JIEO Report 8307
30 November 1998

DoD Guide to Selecting Computer-Based Multimedia Standards, Technologies, Products and Practices

Defense Information Systems Agency
Joint Interoperability and Engineering Organization
Center for Standards
Information Processing Standards Division
10701 Parkridge Boulevard
Reston, Virginia 20191-4357

(This page intentionally left blank)

TABLE OF CONTENTS

1. INTRODUCTION

1.1 ORIGIN

This document was prepared by the Defense Information Systems Agency (DISA), in cooperation with the Services and other activities within the Department of Defense (DoD). This cooperation was effected through the Multimedia/Optical Disk Technology Standards Working Group (M/ODT SWG) of the DoD Standards Coordinating Committee. The M/ODT SWG is chartered to:

- Influence the development of federal, national and international standards which satisfy multimedia functional requirements
- Develop and insert DoD enhancements into federal, national and international multimedia standards
- Develop DoD multimedia specifications where DoD requirements are not otherwise incorporated in federal, national, international or industry standards
- Ensure compatibility and interoperability for multimedia authoring, storage/retrieval, and playback (user) systems through coordination with industry.

The charter of the M/ODT SWG aligns its multimedia activities with computer-controlled multimedia, and its efforts are managed as part of the Information Processing Standards for Computers (IPSC) standardization area within the DoD Information Technology Standards Program implemented by the DISA Joint Interoperability and Engineering Organization (JIEO) Center for Standards (CFS).

1.2 PURPOSE AND SCOPE

The goal of this document is to provide guidance and to serve as a reference source for information system architects and design engineers, acquisition managers, and procurement officials to select appropriate multimedia standards for their projects.

Supporting objectives include enhancing the ease of integration of multimedia products and platforms (often a nontrivial task) and improving interoperability across the DoD enterprise, based on tenets and guidance associated with the Defense Information Infrastructure (DII). This includes conforming to and supporting the relevant standards, technologies, products and practices of the DII Common Operating Environment (COE). Within the DII COE, multimedia must be supported by hardware platforms, the operating system and kernel, infrastructure services, and common support applications. The topics covered by this handbook address this range of required support. Specific multimedia requirements of mission support applications (e.g., Continuous Acquisition and Life Cycle Support (CALS)) are, however, beyond the scope of this handbook.

This handbook defines multimedia, presents an overview of multimedia standards, technologies and products relevant to the DII, and provides detailed discussions of the four main media types – audio, imagery, text, and video – commonly combined in multimedia. It also explains the relationship of various DoD documents to the guidance provided in this handbook (Joint Technical Architecture (JTA), Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance (C4ISR) Architecture Framework, Technical Architecture Framework for Information Management (TAFIM), Adopted Information Technology Standards (AITS), and Information Technology Standards Guidance (ITSG)).

This handbook represents the technical architectural component of the C4ISR Architecture Framework. As such, it replaces both JIEO Report 8303, Multimedia Extensions to the DoD Minimum Desktop Configuration, and JIEO Report 8306, Guidelines for Selecting Multimedia Computing Equipment. In replacing these two reports, this handbook broadens the scope of guidance on multimedia systems and components acquisition to include guidance on implementation. In addition, this handbook both reflects current and emerging technology and also provides a framework for incorporation of future technological developments.

Although multimedia can be defined to include non-computer-based audio/visual materials, the scope of this handbook is limited to multimedia acquired, processed, distributed or presented through computer-based means, to fall within the charter of the M/ODT SWG.

This handbook provides information and guidance that is needed to comply with the Office of Management and Budget (OMB) Memorandum M-97-16 and as required by the Clinger-Cohen Act of 1996 (Public Law 104-106). OMB Memorandum M-97-16 requires that Federal agencies develop and implement Information Technology Architectures for their agencies. Section 5125(b)(2) of the Clinger-Cohen Act (also known as the Information Technology Management Reform Act (ITMRA)) responds to the OMB memorandum and specifically requires DoD to develop and implement an Integrated Information Technology Architecture. The information contained in this handbook will contribute directly to the development of standards profiles that form the foundation for required Information Technology Architectures.

This document is for guidance only, and cannot be cited as a requirement in a contract. It will evolve as guidelines and new standards mature.

