5 INTERFACE STANDARDS AND STRATEGY

EASI/ED will communicate with a large and diverse population of customers and partners, and a wide variety of technologies will be employed to facilitate this communication. In addition, it is anticipated that the data received by EASI/ED will be distributed to multiple databases for processing. The purpose of this section is to recommend public and/or proprietary standards for each of the interfaces EASI/ED has with its customers and partners, and then to recommend a strategy for synchronizing data within EASI/ED. Figure 5-1 illustrates these concepts.

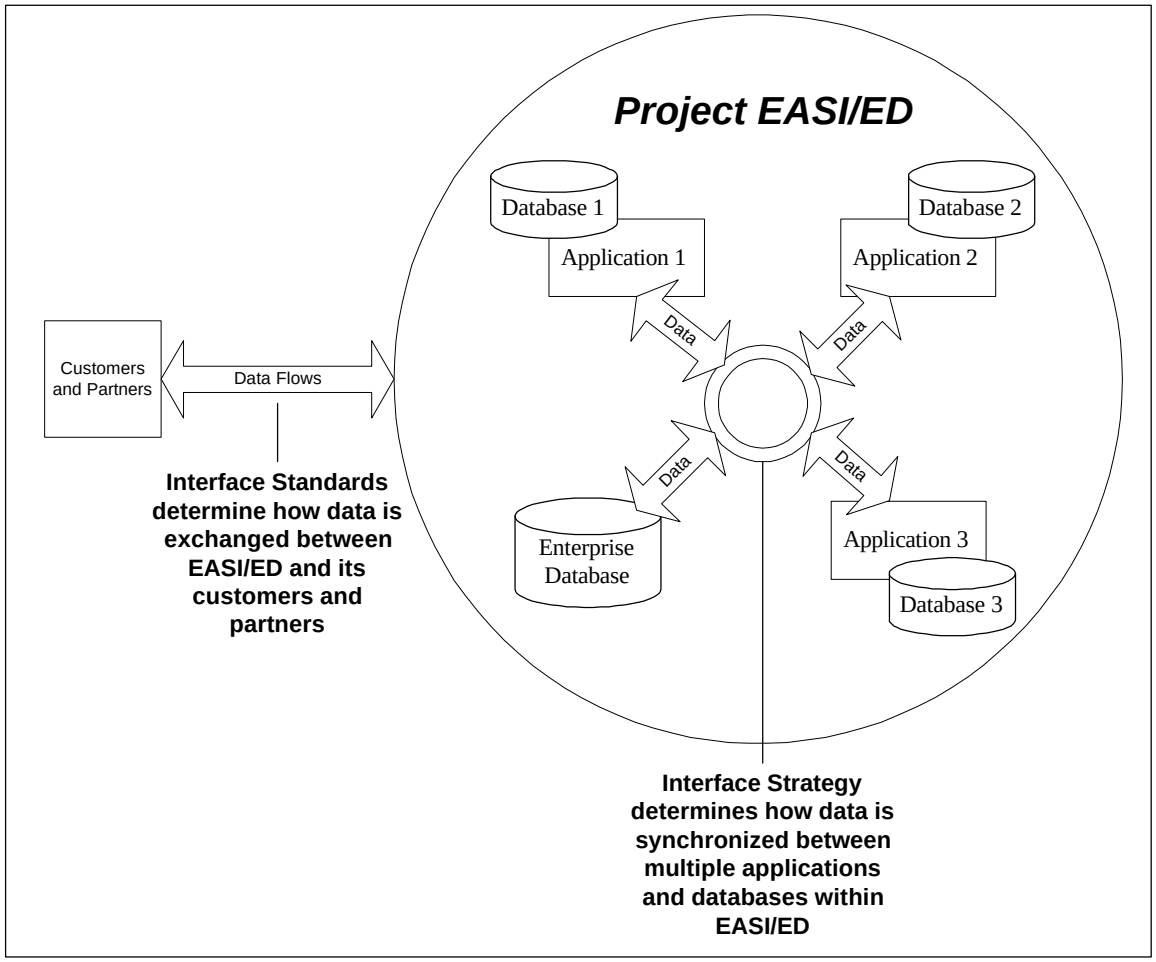


Figure 5-1: Interface Standards and Strategy

Subsection 5.1 presents the Project EASI/ED Interface Standards. Subsection 5.2 presents the Project EASI/ED Interface Strategy.

5.1 Project EASI/ED Interface Standards

The purpose of this subsection is to present the standards recommended for the data flows, or interfaces, between EASI/ED and its partners and customers. The interfaces used for this analysis are the data flows defined in the *Project EASI/ED ASDD*, Version 2.0. For each interface defined in the ASDD, one or more interface standards were chosen.

EASI/ED will interface with partners and customers of different sizes, technical capabilities, requirements, and needs. In addition, the interfaces themselves contain data representing a wide variety of business purposes. Any future development of the Project EASI/ED system must take these factors into consideration. This document serves as a foundation upon which future developers of the EASI/ED system can build to design and develop the interfaces for EASI/ED.

When determining the interface standards for EASI/ED, the needs and capabilities of EASI/ED's partners and customers, the type of data being transmitted, and the appropriate transmission media to be used were considered. The factors that influenced the decision to recommend a specific standard for a specific interface, i.e., a data flow between a specific trading partner and EASI/ED, included the following:

- Technical and administrative capabilities of the trading partner with whom EASI/ED will be communicating;
- Volume of data flow between the trading partner and EASI/ED;
- Business purpose of the data being transmitted; and
- Frequency of data flow between EASI/ED and the trading partner.

Using these factors, the data interchange transmission media and standards were determined for each of the interfaces EASI/ED has with its partners and customers. Whenever possible, mainstream, widely accepted interface standards were chosen. This approach helps reduce the development and implementation efforts for EASI/ED and its partners and customers by leveraging proven transmission media and standards.

All the interface standards recommended in this document relate to an associated transmission medium, i.e., the mechanism through which the data will be transferred. For example, for information being transferred via the World Wide Web (WWW), a recommended standard is hypertext markup language (HTML), a widely accepted and used standard for Web-based documents. Therefore, interface standard recommendations are presented in this document in the context of the applicable transmission medium. Each transmission medium is described at a high level and then each of the applicable interface standard(s) related to that medium is described.

Subsection 5.1.1 presents the assumptions made for the interface standards and strategy analysis. The remainder of this subsection is organized by the recommended transmission media and their associated interface standards. Figure 5-2 presents the six transmission media and the associated interface standards that are recommended for Project EASI/ED.

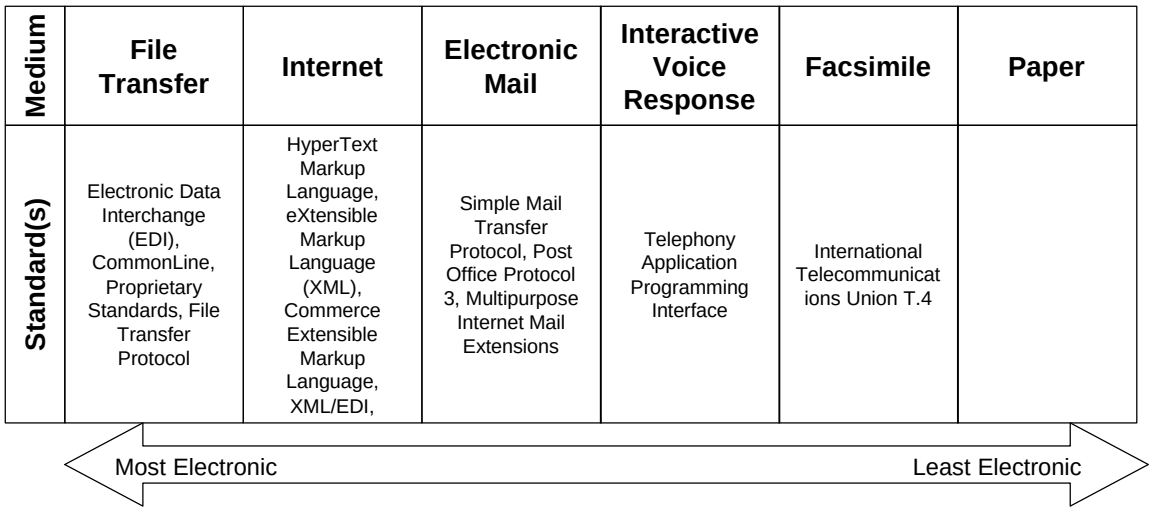


Figure 5-2: Transmission Media and Standards

Figure 5-2 presents the transmission media on a continuum from most to least "electronic". Of the transmission media recommended, file transfer lends itself to allowing EASI/ED and its partners and customers to communicate with minimal manual processing. Paper represents the least electronic and most manually intensive transmission medium.

In cases where there are more than one applicable interface standard for the transmission medium being discussed, the transmission medium subsection is further divided into subsections that address each of the applicable interface standards. The interface standards are presented in the following subsections:

- Subsection 5.1.2 File Transfer
- Subsection 5.1.3 Internet
- Subsection 5.1.4 Electronic Mail
- Subsection 5.1.5 Interactive Voice Response
- Subsection 5.1.6 Facsimile
- Subsection 5.1.6 Paper

5.1.1 Assumptions

The following represents the assumptions made when recommending the Project EASI/ED interface standards.

