, instead of being hardwired into programs, thus system administrators at the funding agency can access and customize them. Similarly, system administrators can customize proposal labels. Therefore, we believe it is easy to extend the solution to support other languages.

A limitation of the NSFC solution is that in developing countries, Internet and network infrastructure can be a hurdle for the two-tiered approach. For example, institutions without fast or reliable Internet access cannot run IRIS. Instead, they must use a standalone version (MiniIRIS, shown in Figure 2) to handle proposal submission and data extraction. Furthermore, as an

Internet-based system, IRIS is subject to network infrastructure constraints, and is far more complicated in system installation and technical support than a standalone system.

Another limitation is that, from an operational point of view, our solution is based a one-to-many relationship between one particular funding agency and all universities and research institutions as fund recipient organizations. However, where there are multiple funding agencies, a local IRIS needs to communicate with multiple e-document management systems in a many-to-many relationship. It would be ideal to create a generic version of IRIS, instead of one for each funding organization. However, technological and organizational challenges need to be overcome to develop and deploy systems with such flexibility and scalability. For example, both in administrative procedures and information requirements for project management would need to be standardized.

In summary, our e-document management solution proved effective in addressing several unique challenges faced by NSFC. It is based on a two-tiered approach with the use of Web services for system integration between the two levels and the use of XML as the intermediate format of information representation. Moreover, multiple submission channels are used to deal with the Internet accessibility constraint, which is typical in developing countries. The innovative application of Web services and XML for e-document management is flexible and generalizeable to fit the need of a broad range of e-document management tasks.

Editor's Note: This article was received on May 30, 2004. It was with the authors for approximately 5 months for 1 revision and was published on December 22, 2004.

REFERENCES

Bonett, D. G. and R. F. Deckro (1993) "A Multinomial Project Evaluation and Review Technique for Information Systems Analysis and Design", *Information & Management* (25) 1, pp. 51-55.

Casavant, T. L., T. A. Braun, S. Kaliannan, T. E. Scheetz et al. (2001) "A Parallel/Distributed Architecture for Hierarchically Heterogeneous Web-Based Cooperative Applications", *Future Generation Computer Systems* (17) 6, pp. 783-793.

Chang, C.-H., C.-N. Hsu, and S.-C. Lui (2003) "Automatic Information Extraction from Semi-Structured Web Pages by Pattern Discovery", *Decision Support Systems* (35) 1, pp. 129-147.

Chin, A. G. (2001) *Text Databases and Document Management: Theory and Practice*. Hershey, PA: Idea Group Publishing.

Eloranta, E., A. Hameri, and M. Lahti (2001) "Improved Project Management Through Improved Document Management", *Computers in Industry* (45) 3, pp. 231-243.

Ferris, C. and J. Farrell (2003) "What Are Web Services?", *Communications of the ACM* (46) 6, pp. 31.

Foo, S. and E.-P. Lim (1997) "A Hypermedia Database to Manage World-Wide-Web Documents", *Information & Management* (31) 5, pp. 235-249.

Hao, X., J. T. L. Wang, and P. A. Ng (1996) "Information Extraction from the Structured Part of Office Documents", *Information Sciences* (91) 3-4, pp. 245-274.

Iocchi, L. (1999) "The Web-OEM Approach to Web Information Extraction", *Journal of Network and Computer Applications* (22) 4, pp. 259-269.

Jain, H. K. and S. Purao (1991) "Distributed Application Development: SDLC revisited", *Information & Management* (20) 4, pp. 247-255.

Jiang, J. J. and G. Klein (1999) "Project Selection Criteria by Strategic Orientation", *Information & Management* (36) 2, pp. 63-75.

Jones, J. I. and K. R. Morrison (1993) "Work Flow and Electronic Document Management", *Computers & Industrial Engineering* (25) 1-4, pp. 527-531.

Lambrix, P. and L. Padgham (2000) "Conceptual Modeling in a Document Management Environment Using Part-Of Reasoning in Description Logics", *Data & Knowledge Engineering* (32) 1, pp. 51-86.

Marble, R. P. (2003) "A System Implementation Study: Management Commitment to Project Management", *Information & Management* (41) 1, pp. 111-123.

Schniederjans, M. J. and R. L. Wilson (1991) "Using the Analytic Hierarchy Process and Goal Programming for Information System Project Selection", *Information & Management* (20) 5, pp. 333-342.

Tian, Q. J., J. Ma, J. Z. Liang, C. W. Kwok et al. (2004) "An Organizational Decision Support System for Effective R&D Project Selection", *Decision Support Systems*.

Zantout, H. and F. Marir (1999) "Document Management Systems from Current Capabilities Towards Intelligent Information Retrieval: An Overview", *International Journal of Information Management* (19) 6, pp. 471-484.