1.3 DEFINITION OF MULTIMEDIA

The following M/ODT SWG definition for multimedia is used in this handbook:

Multimedia: Two or more media types (audio, video, imagery, and text) in which the data streams are electronically manipulated, integrated, and reconstructed to maintain their time relationship (synchronization).

Multimedia data represent information obtained from multiple sources using different types of input devices such as cameras, microphones, scanners, keyboards, and storage devices.

1.4 GENERAL GUIDANCE ON USING THE HANDBOOK

The DoD Guide to Selecting Computer-Based Multimedia Standards and Practices Handbook examines multimedia standards for animation, audio, data, imagery, text, and video. Each section of the guide begins with a table designed to provide recommended and non-recommended standards, technologies and products. Its purpose is to summarize recommendations discussed in the respective section. Each of the chapters addresses specific multimedia standards using the following structured format:

- Overview of the multimedia component
- Summary of standards, technologies, products and practices for that component, with accompanying guidance on selection
- Issues and discussion associated with that component

1.4.1 Standards, Technologies, Products and Practices

Standards are the key to ease of integration and the assurance of interoperability needed for successful exchange of data and information via multimedia. Users can help ensure ease of integration and interoperability of their systems and applications by purchasing and using hardware, software, and implementation practices that conform to specified standards and guidance.

Formal standards, industry specifications, and the technologies and products based on them all must be responsive to users. Users have needs met by standards, technologies, products and practices implemented in both the computer systems and the applications that users apply to their day-to-day activities. A standard or technology generally will not succeed without a viable and widely accepted user application, along with the support of multiple vendors and, in today's marketplace, ease of integration and the ability to support interoperability. Note that the extent of usage of the standard or technology in equipment or applications will determine the success of that standard or technology, although changes in corporate commitments can affect the long-term success of a standard or technology.

Standards and their implementing technologies are developed in two environments. Vendors are involved in both of them. One environment is that of formal national and international standards bodies such as International Telecommunications Union (ITU), Internet Engineering Task Force (IETF), Institute of Electrical and Electronic Engineers (IEEE), International Organization for Standardization (ISO), American National Standards Institute (ANSI), etc. The second environment is more directly market-(i.e., user) driven and results from industry setting a *de facto* standard, either by agreement between companies, implementation by multiple companies, or by the wide acceptance of products from a particular firm in the marketplace. This second environment is driven by the ever-accelerating rate of technology development and advancement that is the norm in today's networking and computing marketplace.

Informal standards often are categorized as being "*de facto*" standards. *De facto* standards are those that, because of their perceived qualities or their degree of market share, have become dominant in the marketplace. It is through purchases that users can most effectively show their support for *de facto* standards.

Development often outpaces formal standards development and ratification cycles due to major support from industry forums. Formal standards can in almost all cases be seen as lagging behind the leading edge of available technology and thus, the marketplace is driven to develop de facto standards. Many problems can only be solved by the use of proprietary implementations prior to the acceptance of the formal standards. Users through their demands and purchases determine whether a proprietary implementation succeeds in the marketplace.

Considering this environment, "industry standards" are determined through the marketplace or through an industry forum.

Finally, the practices used to implement standards, technologies and products can, where leeway exists for alternative implementations, materially affect the ease of integration and interoperability achieved by the implementation. For this reason, the concept of preferred practices must be included in any guidance.

1.4.2 Guidance on Selection of Standards, Technologies, Products and Practices

Earlier DoD guidance documents have recommended selection of technologies and products based on approval status (e.g., emerging, *de facto*) of associated standards. This handbook departs from that approach in an effort to provide more meaningful guidance. This guidance for selection of appropriate technologies, products, and the standards on which they are based, is predicated on the expected level of effort required for implementation. In this way, program managers and executive agents can apply the guidance to make balanced assessments of alternatives appropriate to their respective programs. Note that this guidance does not stand alone as selection criteria for acquisition of particular products. Due diligence is still required on the part of program managers and executive agents, in balancing the particular needs of their respective programs, their mission requirements, and the cost and budgetary impacts of their selection alternatives.