1. The data flows defined in the *Project EASI/ED ASDD* represent all of the interfaces that EASI/ED has with its customers and partners.

2. The standards considered as candidates for the interface standards were those identified in the *Project EASI/ED Common Operating Environment (COE) Document*, July 10, 1998.

3. EASI/ED's partners and customers have varying degrees of technical and administrative capabilities that EASI/ED must accommodate.

4. The volume (the total number of records for a specific data flow received or sent by EASI/ED) for each data flow or transaction was assessed to be either high, medium, or low. The following volumes were used for each rating.

    - High - More 1 million transactions/year
    - Medium - 40,000 – 999,999 transactions/year.
    - Low - Less than 40,000 transactions/year.

    Data flow volumes were those identified by the performance requirements analysis documented in Appendix E.

5. The frequency (how often a data flow is received or sent by EASI/ED) of the data flows were determined to be high, medium, or low. The following frequencies were used for each rating.

    - High - Daily or Weekly
    - Medium - Monthly
    - Low - Less than monthly

    The frequency of the data flows were those identified in the *Project EASI/ED ASDD*.

5.1.2 File Transfer

File transfer is a medium of data exchange by which bulk data is transmitted electronically between trading partners according to specific, previously agreed upon standards. Files can be transmitted across private networks, direct connect lines, or the Internet. The file transfer standards identified and recommended for EASI/ED are Electronic Data Interchange (EDI), CommonLine, Proprietary Standards, and File Transfer Protocol (FTP). The remainder of this subsection presents each of these standards and their applicability to EASI/ED interfaces.

5.1.2.1 Electronic Data Interchange

Description: Electronic Data Interchange (EDI) can be defined as the computer-to-computer exchange of business information in a nationally accepted, standard format. In the United States, the formats followed are those promulgated by the American National Standards Institute (ANSI), Accredited Standards Committee (ASC) X12, commonly referred to as X12. In simplest terms, X12 requires that data are formatted into "transaction sets", the electronic equivalent of a paper business document. For example, the 810 Invoice transaction set is, as its name implies, the equivalent of a paper invoice. Transaction sets are transmitted between trading partners primarily through Value Added Networks (VANs) (interconnected, secure networks used for the transmission of EDI transaction sets). Therefore, EDI prescribes both the medium of interface and the format of the files being transmitted.

Benefits: The purpose of EDI, and any file transfer method, is to replace end-to-end manual and paper-based processes with electronic ones. With standards like EDI, trading partners receive business data knowing exactly where the data is within the transaction set and what the purpose of all the data elements are. The primary benefits of EDI include:

- Speed;
- Elimination of Data Entry;
- Reduction of Errors;
- Standardized Data; and
- Reduction in Costs.

Limitations: Implementing EDI can be a relatively complex endeavor that may require substantial human and financial resources. Therefore, for partners with a relatively small volume or frequency of transactions, EDI may not be cost beneficial. In addition, EDI is viewed by some as an outdated technology. However, EDI is one of the only proven mechanisms by which to transfer bulk data in a nationally accepted standard format. Until other bulk file transfer technologies and standards become more mature, EDI is recommended. However, prior to implementing EDI, the organization must weigh the benefits and costs.

Recommendations: ED has embarked upon several initiatives to implement EDI in its financial aid delivery systems and processes and has implied that EDI is a technology that it plans to pursue on a larger scale. Therefore, wherever possible and where it makes technical and business sense, EDI is recommended as an interface standard. EDI is primarily recommended for those interfaces that are both high to medium volume and high to medium frequency. In addition, EDI is recommended as an interface standard for those partners that have high technical capabilities. The table below identifies the specific transaction sets that are identified for possible usage by EASI/ED. The table also summarizes the type of information for which each of the transaction sets will be used.

EDI Transaction Set (Version 4010)[1]	EASI/ED Usage
135 Student Origination Request and Response	Recommended for the exchange of student loan origination and funding information with Schools, Lenders, Guaranty Agencies, State Agencies, and Private Scholarship Agencies. This transaction set can be used to send or receive origination information for any type of student loan.
144 Student Loan Transfer and Status Verification	Recommended for the exchange of data related to changes in loan status and ownership with Guaranty Agencies, Lenders, and the US Department of the Treasury.
190 Student Enrollment Verification	Recommended for the exchange of data related to the deferment of loans with Guaranty Agencies and Lenders.
191 Student Loan Pre-Claims and Claims	Recommended for the exchange of data related to the assignment and updating of defaulted loans with the US Department of Housing and Urban Development (HUD), Lenders, the US Postal Service (USPS), Collection Agencies, and Guaranty Agencies.

Table 5-1: EDI Transaction Set Use

[1] ANSI X12 Standards Version 4010, December 1997.

EDI Transaction Set (Version 4010)	EASI/ED Usage
198 Loan Verification Information	Recommended for the exchange of data related to aid eligibility, changes in loan status, and loan consolidation with Lenders, Private Scholarship Agencies, the Selective Service, the Immigration and Naturalization Service (INS), Schools, State Unemployment Agencies, Guaranty Agencies, the Direct Loan Lockbox, and the National Payment Center.
810 Invoice	Recommended to request funds and to exchange data about the disbursement of aid by Schools and Lenders.
814 General Request, Response or Confirmation	Recommended for the request of participant demographic information by the Department of Veterans Affairs.
820 Payment Order/Remittance Advice	Recommended for the exchange of data related to the disbursement of aid, the disbursement of administrative cost and expense allowances, and school expenditures with Schools, ED Regional Offices, Lenders, ED/Chief Financial Office (CFO), and ED/Office of Postsecondary Education (OPE)/Accounting and Financial Management Service (AFMS).
821 Financial Information Reporting	Recommended for the exchange of detailed financial data related to income verification, Campus Based Programs authorization amounts, loan repayment information, school audit and program review data, school tuition and aid information, school financial statements, reconciliation information, aid awards, excess cash, delinquent debt, and credit bureau reports with ED/OPE/AFMS, Schools, Health and Human Services (HHS), the Department of Veterans Affairs, the National Payment Center, ED/CFO, the Direct Loan Lockbox, Credit Bureaus, the Department of Defense, and Guaranty Agencies.
824 Application Advice	Recommended for transmitting edit results of origination and disbursement records with Schools.
838 Trading Partner Profile	Recommended for the exchange of data related to Quality Control Reviews with the Office of the Inspector General.

Table 5-1: EDI Transaction Set Use (continued)

Detailed information on the recommended EDI transaction sets on data-flow-by-data-flow basis is presented in Appendix G.

5.1.2.2 CommonLine

Description: CommonLine was developed in an effort to standardize Federal Family Education Loan Program (FFELP) loan-processing procedures and software. The standard was established in February 1995 by a group of FFELP service providers that were brought together by the National Council of Higher Education Loan Programs (NCHELP). These established standards

and procedures were designed to simplify loan application, guarantee, and disbursement processes. The standards provide for standardized file formats and communication services. By utilizing common files and procedures, it enables schools and service providers (guarantors, lenders, and servicers) to communicate and conduct business electronically, through network services provided by CompuServe.

Benefits: Because CommonLine is widely used within the financial aid community, many lenders, guaranty agencies, and schools are familiar with its use. This familiarity can be leveraged by EASI/ED to facilitate interaction with these organizations.

Limitations: CommonLine has four transaction files: Application Send, Change Transaction, Disbursement Roster, and Response. Therefore, not all types of interfaces defined for Project EASI/ED between FFELP partners and customers can be accommodated by CommonLine.

Recommendations: The trend within the FFELP community is for closer interaction between the involved parties to create a single point of contact for FFELP loan information. This trend can be seen in various initiatives such as the establishment of Educational Loan Management (ELM) Resources. All these initiatives rely, to some extent, upon the usage of CommonLine. Therefore, to allow EASI/ED to participate in single point of contact initiatives, CommonLine is a recommended interface standard. The following table provides a summary of how CommonLine may be used by EASI/ED.

CommonLine File (Release 4)[2]	EASI/ED Usage
Application Send File	Recommended for the exchange of data related to the origination of FFELP loans and multi-year promissory notes of FFELP loans with Lenders.
Change Transaction File	Recommended for the exchange of data related to changes to FFELP loan demographic information with Lenders.
Disbursement Roster File/Disbursement Roster Acknowledgment File	Recommended for the exchange of data related to disbursements of FFELP loans, including funding requests and changes to disbursements, with Lenders.

Table 5-2: CommonLine File Usage

Detailed information that presents all the interfaces for which CommonLine is identified as the standard is presented in Appendix G.

5.1.2.3 Proprietary Standards

Description: Proprietary standards are those standards that are prescribed by an organization that determine how its trading partners may interface with it. In most cases, the organization determines both the file formats and the means of transmission, e.g., private networks or the Internet.

Benefits: Proprietary standards allow for the electronic transmission of information rather than through paper, thus reducing processing costs and times for both EASI/ED and its partners.