ABOUT THE AUTHORS

Ou Liu is currently pursuing his Ph.D. degree in the Department of Information Systems, City University of Hong Kong. His current research interests include electronic document management, organizational decision support systems, and model management.

Jian Ma is Associate Professor in the Department of Information Systems, City University of Hong Kong. Previously, he was a Lecturer in the School of Computer Science and Engineering at the University of New South Wales, Australia. His research areas include web-based decision support systems, and object-oriented methods for information system development. His research is published in *IEEE Transactions on Engineering Management, IEEE Transactions on Education, IEEE Transactions on Systems, Man* and Cybernetics, and *Decision Support Systems*.

Ji-Ye Mao is Professor in the School of Business, Renmin University of China. He received his Ph.D. in management information systems from the University of British Columbia. His research interests include human-computer interaction such as design and evaluation of information systems, and user-centered design. His research appears in journals such as *Journal of Management Information Systems, International Journal of Human-Computer Studies,* and *International Journal of Human-Computer Interaction.*

Ron Chi-Wai Kwok is Assistant Professor in the Department of Information Systems, City University of Hong Kong, where he received his Ph.D. He was an Assistant Professor at the School of Management, the State University of New York at Binghamton before returning to City University in 2002. His current research interests include group support system (GSS), technology-based learning, leadership, IS outsourcing and electronic commerce. His papers are published in *Journal of Management Information Systems*, *Communications of ACM, IEEE Transactions on Systems, Man, and Cybernetics*, *Decision Support Systems, European Journal of Information Systems*, *Information & Management.*

Dong Li is a senior manager at the Information Center, National Natural Science Foundation of China. She has been involved in ISIS development and administration for the past several years.

Copyright © 2004 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints or via e-mail from ais@aisnet.org

Communications of the Association for Information Systems

ISSN: 1529-3181

EDITOR-IN-CHIEF

Paul Gray

Claremont Graduate University

AIS SENIOR EDITORIAL BOARD

Detmar Straub Vice President Publications Georgia State University	Paul Gray Editor, CAIS Claremont Graduate University	Sirkka Jarvenpaa Editor, JAIS University of Texas at Austin
Edward A. Stohr Editor-at-Large Stevens Inst. of Technology	Blake Ives Editor, Electronic Publications University of Houston	Reagan Ramsower Editor, ISWorld Net Baylor University

CAIS ADVISORY BOARD

Gordon Davis University of Minnesota	Ken Kraemer Univ. of Calif. at Irvine	M.Lynne Markus Bentley College	Richard Mason Southern Methodist Univ.
Jay Nunamaker University of Arizona	Henk Sol Delft University	Ralph Sprague University of Hawaii	Hugh J. Watson University of Georgia

CAIS SENIOR EDITORS

Steve Alter U. of San Francisco	Chris Holland Manchester Bus. School	Jaak Jurison Fordham University	Jerry Luftman Stevens Inst.of Technology

CAIS EDITORIAL BOARD

Tung Bui University of Hawaii	Fred Davis U.ofArkansas, Fayetteville	Candace Deans University of Richmond	Donna Dufner U.of Nebraska -Omaha
Omar El Sawy Univ. of Southern Calif.	Ali Farhoomand University of Hong Kong	Jane Fedorowicz Bentley College	Brent Gallupe Queens University
Robert L. Glass Computing Trends	Sy Goodman Ga. Inst. of Technology	Joze Gricar University of Maribor	Ake Gronlund University of Umea,
Ruth Guthrie California State Univ.	Alan Hevner Univ. of South Florida	Juhani Iivari Univ. of Oulu	Claudia Loebbecke University of Cologne
Munir Mandviwalla Temple University	Sal March Vanderbilt University	Don McCubbrey University of Denver	Emannuel Monod University of Nantes
John Mooney Pepperdine University	Michael Myers University of Auckland	Seev Neumann Tel Aviv University	Dan Power University of No. Iowa
Ram Ramesh SUNY-Buffalo	Kelley Rainer Auburn University	Maung Sein Agder Univ. College,	Carol Saunders Univ. of Central Florida
Peter Seddon University of Melbourne	Paul Tallon Boston College	Thompson Teo Natl. U. of Singapore	Doug Vogel City Univ. of Hong Kong
Rolf Wigand Uof Arkansas,LittleRock	Upkar Varshney Georgia State Univ.	Vance Wilson U.Wisconsin,Milwaukee	Peter Wolcott U. of Nebraska-Omaha

DEPARTMENTS

Global Diffusion of the Internet. Editors: Peter Wolcott and Sy Goodman	Information Technology and Systems. Editors: Alan Hevner and Sal March
Papers in French Editor: Emmanuel Monod	Information Systems and Healthcare Editor: Vance Wilson

ADMINISTRATIVE PERSONNEL

Eph McLean AIS, Executive Director Georgia State University	Samantha Spears Subscriptions Manager Georgia State University	Reagan Ramsower Publisher, CAIS Baylor University