To facilitate these balanced assessments, this handbook categorizes standards, technologies and products in terms of their suitability to enterprise-wide multimedia applications, using the following hierarchy:

- **Lower Level of Effort –** well supported commercially, meaning that fully supported commercial products will be available within the next twelve months. Generally secure and interoperable with other DoD components, and cost-effective.
    - The standards and technologies in this category have been reviewed and recommended by representatives of various Government organizations, including DoD Services and agencies.
    - Existing products in this category should be available via existing DoD-wide procurement vehicles and enterprise licensing programs, including free use for both Government and commercial applications.
- **Higher Level of Effort –** not well supported commercially or otherwise presenting specific integration, interoperability, cost or security challenges.
    - May be integrated, made secure and made interoperable with other DoD components, but likely only with significant additional time and resource investments.
    - Implementations are expected not to be cost-effective.
    - Typically represent legacy standards, technologies and products.
- **Monitored –** Emerging, new, or revisions to existing, multimedia standards, technologies and products, with maturity and market penetration insufficient to categorize an expected implementation level of effort.
    - Encouraged in proof-of-concept prototypes, in order to ensure appropriate evaluation of their applicability to the DoD enterprise.
    - Use on operational systems expected to pose significant challenges and risks.

Guidance in each area will differentiate where necessary between that for multimedia consumers and that for multimedia authors. This is necessary to simplify authoring system implementation and support, while ensuring DoD multimedia consumers the means to access a reasonable segment of the broad range of multimedia published in non-DoD arenas. This simplification is crucial to expanding multimedia system and component acquisition to the entire DoD enterprise, through such vehicles as the DISA/JIEO Software Enterprise Licensing (SEL) Program. Only through simplification can the needs of the majority of users be met and maintained at reasonable cost, when the scale of enterprise-wide implementations is considered.

This handbook will also, where appropriate, provide separate guidance on implementation practices, using the following hierarchy:

- **Preferred Practices –** specific practices that contribute to ease of integration, interoperability, and security.
- **Discouraged Practices –** these practices should not be adopted for DoD systems, usually for reasons of security and interoperability.

1.4.3 Related Guidance

This handbook should be used with other relevant DoD documents that provide additional guidance and information concerning multimedia standards; a summary of these documents is provided in Appendix D, Overview of Related DoD Guidance. Note that references to JIEO Report 8303, Multimedia Extensions to the DoD Minimum Desktop Configuration, and to JIEO Report 8306, Guidelines for Selecting Multimedia Computing Equipment, are for historical purposes only, since this handbook replaces both previous guidance documents.

Appendix E of this handbook, Selection Criteria Cross-Reference, provides a cross-reference that should be used to identify the standards appropriate to your specific needs.

Appendix F of this handbook, Cost-Benefit Analysis Checklist, provides guidance for those planning the acquisition of multimedia systems or components. This checklist is designed to ensure that key cost and benefit factors are included in the scope of any analysis.

2. MULTIMEDIA DATA INTERCHANGE

Standards for multimedia data interchange include those necessary for electronically conveying text in different languages, compilation of text and other media into compound documents, depiction of text on monitor screens and hardcopy devices, and conveying information to be printed to printing subsystems and via other electronic means. These standards also encompass two-dimensional (2-D) and three-dimensional (3-D) vector graphics, raster depictions of still images and motion video, animated graphics, digital and encoded audio, and mixtures of video and audio information.

2.1 TEXT AND DOCUMENT ENCODING

The term “text” normally refers to words, sentences, and paragraphs, in contrast with “data,” which are defined units, such as “name” and “amount due.” Text information is normally readable by humans without further processing; data normally refers to machine representations that require processing in order to be readable by humans.

Text also may refer to alphanumeric data, such as “name” and “address,” to distinguish it from numeric data, such as “quantity” and “dollar amount”. A page of text occupies about 2,000 to 4,000 bytes, presuming 40-80 characters per line and an average of 50-60 lines per page.

Text encoding is a method of defining character sets as numerical values that are mapped to specific characters, for processing by computing and communications systems.

2.1.1 Text Encoding

Computer representations of text from different languages requires encoding schemes for characters and symbols other than the twenty-six uppercase and twenty-six lowercase Latin characters and ten Arabic numerals of U.S. English. This section provides guidance for that encoding.