[2] CommonLine Release 4, January 29, 1999.

Limitations: Because proprietary standards by definition are proprietary, EASI/ED may have to accommodate numerous standards from numerous organizations and make changes based upon any updates or changes prescribed by the organization that owns the proprietary standard.

Recommendations: The following partners were identified as those that may require proprietary standards for their interface with EASI/ED: IRS, Credit Bureaus, Bureau of Labor Statistics, Selective Service, Social Security Administration, Third Party Repositories, Department of Labor, Immigration and Naturalization Service, USPS, Collection Agencies, Department of Defense, State Agencies, HUD, Department of the Treasury, and the National Student Loan Clearinghouse. The type of data being transmitted is wide and varied. A complete listing of all the data flows for which Proprietary Standards are the recommended interface standard is presented in Appendix G.

5.1.2.4 File Transfer Protocol

Description: FTP allows for the transfer of text and binary files to and from computer systems running FTP servers. FTP is a batch-oriented protocol: a user must transfer the entire file to a local disk and then open it with an appropriate application, rather than accessing the file interactively. Until the advent of Multipurpose Internet Mail Extensions (MIME) (please refer to subsection 5.1.4.3 for further information on MIME) and other binary electronic mail (e-mail) encoding schemes, use of a mutually accessible FTP server was the only practical way for users to exchange files over the Internet.

Benefits: Today, FTP is widely available and used. Most desktop operating systems now include a basic FTP client and clients with a graphical FTP interface also are available. In addition, FTP functionality is built into many Web browsers.

Limitations: FTP provides a standard for transmitting files between trading partners. As such, it is a transmission standard. However, FTP does not prescribe the format of the actual data being transmitted. The file could be in a text format or it may be in a structured format such as a Portable Document Format (PDF) or Microsoft Word file. Therefore, prior to implementing FTP, the trading partners must agree to how the actual data will be formatted in the file.

Recommendations: FTP is considered as an interface standard for EASI/ED because it leverages a readily accessible and well-used interface method. It is recommended for use primarily with interfaces that have relatively low volumes and frequency, but large data sets. For example, a School may need to send information on 1,000 student loans to EASI/ED. The relatively large size of the data set would preclude the use of an interactive medium. With an interactive medium, both the sending and receiving organizations' systems would have to be on-line. The sending organization would send a request (the loan information) and the receiving organization would have to respond to that request. If the sending organization is trying to transmit a very large data set, the time required for the recipient to receive and respond to the request may be quite long. Therefore, an interactive medium may not be practical. However, the volume or frequency is not large enough to warrant the use of EDI or some other bulk file transfer standard that would require a relatively high implementation costs for ED and its partners. The type of data being recommended for FTP usage is wide, varied, and used for virtually every functional area within EASI/ED. In addition, FTP is recommended as an interface standard for a wide variety of partners. A complete listing of all the interfaces for which FTP is the recommended standard is presented in Appendix G.

5.1.3 Internet

The Internet is a vast collection of local, regional, and national networks that are linked together to exchange data and to distribute processing tasks. The networks are interconnected following Transmission Control Protocol/Internet Protocol (TCP/IP). Accessing the Internet is completed through an Internet Service Provider (ISP), e.g., America On-line, that will supply a user with a TCP/IP connection via their hardware. Users utilize specific software called Web browsers, e.g., Microsoft Internet Explorer, to access the Internet from their computers. The Web browser translates the protocol language Hypertext Transport Protocol (HTTP) into readable information or documents, such as text, pictures, and video for the user. The information shared is stored in information packets called Web documents. These documents are created commonly using HTML. It should be noted that there is a distinction between the Internet and the World Wide Web. As stated earlier, the Internet is the collection of the underlying, interconnected networks. The World Wide Web is a set of "domains" or applications on the Internet that communicate with each other using HTTP.

The Internet is one of EASI/ED's preferred interface media. Customers and partners capable of connecting to the Internet should communicate with EASI/ED via the World Wide Web. Interfaces involving individual-level interactions, such as the on-line completion of a FAFSA, status updates, updates of demographic information, and on-line counseling are good candidates for the Internet. In addition, EASI/ED could use the Web as a repository for forms and/or information-only documents in the form of downloadable files.

The Internet is a rapidly changing technology. It has evolved from being a basic information tool to a strong interactive information exchange, research tools, and a business process medium. The remainder of this subsection discusses traditional and emerging Internet standards, primarily for "marking-up" documents for presentation on the Web. The emerging standards are presented primarily for informational purposes and to highlight those technologies that may be of particular interest or provide future benefits to EASI/ED. However, because they are emerging standards, they are not recommended as interface standards for the near term.

5.1.3.1 Hypertext Markup Language

Description: HTML is a simple markup language used to create hypertext documents that are portable from one platform to another. Through the use of tags, markup languages define the manner by which the document should be presented by an application. HTML is one of many markup language standards.

HTML is a World Wide Web Consortium (W3C) specification. The W3C is an international industry consortium, led by Tim Berners-Lee, creator of the World Wide Web. The purpose of the W3C is to "lead the World Wide Web to its full potential by developing common protocols that promote its evolution and ensure its interoperability."[3] Services provided by the Consortium include providing a repository of information about the World Wide Web for developers and users, providing reference codes for implementations to embody and promote standards, maintaining and recommending HTML standards, and reviewing various prototype and sample applications to demonstrate use of new technology. The W3C is vendor neutral, working with the global community to produce specifications and reference software that is made freely available throughout the world.

[3] http://www.w3.org/Consortium/

Benefits: HTML is recommended by the W3C and is the most widely used and supported markup language for the purpose of displaying information on the World Wide Web.

Limitations: HTML has one hypertext linking type and developers are restricted in designing documents due to rigid Document Type Definitions (DTDs) for HTML.

Recommendations: HTML 3.2 is an W3C specification developed in early 1996 together with Internet involved organizations such as International Business Machine (IBM), Microsoft, Netscape Communications Corporation, Novell, SoftQuad, Spyglass, and Sun Microsystems. HTML 3.2 is not the latest version of HTML. HTML 4.0 is the most current version of the language, builds on earlier versions, specifically HTML 3.2 and HTML 2.0, and includes new or enhanced features. However, because it is the most recent version of HTML, very few Web browsers can support it currently. Therefore, for the near term, EASI/ED should support HTML 3.2 as its HTML standard. As HTML 4.0 becomes more common place and more browsers are able to support it, EASI/ED should re-evaluate, at that time, whether HTML 4.0 should be supported as an interface standard.

For the near term, until other Web-based technologies, such as those discussed below, evolve and become more mature, all the Web-based interfaces identified for Project EASI/ED leverage HTML as the interface standard. The types of data flows for which HTML is identified as a standard are wide and varied. In addition, HTML is a recommended as an interface standard for virtually all of EASI/ED's partners and customers. A detailed listing of all the interfaces for which HTML is identified as the standard is presented in Appendix G.

5.1.3.2 eXtensible Markup Language

Description: The eXtensible Markup Language (XML) is not in itself a single markup language like HTML. It is instead a set of rules for designing a markup language, or adding extensions to present markup languages like HTML. XML was developed using the Standard Generalized Markup Language (SGML), which is codified as ISO 8879, the standard from the International Organization for Standards (ISO) for defining descriptions of the structure and content of different types of electronic documents. HTML itself is a "document type" defined in SGML, and is being reformulated into XML to make it more adaptable across all platforms. According to the XML specification, XML was designed "to make it easy and straightforward to use SGML on the Web: easy to define document types, easy to author and manage SGML-defined documents, and easy to transmit and share them across the Web." Figure 5-3 illustrates the relationships between SGML, HTML, and XML.

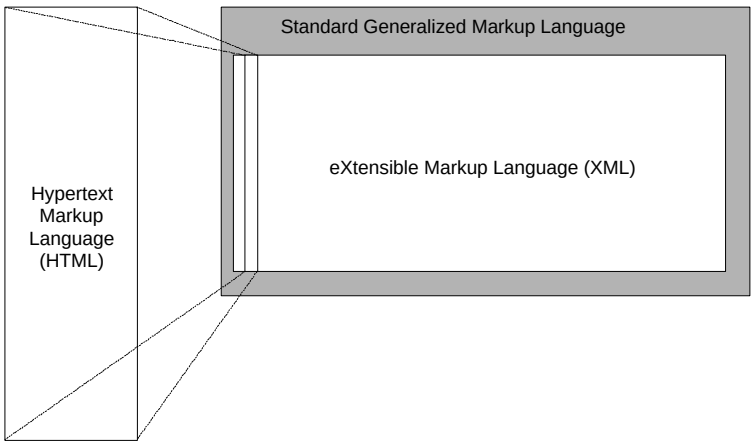


Figure 5-3: SGML, XML and HTML

As shown in Figure 5-3, SGML can be used to describe literally thousands of different document types, and XML is a subset of SGML that is optimized for defining new document types and supporting delivery and interoperability over the Web. HTML is just a single document type among the countless number that can be defined using SGML.

Benefits: Some of the benefits of XML include the following.

- Web developers can design their own document types, as opposed to being restricted to using HTML. DTDs can be tailored to a specific purpose, so that the cumbersome "tricks" that are required of HTML to achieve special effects should no longer be required. Therefore, Web developers can invent their own markup elements.

- The hypertext linking features of XML are vastly superior to those of HTML. Specifically, unlike HTML, XML allows Web developers to specify the *type* of link using one of the following five types: Simple, Extended, Locator, Group, or Document. Each of these types has multiple attributes that greatly refine the control that developers have.

- XML files are true SGML files as well, allowing them to be utilized in other environments beside the Web.

Limitations: While XML provides users with great flexibility, it may pose problems for interchange and software integration. For example, the possibility exists for documents to use the same tag name in different contexts. A <PART> tag in an illustrated parts catalog identifies something quite different than a <PART> in a dramatic play. Within a single document, the term "title" may refer to the document itself, the name of a book, and the formal appellation associated with its author (e.g., "Dr."). This potential collision over different uses of the same name poses problems for anyone developing XML- based software and applications.

Recommendations: XML is still in the experimental stage, and there may be subtle differences as to how various Web browsers interpret XML documents. Widespread XML usage will only be possible when XML enabled browsers are as robust and widely available as HTML enabled browsers. The open-source beta of Netscape 5 (code named "Gecko") and Beta 2 of Microsoft Internet Explorer 5 both nominally support XML. However, it appears that Microsoft is retaining

its own proprietary extensions, which has a potentially dramatic impact on the portability of XML documents. Therefore, for the short term, XML is not recommended as an interface standard for EASI/ED. XML is described and presented in this document as a standard whose development EASI/ED should monitor. As XML becomes more widely used and accepted, EASI/ED should re-evaluate when and where it may be applicable and beneficial.

5.1.3.3 Commerce Extensible Markup Language

Description: Commerce Extensible Markup Language (cXML) is an open Internet-based standard that facilitates Electronic Commerce (EC) over the Web. It is a subset of XML that defines standard sets of data for Internet-based EC. cXML is a new set of DTDs for XML. It has been developed in conjunction with more than 40 leading EC business companies such as Cisco Systems, VISA International, US West, Bristol-Myers Squibb, the State of California, Chevron, Merck & Co., and many others.

cXML is an explicit meta-language used to describe the characteristics of items available for sale. It enables the development of "intelligent shopping agents" that help facilitate the laborious and tedious tasks of corporate purchasing. By programming the characteristics of an item sought into request messages and releasing these messages to the network, the request will return exactly what the user wants or nothing at all, which in itself is sometimes important to know. cXML can be thought of as "bar coding" for the Web but with a far richer set of attributes to uniquely identify and describe products.

Benefits: cXML provides real time information exchange. It reduces on-line business trading costs by facilitating the exchange of content and transactions over the Internet. cXML has a lower-cost of implementation than some traditional EC mechanisms such as EDI because of its XML base and ability to leverage existing HTML infrastructure and software.

cXML provides an infrastructure that will streamline the process of digitally exchanging catalog content and transactions in a secure manner. cXML supports all supplier content and catalog models, including buyer-managed, supplier-managed, content management services, electronic marketplaces, and Web-based sourcing organizations. This will allow suppliers to provide customers with selective access to personalized catalog content while maintaining their unique branding and competitive differentiation.

Additionally, cXML defines a request/response process for the exchange of transaction information. These business processes include purchase orders, change orders, acknowledgments, status updates, ship notifications, and payment transactions.

Limitations: Since cXML is still undergoing development, it is difficult to evaluate what the limitations will be. However, the limitations presented for XML also apply to cXML.

Recommendations: cXML is described and presented in this document as a standard whose development EASI/ED should monitor. If cXML indeed fulfills the specifications presented by the developers and becomes widely used and accepted, EASI/ED should evaluate this standard then and decide were it will be beneficial and applicable.

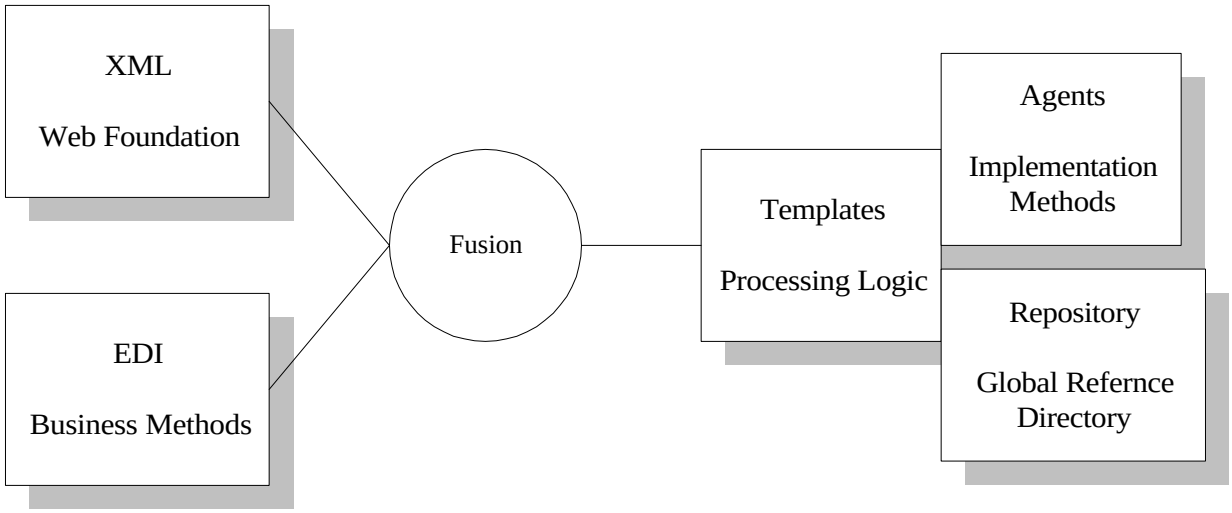
5.1.3.4 XML/EDI

Description: XML/EDI is an emerging technology that leverages the Internet, EC, and EDI to form a framework that provides formal interfaces for commercial EC components to inter-operate. These interfaces will be open, yet standardized. The business model may be comprised

of ad-hoc interactions between small groups, or agreed upon national or international frameworks.

XML/EDI is the fusion of five technological components: XML, EDI, Templates, Agents, and Repositories. The combination of these capabilities is what makes XML/EDI so powerful.

Each component adds unique tools that leverage the other components. Figure 5-4 illustrates the interrelationships of these components.



Source: XML/EDI Group, August 1997

Figure 5-4: XML/EDI: The Fusion of Technologies

- **XML** provides the format for the structured presentation of the document on the Web. It provides a file format for the data, and a schema for describing the data structure

- **EDI** provides the structure or order of the actual data. In other words, all data is ordered according to EDI transaction sets formats.

- **Templates** are the rules that provide a basis for processing an XML document that contains a structured EDI transaction. They either are referenced by or travel along inside the XML document as a special section, and can be easily read and interpreted because they look like a spreadsheet or a table in layout and content. Templates are supplemented by DTDs. DTDs enable transaction interoperability and rules-enabled processing which may include presentation of transactions. DTDs let two organizations understand each other's data. Templates, on the other hand, define what happens to this data.

- **Agents** are basically translators to interpret XML/EDI documents and can be likened to a browser. As a browser is necessary to interpret traditional HTML documents, Agents are required to interpret and process XML/EDI documents. Agents interpret the Templates to perform the work needed, interact with the transaction and the user to create new templates for each new specific task, or look up and attach the right template from repositories for existing jobs. They also can reference DTDs to determine display characteristics for forms.

- **Repositories** are directories referenced by agents to determine how XML/EDI documents should be interpreted. The directory contains tags and templates used by the XML/EDI document that are required for the processing of the document. The agent queries this directory to determine the processing logic. When any new tag or a template is developed, the processing logic or steps are placed in these repositories. These repositories are primarily responsible for providing standardized ways to understand XML/EDI documents.

Benefits: The primary strength of XML/EDI is that it can encode a document's information very precisely and in a very rich structure. Through the use of XML tags and DTDs, XML/EDI transactions are self-describing. XML/EDI documents serve as a structure that contains data and will include instructions on how the transaction should be processed or displayed. Much like when a file is double clicked within Microsoft Windows Explorer and the application automatically launches, an XML/EDI document will have transaction status self-contained for applications (or users) to set/interrogate. The document will be able to find the application (or user) by using the search, classifying, and routing mechanisms embedded in the templates and interpreted by the agent instead of the applications (or user) having to find it. XML/EDI documents do not need to be exchanged via the Internet since XML/EDI does not rely on Internet browsers to communicate. These documents can be transported using e-mail, over a network, or by any file transfer means.

Limitations: XML/EDI is fairly new technology and has not been implemented in many organizations.

Recommendations: XML/EDI is still a relatively new technology and is not yet widely used. It is being presented in this document as a technology whose development EASI/ED should monitor. Should XML/EDI become more widely used, EASI/ED should determine, at that time, whether XML/EDI is a technology that should be implemented. If should be noted, however, that EASI/ED may use XML/EDI for data exchange wherever EDI is recommended as an interface standard in subsection 5.1.2.1. In addition, all high volume interfaces that require batch processing by either ED or its partners are good candidates for XML/EDI implementation.