2.1.2 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	**SOURCE**
Lower Level of Effort	
Unicode 2.0 or 8-bit ASCII, depending on application and operating system	DII COE, AITS/ITSG
Higher Level of Effort	
UCS, ISO 8-bit, 7-bit ASCII, Baudot	AITS/ITSG, RFC 821 (SMTP), JANAP 128

PRACTICES	**SOURCE**
Preferred	
ISO 8859-1 (Latin-1) alphabet mapping	DII COE, AITS/ITSG
Discouraged	
Authoring with UCS, ISO 8-bit, 7-bit ASCII, Baudot or other mapping	AITS/ITSG, RFC 821 (SMTP), JANAP 128

2.1.3 Overview

2.1.3.1 ASCII and ISO 8859-1

The American National Standards Institute (ANSI) defined the American Standard Code for Information Interchange (ASCII) as a standard character set in 7-bit (128-character) encoding. Later, an 8-bit (256-character) encoding was added. Virtually all PCs and office automation systems process ASCII-encoded information, with other languages transliterated into ASCII format where required. For this reason, ASCII is will require less effort to integrate and make interoperable. However, the need to represent more than the basic English characters (i.e., including punctuation marks, etc.) requires the use of 8-bit ASCII.

ISO 8859-1, known as Latin-1, is the preferred implementation of ASCII-based 8-bit codes for representing character data. The Latin-1 standard is built on the ANSI ASCII standard. Text referred to as “ASCII” is usually a 7-bit subset of the ISO standard (i.e., using only the first 128 code values, 0 - 127).

Note that with ongoing efforts associated with conversion to European Monetary Union currencies, the standard may be modified to include a character representing the new currency unit.

	¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯
°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

Figure 2.1.3.1 ISO 8859-1 Character Set

2.1.3.2 Unicode 2.0 and UCS-2

Unicode 2.0 is a text encoding standard which grew out of the need to represent ligatures, diacritical marks, special characters and other features of non-English languages. Since Windows NT uses Unicode for long filenames, it is a mandated standard, though ASCII may be applied for other uses. Its proponent is the Unicode Consortium, a combined industry-government group with representation from all major computer software firms (e.g., Apple, Microsoft, Sun, IBM, and HP). The Unicode 2.0 standard contains nearly 40,000 distinctly coded characters derived from the principal written languages of the Americas, Europe, the Middle East, Africa, India, Asia, and the Pacific.

UCS-2 is a specific 16-bit (i.e., two byte, hence the -2 identifier) implementation of the Unicode standard. UCS-2 is independent of byte order (i.e., big-endian or little-endian); the mandated guidance follows operating system practice for byte order. Regardless of the implementation, it is inappropriate to treat any portion of the character code independently; rather, the entire construct must be included in the processing.

Unicode implementations are also available in 7-bit, 8-bit, and a 16-bit extended form. However, the difference between Unicode and ASCII is that ASCII is character-based, whereas Unicode carries the added concept of glyphs – visual depictions of characters. In ASCII parlance, a font is composed of characters; in Unicode parlance, a font contains glyphs. Unicode does not attempt to encode all glyphs, presuming that the number of characters to be encoded is less than the number of glyphs. Unicode characters, however, are a mixture of characters and glyphs, largely because the two have been confused in legacy implementations of text encoding. Another complicating factor is that Unicode characters do not identify the associated language; separate standards are required for that identification, such as:

- ISO/IEC 6429 Based Language Tagging Proposal
- ISO 639 Code for the Representation of Names of Languages
- Committee Draft ISO 639-2 on Extended Language Codes

For this reason, Unicode support in multimedia systems should be limited in the near term to those applications requiring portability across national language boundaries, and filenames, directory objects, and other operating system constructs.

7-bit ASCII coding schemes are discouraged, because of the limitations imposed on character set content.

2.1.3.3 Baudot

Baudot encoding is an obsolete 5-bit code for teletype applications. In order to represent all the characters of the English language, as well as numbers and special symbols and control characters, a shift scheme was used to overload the encoding, assigning two definitions to most codes. In current implementations, Baudot-encoded text is converted to ASCII prior to display.