5.1.4 Electronic Mail

Description: Electronic mail (e-mail) is one of the most commonly used electronic communications media. It allows an individual or an organization to exchange messages electronically using the Internet or other mail transfer protocols. For most e-mail, a file can also be attached to an e-mail text message.

Benefits: E-mail provides a relatively easy way for partners and customers to electronically communicate with EASI/ED, especially for those partners and customers that may not have access to more sophisticated electronic transmission media.

Limitations: Because e-mail is an electronic transmission medium, it requires participants to have access to, at a minimum, a personal computer, an e-mail application, and an Internet or other mail transfer service provider.

Recommendations: EASI/ED's partners and customers with the technical capabilities to use e-mail could use this facility to send and/or receive requests for information, responses, and/or confirmations in the financial aid process. Some examples of the types of interfaces include Cohort Default Rate appeal and response between Schools and EASI/ED, and customer service

requests and responses from a customer. E-mail is recommended as an interface standard for a wide variety of partners and customers.

The following subsections discuss the recommended standards for e-mail message transmission and transmission of a file(s) attached within the body of the message.

5.1.4.1 Simple Mail Transfer Protocol

Simple Mail Transfer Protocol (SMTP) provides the protocol for the transmission of e-mail directly from the sending user's host machine to the receiving user's host machine. The purpose of SMTP is to facilitate the transfer of e-mail reliably and efficiently. SMTP provides the protocol by which e-mail messages are encoded by the sender and subsequently decoded by the recipient.

In all cases where the interface between EASI/ED and its trading partner is through e-mail, SMTP is the recommended interface standard. SMTP is used to transmit a wide variety of data between EASI/ED and a large majority of its partners and customers. A detailed listing that provides all the interfaces for which SMTP is the recommended interface standard is presented in Appendix G.

5.1.4.2 Post Office Protocol 3

Post Office Protocol 3 (POP3) is the most widely implemented protocol for the retrieval of Internet based mail messages by a mail client and is supported by the major Internet mail client applications. It is used in conjunction with SMTP. POP3 provides the mechanism for a server to store and serve mail for various client machines that are not connected to the Internet 24 hours a day. In essence, POP3 provides the ability to create an electronic post office box as e-mail is held in the POP server until the user logs-in and retrieves it.

POP3 is not an interface standard, per se, and, therefore, is not recommended as an interface standard for EASI/ED. It is a standard used by the recipient of an e-mail message to retrieve and read the e-mail. It is being presented in this document to provide a holistic view of the standards necessary to send and receive e-mail. POP3 will be used by EASI/ED to receive e-mail messages from its partners and customers.

5.1.4.3 Multipurpose Internet Mail Extensions

While SMTP provides a standard to send e-mail message, Multipurpose Internet Mail Extensions (MIME) provides a standard for how documents can be attached to e-mail messages.

SMTP and MIME provide a well-supported set of standards to establish interoperability. SMTP and MIME, used in conjunction, will provide sufficient product support and interoperability to build a messaging system to meet EASI/ED functional requirements.

MIME provides the standards for how attachments may be sent with an e-mail. However, the actual attachment could be in various formats such as a Portable Document Format (PDF) file. Although PDF is a proprietary standard of Adobe Inc., it is a widely used, industry de facto standard that provides a platform-independent ability to exchange images of documents. PDF is used when documents must look exactly the same across all computer platforms and browsers such as IRS tax forms, application instructions, and brochures. PDF files are compact and can be viewed by anyone who has the free Acrobat Reader software. Many federal agencies use PDF as the file format for downloadable files on the Internet. EASI/ED may use PDF files for posting

instructions, applications, and books on the Internet. Since PDF is a standard for the structure of the data rather than how these files are being transmitted to and from EASI/ED, it is not specifically defined as an interface standard. However, in some cases, data flows have been identified with PDF as an interface standard. This was done to ensure that the format of the data being transmitted was not lost when determining an interface standard. Most of the data flows that are recommended for use with PDF are related to the exchange of correspondence between EASI/ED and its trading partners (for a detailed listing of the interfaces for which PDF is the recommended standard, please refer to Appendix G).

A variation of MIME is the Secure Multipurpose Internet Mail Extensions (S/MIME) standard. RSA Inc., a company founded by cryptographers that invented the most prevalent form of public key encryption, promotes S/MIME. The S/MIME method digitally signs an e-mail message and encapsulates it into an S/MIME envelope, where it is encrypted and sent to the receiver. S/MIME uses RSA's public key encryption service to encrypt the message and specifies that X.509 digital certificate standards be used to validate the public keys. Support for S/MIME is growing with its adoption in Microsoft's Exchange, Lotus Notes, Novell's GroupWise, Netscape's Communicator, and other messaging platforms. RSA is promoting its interoperability for inter-organizational use.

However, S/MIME's interoperability has encountered some hurdles attributed to problems in key and certificate management and the computation of digital signatures between various applications, products, and vendors. Therefore, until use of S/MIME increases, it is not a recommended interface standard for EASI/ED.

5.1.5 Interactive Voice Response

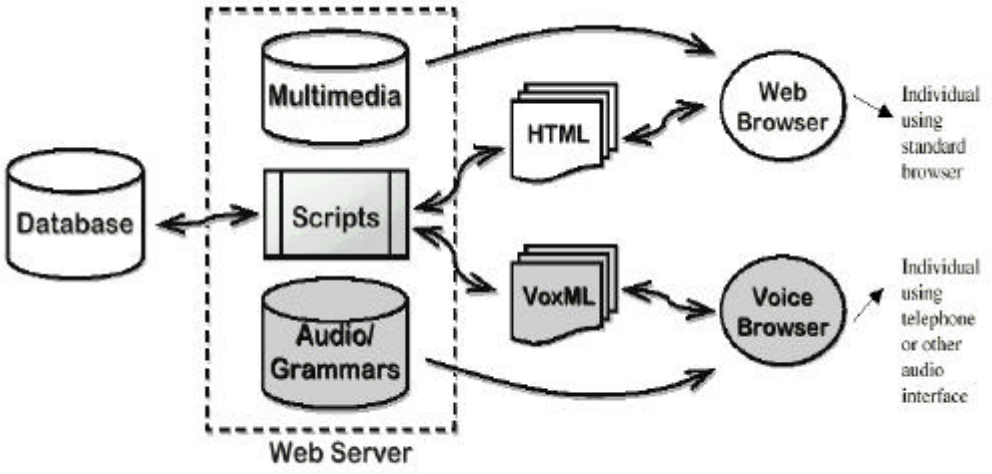
Telephonic communication is another transmission medium that may be widely used by EASI/ED and its partners and customers. This type of interface will generally be used for discussions, counseling, inquiries, status requests and responses, and other interactive interfaces.

For automation purposes, an Interactive Voice Response (IVR) system is envisioned as an integral part of EASI/ED. IVR is used to access and obtain information using limited human intervention. An IVR system can access data repositories to retrieve and update account information. It can be used for filling requests, producing reports, and interfacing with host systems. From a user standpoint, IVR systems allow users to access an organization's databases using a telephone's keypads that allows the user to navigate through the system to obtain and provide the necessary information. Most IVR systems also allow the user to connect to a customer service representative should the user require assistance additional to that provided by the system itself. The most widely used standard for IVR systems is currently Telephony Application Programming Interface (TAPI). Detailed information that describes the interfaces for which TAPI is the recommended standard is provided in Appendix G.

Voice XML or **VoxML** is an emerging markup language developed by Motorola for voice applications that allows developers to simply and easily add speech interfaces to their Web applications or content. Developing VoxML applications is similar to developing Web applications. VoxML applications are written in the form of the dialogues that allow users to interact by simply talking with the application. These dialogues are navigation and input commands produced via speech recognition of the end-user's voice, and output produced via text-to-speech technology or recorded audio samples. The voice browsers interpret the dialogues just as Web browsers interpret HTML pages. People can access Internet information from any telephone or voice-enabled device.

The VoxML language is based on the XML standard. As such, the language follows all of the syntactic rules of XML with semantics that support the creation of interactive speech applications.

The VoxML applications development environment consists of a Web server that hosts voice applications written in the VoxML language, and a voice browser. The VoxML application server is just an HTTP server that is typically located at the content provider's site. The VoxML contents are fetched from the VoxML application server over the Internet/Intranet by a VoxML voice browser. That voice browser can be running in the desktop development environment for users who have a PC. In this case, the user would access the voice browser via microphone and headphones connected to the PC. The voice browser can also be running in some other VoxML access server or device. This is particularly useful when the user is accessing the system via telephone. The telephone user dials into the access server that runs the voice browser for him. Figure 5-5 provides a graphical representation of how VoxML can be used.



Source: Motorola, Inc. 1999.

Figure 5-5: VoxML

A positive attribute is that users that are not able to connect to the Internet via a computer can provide and receive information to and from a system using a telephone. However, VoxML is still an emerging technology that does not yet have widespread usage. Therefore, VoxML is not recommended as an interface standards for EASI/ED at this time. However, ED should follow developments related to VoxML usage to ensure that future opportunities can be capitalized if VoxML becomes more widely used.

5.1.6 Facsimile

Description: Facsimile services allow users to transmit and receive documents via telephone lines. In recent years, it has become possible to fax documents using electronic mail (e-mail) or the Internet. However, if both the transmitting and receiving organizations have on-line capabilities (e-mail and/or Internet), then the preferred method of transmission is on-line exchange rather than fax because on-line interaction will more easily facilitate the electronic exchange and processing of data. For example, if a partner enters information via a Web page,

the data can be directly input into the receiving organization's systems. The subsequent response could also be facilitated through the electronic Web environment.

Benefits: Since few resources are required for this mode of transmission, i.e., a facsimile machine and a telephone line, partners and customers with limited technical capabilities can easily communicate with EASI/ED through this medium.

Limitations: Interaction by facsimile may be suitable only for low volume and low frequency interchanges of information since sending facsimile documents requires more manual processing than other transmission media.

Recommendations: The recommended standard for facsimile transmission with EASI/ED is International Telecommunications Union (ITU) T.4. The standard provides a transmission protocol for facsimile machines that use Group 3 compression. Although ITU T.4 is not the most current version of this standard, it is most widely in use. The most current standard, ITU T.6, is used to communicate with the machines that use Group 4 compression. Within the next year or two, ITU T.6 may become the industry standard for facsimile transmission and reception, but, in the near term, ITU T.4 is the most widely used and accepted standard. Therefore, this version of the standard is the recommended one for EASI/ED. As ITU T.6 becomes more widely used, EASI/ED should consider implementing ITU T.6 as well. Detailed information that describes for which interfaces ITU T.4 is the recommended standard is presented in Appendix G.

5.1.7 Paper

All the transmission media discussed above are more or less electronic and require access to technology. However, to accommodate those partners and customers that do not possess electronic capabilities, the use of paper must be incorporated into the design of the EASI/ED system to provide universal access to EASI/ED. Therefore, paper is a recommended interface medium for EASI/ED. Detailed information that describes the interfaces for which paper is the recommended standard is presented in Appendix G.

5.2 Project EASI/ED Interface Strategy

The purpose of this subsection is to present an interface strategy for EASI/ED. The interface strategy describes how data that is received from EASI/ED's trading partners should be synchronized among the physical applications that will constitute EASI/ED.

As illustrated in Figure 5-6, an external source may send in new data or an update to existing data to EASI/ED. Since EASI/ED has multiple databases used by multiple physical applications, the update should be relayed to all the applications that need this incoming data. Depending on what the application needs the data for, the update may need to be relayed immediately, or the update could wait until some scheduled event (e.g., a weekly or monthly update of all changes).

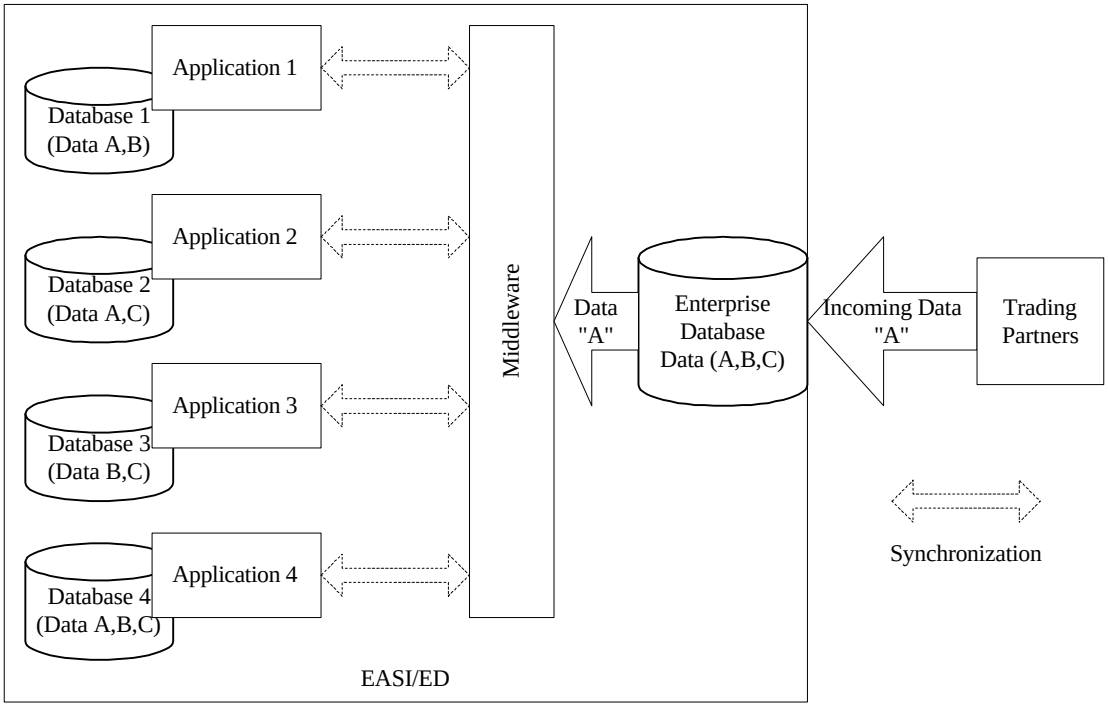


Figure 5-6: Project EASI/ED Interface Strategy Example

For example, assume data A is updated and the change is sent from an external source to EASI/ED's enterprise database. EASI/ED's Application 1 and Application 2 need this data instantaneously (real time), whereas Application 4 uses data A only for processes performed on an annual basis.
This synchronization of data between multiple applications is implemented using software tools referred to as "middleware". Middleware is crucial to the success of an enterprise distributed computing environment, that may include host-based systems, conventional client/server systems, and Internet-based applications. This subsection will introduce the different types of middleware and make some recommendations on what middleware SFA is likely to require to successfully implement EASI/ED.

This subsection contains the following subsections:

- Subsection 5.2.1 Assumptions
- Subsection 5.2.2 Timeliness Requirements

- Subsection 5.2.3 Synchronization Techniques
- Subsection 5.2.4 Recommendations

5.2.1 Assumptions

The following represents the assumptions under which the interface strategy analysis was performed.

1. The *Project EASI/ED ASDD, version 2.0*, comprises all the processes that need to be examined within the scope of this analysis.

2. The *Project EASI/ED LDMD, version 2.0*, comprises all the data entities that need to be examined within the scope of this analysis.

5.2.2 Timeliness Requirements

A key concept when considering data synchronization is that, depending on what use is being made of the data, the timeliness requirements will vary. Analysis was conducted to examine how up-to-date the data used by various EASI/ED business processes needs to be.

There are two broad categories of timeliness requirements:

- **Real Time Requirement**. A real time requirement exists when a set of data is needed by a process immediately as the data changes or is inserted in the data store. A process should be able to access the most current information without any delay. Examples include loan servicing, customer service, and repayment counseling.

- **Non-Real Time Requirement**. A non-real time requirement exists when a set of data is needed by a process, but not until the process is ready to use the data. For example, the aid organization default rate management process may need the latest aid program information, but needs it only as of a certain point in the year since the default rate management process is executed annually. There are many non-real time scales. However, analysis of ED processes indicated that two non-real time scales were prevalent:

    a) **Daily updates:** Processes need the updated information within 24 hours after it changes. The core aid delivery processes such as aid application and aid origination and disbursement functions need to receive updated information on daily basis.

    b) **As of:** Processes that update the data on a periodic basis ("as of" a specific date, no later than), such as the annual calculation of aid organization default rates, and quarterly interest and special allowance calculations.

For details of the timeliness requirements gathered for this task, refer to Appendix H. Timeliness requirements are a primary factor in determining the synchronization technique(s) that should be implemented in an enterprise. These synchronization techniques are discussed in the following subsection.

5.2.3 Synchronization Techniques

There are several techniques to accomplish synchronization in any given enterprise environment. However, as shown in Figure 5-7, these can be grouped into two categories:

- **Data Synchronization**: Data synchronization is, in essence, a database-to-database communication to exchange data. The propagation of changes from the source database to target databases can be accomplished by using either database scripts such as custom developed stored procedures, or an add-on utility such as a commercial off the shelf (COTS) data extraction/transformation tool. The target databases may be on different platforms than the source database.

- **Functional Synchronization**: Functional synchronization is a technique in which information exchange between two data stores in achieved indirectly via applications resident on both ends, rather than directly through database-to-database mechanisms.