2.2 DOCUMENT ENCODING

Electronic messaging and other applications may require that all information, including binary data or compound documents, be encoded as text for transmission (i.e., for compatibility with text-only transmission media).

2.2.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
Uuencode	RFC 821 (SMTP)
MIME (RFC 1521), Secure-MIME (S/MIME)	GCCS
Higher Level of Effort	
BinHex (Mac .hqx)	Apple Computer

PRACTICES	SOURCE
Preferred	
Use the grave accent (`) instead of a space character in uuencoded text	RFC 821 (SMTP)
Discouraged	
Using space characters in uuencoded text	RFC 821 (SMTP)

2.2.2 Overview

2.2.2.1 Uuencode

Uuencode (Unix-to-Unix encoding from Unix-to-Unix Communications Protocol – UUCP) is a 7-bit ASCII encoding scheme which is the predominant scheme used by many SMTP (RFC 821) e-mail applications to convert attachments to ASCII code which can be conveyed under the rules of RFC 821. If normal uuencode practice is followed, the name of the file (and its extension) can be included as plaintext at the beginning of the encoded data stream. If the filename extension has been associated with a particular content type or source application, then that content type can be recognized and the associated application can automatically be launched as a viewer or editor. Unfortunately, not all users and applications do so.

More importantly, there are two distinct implementations of uuencode. The more modern implementation uses a grave accent (`) instead of a space. This is crucial should the transfer software strip trailing blanks before sending. Obviously, should trailing blanks representing encoded data be removed, the entire encoding will be corrupted.

These and other flaws in the uuencode scheme led to development of a base 64 encoding scheme instead of the 7-bit ASCII scheme. This new scheme is part of the basis for Multipurpose Internet Mail Extensions (MIME and S/MIME).

2.2.2.2 MIME and S/MIME

The Multipurpose Internet Mail Extensions (MIME, RFC 1521) standard was designed to address the shortcomings of uuencode. MIME provides a more robust encoding scheme designed to survive passage through extant Internet gateways. MIME also typically includes a reference to the content type of (source application for) the data encoded. This ensures portability across platforms, without requiring mapping of individual filename extensions to particular source applications.

Secure MIME (S/MIME) is a draft Internet standard for incorporating public-key cryptographic schemes (specifically, the Diffie-Hellman key exchange algorithm) as part of the MIME encoding. Note that Diffie-Hellman is a royalty-free algorithm; previous incarnations of S/MIME were based on proprietary algorithms from RSA Data Security, Inc. (RSA), which would have required royalty payments to RSA by S/MIME users. However, the Diffie-Hellman algorithm is preferred as part of the proposed standard, not required, in order to accommodate existing S/MIME implementations which had been based on RSA algorithms. A more recent edition of the standard, Version 3, is now under consideration. Version 3 includes provision for Extended Security Services, which would permit support for both U.S. Government cryptographic algorithms from the Defense Message System and other S/MIME algorithms, improving interoperability support.

A separate, related working group is preparing a proposed standard, PGP/MIME, based on the Pretty-Good-Privacy (PGP) cryptographic algorithms rather than those being covered by the S/MIME approach. For example, PGP/MIME uses the Digital Signature Standard (DSS) for user authentication, whereas S/MIME relies solely on X.509 certificates. Some combination of the two standards is likely, since a primary reason for the PGP/MIME venture was to afford users a royalty-free standard. With the decision of the Internet Engineering Task Force (IETF) to base S/MIME on Diffie-Hellman rather than RSA algorithms, some of the need for PGP/MIME has disappeared.

2.2.2.3 BinHex

BinHex is a proprietary binary data and text encoding scheme promulgated by Apple. It is the default encoding scheme used by Macintosh applications and operating systems. BinHex handles the multiple forks (i.e., resource and data) used by the Macintosh file system, so is useful in transmitting files from Macintosh to Macintosh. Although freeware and shareware decoders such as `BinHex13.exe` and StuffIt Expander are available for PC systems, they are not widely distributed or installed. This makes documents encoded with BinHex difficult to exchange freely throughout the enterprise.

2.2.3 Issues/Discussion

For character sets, each language needs a programming environment to handle conversion, sorting, and string handling to support proper localization and internationalization.

Lack of broad support for S/MIME hampers its application across the enterprise.