Each of these categories of synchronization can be used to exchange information between target and source data repositories in two modes. The choice of mode of synchronization is primarily governed by the timeliness requirements of the target application. The two synchronization modes are:

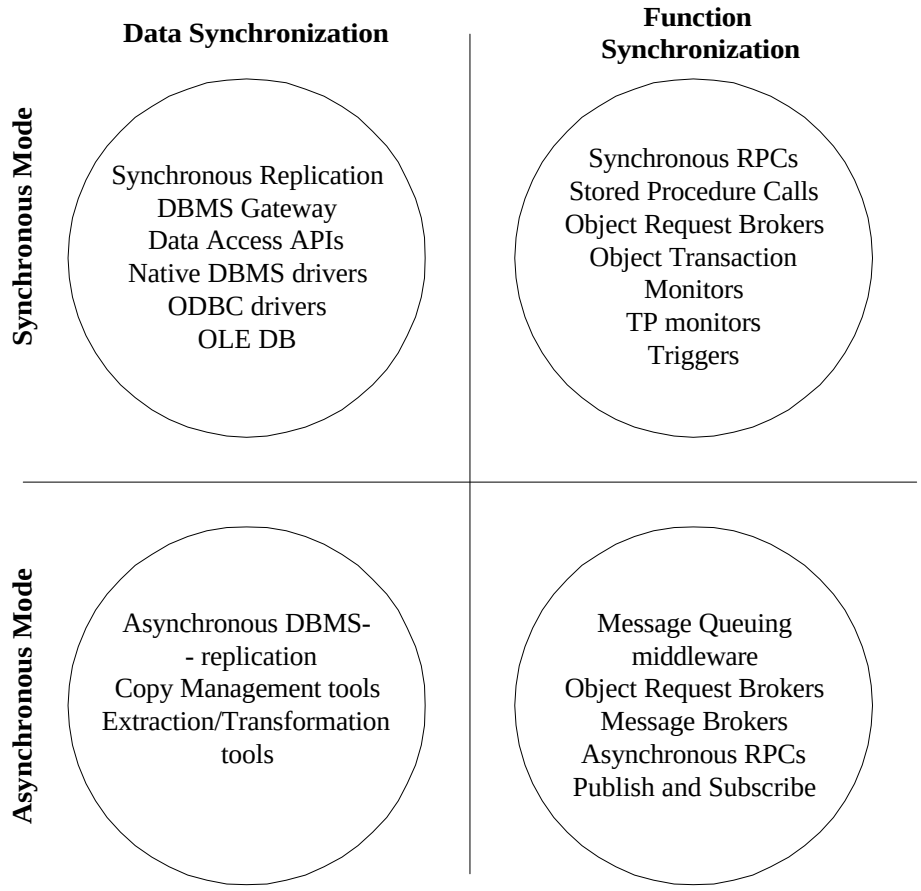


Figure 5-7: Synchronization Techniques

- **Synchronous Mode**: Data is exchanged immediately, either when the request for information comes in from a target application to the source application, or when an update comes in from other systems to the source application. The updates are then communicated simultaneously to all other related databases. This type of update uses a two-phase commit technique, which requires that all target databases be able to commit

the changes before any processing on the source system can proceed. This typically results in contention in both the source and target systems, which can have a significant negative impact on processing throughput.

- **Asynchronous Mode**: Changes to the source system are logged and shared with other systems, either continuously or at set intervals. The difference between synchronous and continuous asynchronous modes is that with the latter, processing in the source system continues whether or not the updates occur to the target system. If for some reason the updates to the target systems fail, updates to the source system continue to be logged until the transactions can be replicated to the target systems. This method is more scalable because it can tolerate communications interruptions that may occur over an extended communications network such as a WAN.

Functional synchronization is gaining popularity as an alternative to data synchronization. Where data synchronization only addresses data, functional synchronization also addresses business rules. With functional synchronization, an application program may extract data from multiple application systems into its local database and filter it using the business logic embedded for a particular process, so that all the edits are made locally without disturbing the integrity of the source databases. The best uses for functional synchronization are for linking different types of applications that were developed using different design or development tools. Functional synchronization is useful with complex applications that are troublesome to document and with dynamic applications that are updated frequently. Some of the commonly used functional and data synchronization techniques are described below.

5.2.3.1 Synchronization Approaches

As depicted in Figure 5-7, there are two different synchronization techniques, each with two different modes of implementation, resulting in a total of four possible approaches to implementing synchronization. These approaches are:

- Data Synchronization – Synchronous Mode (DS)
- Data Synchronization – Asynchronous Mode (DA)
- Functional Synchronization – Synchronous Mode (FS)
- Functional Synchronization – Asynchronous Mode (FA)

Data Synchronization – Synchronous Mode (DS)

This approach is sometimes called the “Real Time Data Reconciliation” technique. Copies are made of a source database and then transactional updates that occur on the source database is communicated to all targets, so the target databases end up with the same data. As mentioned earlier, this is usually accomplished using the two-phase commit technique, which requires that the update source remains locked until all the target databases acknowledge that they have committed their updates. This approach is often used with very high-value transactions, e.g., electronic funds transfers, because of the overriding need to maintain a consistent snapshot of the data. It demands that the organization assign updating privileges to each target and source and establish the location of the system of record. This approach is utilized if the data is required by a process on a real time basis and gets loaded using a batch process.

Some of the examples of DS middleware tools are:

- **Synchronous Replication:** Synchronous replication provides "tight consistency", in that the original update is made to wait until all the copies are updated within that single, originating global transaction. Not only are all copies at a given target site consistent among themselves, they are also consistent with all other copies at all other sites in the network. Synchronous replication is accomplished through the use of multi-site coordinated commit processing (two-phase commit protocol) or can be trigger-based. Some of the software tools available are InfoDirector and Sybase Replication Server (which offers both synchronous and asynchronous replication).

- **Open Database Connectivity (ODBC) Drivers**: These drivers are middleware that establish a least-common-denominator base of functionality for data exchange between data sources.

- **Data Access APIs**: As an alternative to using ODBC, APIs could be used to write database drivers. These APIs provide programmers with universal access to a wide range of relational databases. One of the emerging driver standards is the Java Database Connectivity (JDBC) standard developed by Sun and JavaSoft.

- **DBMS Gateways**: This type middleware supports development of applications that access legacy databases. With the increasing interest in distributed systems, DBMS Gateways have gained popularity. These products provide consistent connections between multiple databases to exchange information. Some of the commonly used DBMS Gateway middleware tools are dbAnywhere, Cerebellum, and Sybase web.sql.

- **Object Linking and Embedding (OLE) DB:** OLE DB is the fundamental Component Object Model (COM) building block for storing and retrieving records, and defines interfaces for accessing and manipulating all types of data. Some of the middleware tools based on OLE DB are : Acceler8-DB and DataGate/400, Connect OLE DB, and SequeLink OLE DB.

Data Synchronization – Asynchronous Mode (DA)

This data synchronization technique actually comprises two somewhat different techniques: "Batch Data Reconciliation" and "Log based replication". Both distribute data (whether updates, inserts or deletes) from source systems to target systems at either set intervals or set times. However, not all the databases will receive the updates at the same time, because the timing of distribution varies according to the timeliness requirements.

Some of the DA middleware tools are:

- **Asynchronous DBMS Replication**: Asynchronous replication provides "loose consistency", in that the latency before data consistency is achieved is always greater than zero. Replication occurs asynchronously with respect to the originating transaction or event. It is accomplished through either a complete or incremental refresh, or through the continuous propagation of individual database changes. When refresh technology is used, extracts from primary sources are scheduled and executed, data merging and/or transformation may occur, and target copies are loaded. When the propagation of database changes is used, the propagated change can be a software transaction, or a subset or superset thereof. With this technology, the target copy is initialized to bring it

to a consistent, baseline state with its primary source, and then only the database changes to the source system are propagated. Each originating software transaction either preserves its discrete identity during the replication process, becomes a part of an aggregated whole that is applied as a replicated set, or totally losses its identity during the replication process. If each originating software transaction preserves its discrete identity during the replication process, this is called transaction-based asynchronous replication. If each originating software transaction becomes part of an aggregated block of transactions that is applied as a unit, it is called transaction-consistent asynchronous replication. Some of the tools available are InfoReplicator, Lotus Notes, and Sybase Replication Server.

- **Data Extraction/Transformation Tools**: Data extraction is a process of selecting data from various source systems, and then transforming and restructuring that data so that it can be applied to one or more target systems. Data extraction tools provides the facility to specify which data is to be extracted and provides access to all the source databases. Data transformation tools, on the other hand, perform such functions as combining several data elements into one, summarization, computation, and generation of new attributes. Generally, both functions are included in single middleware tools such as Platinum InfoRefiner, IBM DataPropagator, Constellar's Constellar Hub, and Praxis International's OmniEnterprise.

Functional Synchronization – Synchronous Mode (FS)

This approach to synchronization facilitates real time distribution of data using application-level communication rather than direct database-to-database conversation.

Some of the examples of FS middleware are:

- **Synchronous Remote Procedure Calls (RPC):** Synchronous RPCs allow an application running on one computer system to call a procedure that executes on another system. The RPC middleware provides an API that the programmer uses to initiate and respond to the procedure calls in the client and server applications, and then manages the details of transferring the calls across the network. Examples of commonly used RPC middleware is Sun RPC (used on the Solaris Operating System), and Microsoft RPC (used on the Windows platform).

- **Object Request Brokers:** With the advent of the Common Object Request Broker Architecture (CORBA) standard, a new type of distributed environment has become possible. Unlike Microsoft's Distributed COM, which is based largely on RPCs, CORBA is based on servers called Object Request Brokers (ORBs). These ORBs handle requests from clients or other ORBs and instantiate objects to provide services. Because only the ORB needs to know the information necessary to instantiate objects, the configuration management overhead of distributing that information to each client (a major shortcoming of DCOM) is avoided. ORBs are being built into most industry-leading application servers, such as the Oracle Web Application Server and the Sun/NetDynamics application server. Furthermore, a new class of "enterprise integration" tools has arisen for implementing functional synchronization that support CORBA and some include an ORB (most also provide support for messaging and/or DCOM as well). An example of an ORB-enabled enterprise integration tool is TIB/ActiveEnterprise from TIBCO. Commercially available ORBs include the Inprise

Visibroker and Iona's Orbix. It should be noted that ORBs can operate asynchronously as well as synchronously.

- **Transaction Processing (TP) monitors**: TP monitors (also called transaction managers) have their roots in early mainframe processing. In a database system, a transaction represents a group of commands that must succeed or fail as a single unit. For every action performed, verification the action has occurred correctly is required before execution can proceed. This management of transactions is performed by TP monitors. Some of the examples of TP monitors are IBM's CICS and Encina, BEA's Tuxedo, and iTRAN, which provides vendor independent TP monitoring APIs and libraries. Two standards have been developed which will facilitate support for TP products in heterogeneous environments: the International Organization for Standardization (ISO TP) standards and the Distributed Transaction Processing (DTP) model from The Open Group (formerly X/Open).

- **Object Transaction Monitors:** Although ORBs support transaction semantics (CORBA Transaction Services, for example), they lack many of the services required for high scalability, including efficient memory and state management, load balancing, and resource sharing. A second generation of ORB product is emerging as a new standard for application deployment. Object transaction monitors (OTMs) combine the scalability and robustness of TP monitors with the flexibility of component models. OTMs are designed to support the deployment of high-volume, mission-critical applications. They support distributed transactions, and they support highly efficient resource sharing. An OTM supports one or more of the three standard distributed component models (CORBA, EJB, and COM). Examples of CORBA OTMs include BEA's M3, Iona's OrbixOTM, Inprise's VisiBroker/ITS, and IBM's Component Broker. EJB OTMs include WebLogic's Tengah, Secant's Extreme, and Persistence's PowerTier. COM OTMs include Microsoft Transaction Server (MTS) and Sybase Jaguar CTS.

- **Stored Procedure Calls**: Most DBMS products allow developers to write executable modules that are stored in the DBMS. These procedures are called interactively by applications using the DBMS, or directly by a DBMS user. Some of the DBMS products that support stored procedures are Oracle, MS SQL Server, DB2, and Sybase.

- **Triggers**: Triggers are specialized stored procedures that get executed in response to an event in the database, particularly an insert, update or delete operation. In the context od data synchronization, the triggers are then used to transmit information from one place to other in response to an event.

Functional Synchronization – Asynchronous Mode (FA)

This synchronization approach provides inter-application communication on an asynchronous basis. The application software provides an API that the programmer uses to create and accept messages in the client and server applications. At run time, the software manages reliable delivery as the messages are transferred between the systems.

Some of the commonly used middleware techniques are:

- **Asynchronous RPCs**: RPCs can also be implemented to perform asynchronous communication by instructing the server to respond to a request on timely basis and providing a decision making facility to determine whether a response is required. Examples of widely-used RPCs are the Sun RPC and the Microsoft RPC.

- **Message Oriented Middleware (MOM)**: MOM is a specific class of middleware that operates on the principles of message passing and/or message queuing. MOM is generally based on exchanging information asynchronously applications. Applications communicate by writing messages (information packets that contain the address of the sending and receiving application, and the information to be exchanged) into message queues and reading messages from the message queues that have been sent by servers. From the sender's prospective, the message has been sent as soon as it has been deposited in the appropriate queue, even though the reply from the recipient has yet to be received. MOM allows applications to communicate over a network without a dedicated and persistent network connection. MOM products generally are implemented using one of two mechanisms:

  1. **Publish and Subscribe Message Oriented Middleware:** The message is published (that is, sent) to a public service, and the target destination awaits messages from a service in which it has expressed an interest (subscribed). The source and target need have no knowledge of each other, only of the service that intermediates.

  2. **Message Queuing:** The message is sent to the target's queue, then the target checks the queue on its own schedule for messages it needs.

  MOM offers many advantages in a distributed database and computing environment, because it is DBMS-independent and provides communication between relational databases from different vendors, between relational and non-relational databases, and between databases and non-database data stores (e.g., VSAM files). Examples of MOM middleware tools include BEA's Top End, IBM's MQ Series, Talarian's SmartSocket, PeerLogic's PIPES Platform, and Level 8 System's XPIC.

- **Message Brokers**: A message broker is middleware acting as a traffic controller between source and target. Message brokers "share" messages in the sense that the source need transmit only one message and the broker will deliver one or many versions of it to one or many target applications. Message brokers typically incorporate data transformation capabilities, content-based routing, and business rule repositories, so that they serve as the hub of a distributed enterprise. Examples of message brokers include CAI Impact!, Glotech's MBS, SAGA's EntireX, NEON's NEONet, Vitria's BusinessWare, and Active Software's ActiveWorks.

5.3.4 Recommendations

Synchronization is a technique to either replicate or distribute data among multiple data stores or applications. Data must be available for multiple applications to perform required tasks at the right time. One particular synchronization approach may not be able to fulfill all the synchronization needs of an enterprise. There are several factors that govern the selection of an approach. These factors and the synchronization selection process are discussed in more detail in Appendix H. The result of the analysis described in Appendix H is summarized as in Figure 5-8.

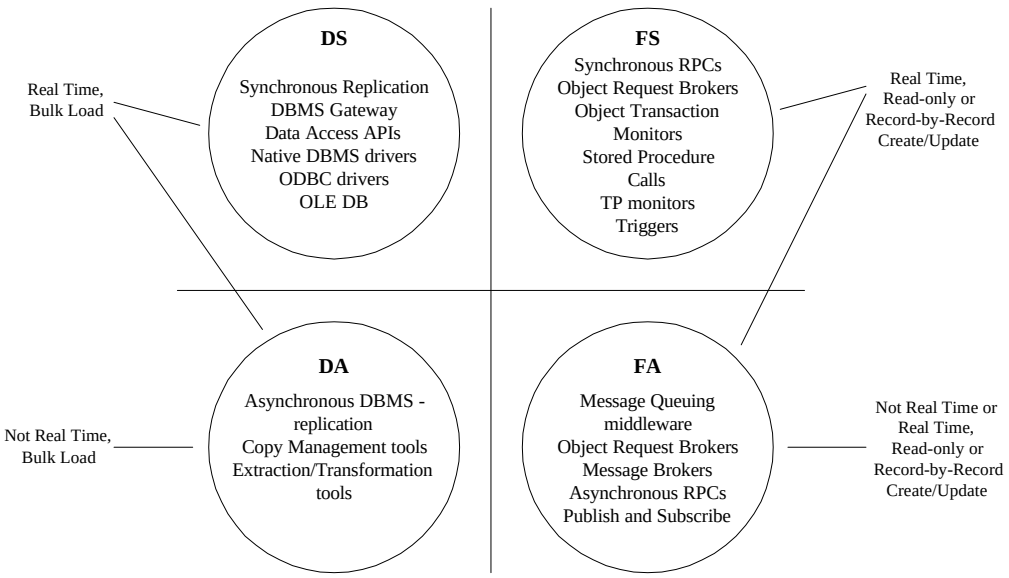


Figure 5-8: Project EASI/ED Synchronization Recommendations

The possible permutations that can result from the process contained in Appendix H are as follows:

- Either the FS or FA approach should be used for real time synchronization where access is either read-only or record-by-record updates.
- The FA approach should be used for non-real time synchronization where access is either read-only or record-by-record updates.
- Either the DS or DA approach should be used for real time synchronization where access is bulk load of data.
- The DA approach should be used for non-real time synchronization where access is bulk load of data.

As shown in Figure 5-8, each synchronization approach has some benefits and there is no particular order of importance of these approaches with respect to Project EASI/ED. For Project EASI/ED synchronization purposes, therefore, it is important to implement a blend of technologies that can provide all four types of approaches. There may not be a single suite of products that will meet all the needs, but functional synchronization tools in conjunction with the database tools supporting data synchronization can provide the functionality of all four quadrants.

Another important aspect of successful implementation of synchronization using middleware tools is the management of the applied techniques. A key challenge for middleware currently is the lack of standards for the information exchanged between the source and target to achieve synchronization. Specifically, the newer technologies such as MOM have no standard formats for messages. Every vendor has its own format and design for messages.

The challenge is to keep synchronization operating smoothly while accommodating changes in technology and in data processing needs. Through disciplined configuration management and

periodic technology assessments, synchronization management ensures that each synchronization technique remains current and aligned with changing business and technology requirements. Adopting this strategy will ensure that the ED's existing and future application systems are able to consistently and reliably share the data necessary to support the enterprise-wide vision of Project EASI/ED.