Proliferation of MIME application types complicates its use and dilutes the utility of the standard.

Text encoding will be used in virtually all applications; even images can have embedded, encoded text identifiers.

Target presentation systems and viewers may not have the required support for specific text encoding.

2.3 COMPOUND DOCUMENT FORMAT

Compound documents provide a mechanism for packaging information from different media sources and for displaying multimedia information through a common interface.

2.3.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
HTML 3.2	GCCS, JTA, AITS/ITSG
PDF 2.0 (authoring), PDF 3.0 (viewing), Word 6, PowerPoint 4, Excel 5, ASCII Text (or others with freeware viewers or viewers bundled with the operating system)	GCCS, JTA, DII COE
Higher Level of Effort	
Postscript, WP6.x, Word 2, RTF, Freelance 96, SGML, PDF 2.0 (viewing)	GCCS, JTA, AITS/ITSG
HTML 2.0, dBase 4, WP5.x, Lotus 1-2-3 3.x, Freelance 2, Microsoft Office 95 or Office 97 below SR-1 (and others without freeware viewers)	GCCS, JTA, AITS/ITSG
Monitored	
ISO HyTime, XML, Word 97 SR-1, PowerPoint 97 SR-1, Excel 97 SR-1, HTML 4.0	ISO, W3C, Microsoft

PRACTICES	**SOURCE**
Preferred	
Preferred font color	DII COE
Preferred typefaces, point size, and capitalization	DII COE
Hyperlinks distinguished from regular underlined text	DII COE
Discouraged	
Proprietary extensions to HTML	DII COE

2.3.2 Overview

2.3.2.1 HTML

HyperText Markup Language (HTML) is an informal Internet standard defined by RFC 1866. HTML consists of a set of tags that conform to Standard Generalized Markup Language (SGML) rules and conventions. The HTML tag set can be used as the basis to define a Data-Type Definition (DTD) that is consistent with SGML syntax. By defining HTML in an SGML DTD, HTML becomes an SGML application.

The HTML document type contains relatively general semantics for representing information for linking of data and document with a limited SGML tag set and limited formatted capability. Moreover, simplicity was the guide in development so that multiple browsers and editors could be used on multiple platforms. The following list gives some idea of the specific uses available: hypertext news, mail, on-line documentation, and menus of options, database query results, and simply structured documents with in-line graphics. HTML has the capability to allow networked hypertext to use text, sound, movie, and images in a variety of formats.

Development and promulgation of new HTML standards is the responsibility of the World Wide Web Consortium (W3C).

The preferred implementations of HTML are governed by guidance promulgated in support of the DII COE.

2.3.2.2 PostScript and Portable Document Format (PDF)

PostScript and PDF are informal vendor specifications published by Adobe Systems. PostScript is a page description language designed for presentation of complete, formatted, final-form page images on output printing devices. It heavily influenced the ISO/IEC standard 10180, Standardized Page Description Language. PDF is an application of PostScript that includes freeware viewing tools and supporting authoring tools.

Portable Document Format (PDF) is a proprietary specification developed by Adobe Systems Incorporated. PDF is a file format used to represent a document in final form independent of the application software, hardware, and operating system used to create it. It offers precise control over the documents appearance and permits the document to retain that same appearance over different platforms. The document may contain any combination of text, graphics, and images in a device independent and resolution independent form. It may also contain information possible only in an electronic representation, such as a set of hyperlinks, thumbnail icons of pages, chapter outlines and page annotations. The PDF file contents may be searched and archived.

Adobe has asserted that PDF is an open standard based the availability of the published format. The Adobe implementation of PDF is called Acrobat and is being used on the Internet for final form presentation of a document on a platform as well as on paper. The US CALS Industry Steering Group, Hypermedia Information Systems Special Project Committee has recommended the use of PDF as a standard for delivery of non-revisable documents.

PDF is both complete and extensible. It provides a compressed cross-platform file that retains fonts, color, and layout providing for electronic distribution of paper media format in soft-copy form. A “plug-in” capability allows for PDF producers to add private data and to use that data to define new actions, keys, and behavior.

Numerous software applications support the input, processing and generation of the PDF format. Browsers are available at no cost allowing PDF documents to be read on Macintosh, UNIX, and Windows platforms.

Adobe is working closely with several vendors that plan to introduce image-display devices that can easily produce multimedia presentations from PDF files.

The availability of screen fonts on the receiving platform to match those in the PDF document has posed a difficulty. Adobe is employing their Multiple Master Font technology to produce font characters from a generic font that are fitted on the fly to the font metrics of the original (though unavailable on the presentation platform) font. In some ways PDF is considered in competition with SGML, which is a document mark-up language, and HTML, which is a specific implementation of SGML and a de facto standard for information delivery on the Internet.

2.3.2.3 Microsoft Office Products

Use of Microsoft’s Word document, PowerPoint presentation and Excel spreadsheet applications are preferred based on their inclusion in the DII COE suite of Common Support Applications. Until all DoD installations have received the current release (i.e., 97 SR-1), the least-common-denominator standards are Word 6, PowerPoint 4 and Excel 5. For users without these products, freeware viewers are available from Microsoft and other providers. Note that use of Office 95 and Office 97 products not current with the SR-1 upgrade to Office 97 can create interoperability and other difficulties.

2.3.2.4 ASCII Text

Defined in the previous section, ASCII text provides a mechanism for exchanging any textual content that does not require the inclusion of formatting specifications. Most operating systems, furthermore, have software bundled with them for browsing ASCII text (e.g., Microsoft QuickView, Notepad and WordPad).

2.3.2.5 Corel, Lotus and Other Compound Document Tools

Legacy installations of Corel's WordPerfect, Lotus 1-2-3 spreadsheet and Freelance presentation, and dBase database applications remain within the Department. Limited availability of freeware viewers complicates their use for multimedia applications between organizations. The more advanced versions of WordPerfect and Freelance provide better interoperability, though freeware viewers covering all formats are still not available.

2.3.2.6 RTF

Rich Text Format is an informal vendor specification published by Microsoft, based on an earlier IBM standard. RTF text is a form of encoding various text formatting properties, document structures, and document properties using the printable ASCII character set. Special characters can also be encoded, although RTF does not prevent the use of character codes outside the ASCII printable set. In addition, differences between implementations of RTF, especially in the area of special word processing features and graphics representations, may mean that not all features of a document can be preserved across an RTF conversion.

2.3.2.7 SGML

Standard Generalized Markup Language is a formal standard defined by ISO/IEC 8879, FIPS 152, and MIL-M-28001B. SGML is a meta-language that allows users to define, in machine-readable form, the structure and content of any class of documents. The standard specifies a method for creating document hierarchy models in which every element in a document fits into a logical, predictable structure.

SGML is able to separate the logical and physical structure of text. In this way, the standard is able to distinguish between the role of piece of text (e.g., caption, title, chapter, and index) and its appearance (e.g., typeface, font, size, and margin). This permits text to be tagged with descriptive markup, enhancing its functionality. By providing the ability to associate processing instructions with document markup, SGML includes a mechanism for referencing non-text forms within a document. By providing tags that enable query and hypertext capabilities, SGML is a standard that allows the production of intelligent documents for distribution and use on CD-ROM and other random access media. The SGML standard is useful to organizations that exchange information among systems, applications, departments, and users.

Use of SGML has largely been replaced in multimedia practice by one of its Data-Type Definitions (DTDs), HTML.

2.3.3 Issues/Discussion

All frameworks for assembling multimedia into compound documents in widespread use are based on proprietary standards. For this reason, no standard mechanism is mandated or preferred for which no freeware viewer is available.

Presentations (i.e., compound documents with primarily graphical or bullet-chart content) should be built using PowerPoint 4. If PowerPoint 97 is used, the file should be saved as PowerPoint 4 and its contents confirmed to not contain constructs that are incompatible with PowerPoint 4.

Compound documents with primarily text content should be built using Word.

Compound documents based on tables of data and graphical charts from analysis of that data should be prepared in Excel, though the results will be more portable if embedded in either a PowerPoint presentation or Word document.

All frameworks for assembling multimedia into compound documents in widespread use are based on proprietary standards. For this reason, care should be taken to ensure that the recipient of the compound document has at least access to a freeware viewer for that particular product’s format.

Mission specific standards, technologies and products include the ACR-NEMA (American College of Radiology and National Electrical Manufacturers Association joint committee) standard DICOM 3.0 (Digital Imaging and Communications in Medicine Version 3.0) for medical imaging, and MIL-STD-1840 and MIL-PRF-28001 for CALS (Computer-Aided Logistics Support).

2.3.4 Monitored and Related Standards, Technologies, Products and Practices

Portable document programs such as Envoy and Adobe’s Acrobat are not designed to replace HTML but rather to create a synergy between the two formats. All such products make it easier to move between portable documents and HTML pages by supporting more robust hypertext links. Those listed in this section have not, however, found widespread use.

2.3.4.1 OpenDoc

Several consortia are exploring OpenDoc and other *de jure* and *de facto* compound document standards, but none have received widespread incorporation in COTS products.

2.3.4.2 Hypermedia/Time-based Structuring Language (HyTime) and Extensible Markup Language (XML)

HyTime, like HTML, is an SGML DTD. Defined by ISO/IEC 10744, it is designed to incorporate many of the features now available in proprietary implementations of HTML 3.x, to which it generally corresponds. A companion DTD, called Extensible Markup Language (XML), is being coordinated specifically for improved support for multimedia. Neither has yet found widespread use. A style-sheet language companion DTD to XML, Extensible Style Language (ESL), has also been submitted to the W3C for consideration.

2.3.4.3 SMDL

Standard Music Description Language is a draft Internet standard defined by ISO/IEC 10743. SMDL defines a language for the representation of music information, either alone, or in conjunction with text, graphics, or other information needed for publishing or business purposes. Multimedia time sequencing information is also supported. SMDL is a HyTime application conforming to ISO/IEC 10744, Hypermedia/Time-based Structuring Language. SMDL is also an SGML application conforming to ISO 8879, SGML.

2.3.4.4 Common Ground

Common Ground is a proprietary electronic page viewer acquired by Hummingbird Inc., in 1995. This format specification is marketed as an aid to internal company communication, rather than a commercial publishing application. For simple electronic distribution of everyday documents, Common Ground offers a competitive alternative to Adobe Acrobat and PDF. The product comes with viewers for Mac, PC and UNIX, and it is independent of networks, media, and desktop platforms. Documents may be distributed by floppy disk, CD-ROM, or over networks. Delays in production of Common Ground have allowed Acrobat to gain momentum and become the de facto standard.

2.3.4.5 Envoy

Another contender in the portable document arena is Tumbleweed Software’s Envoy. It was licensed to WordPerfect and today is licensed by Corel with its Corel Office Suite for Windows and with WordPerfect. Envoy 7.0’s viewer is more robust than its competitors’, letting the user easily create indexed documents with annotations, tables of contents, links to other Envoy documents and to Web sites. The program’s ability to embed Envoy files in many applications, including some E-mail packages, and its fairly compact file sizes have earned it a loyal following.

2.4 DOCUMENT COMPRESSION

Compression of any files larger than 100Kbytes should be routine practice. Key compression algorithms are patented (IBM has a patent on the arithmetic code used in JBIG and many JPEG images). Their widespread use generally ignores this fact. The consequences are beyond the scope of this handbook.

2.4.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort PKZIP (GZIP, ZIP)	GCCS
Higher Level of Effort LZH, LZW, Z, StuffIt (Mac .sit and .sea), ARC, ARJ	Apple Computer, US Patent Literature
Monitored X/Open C604 Wavelet compression, fractal compression	The Open Group

PRACTICES	SOURCE
Preferred Compression of all documents larger than 100KB	RFC 821 (SMTP)
Discouraged Use of other patented algorithms	DII COE

2.4.2 Overview

The following compression and related standards are applicable to compound document interchange.

2.4.3 PKZIP

GZIP and ZIP (Lempel-Ziv 1977 or LZ77 with hashing, plus secondary static Huffman coding on a block basis) are implementations of compression and encoding standards developed and promulgated by PKWare, Inc. (owned by Phil Katz). Huffman coding is a statistical technique that gives a reduction in the average length used to represent the coded character data. GZIP is the GNU Unix implementation; it is frequently seen as documents with .gzip or .gz file extensions. Expect compression up to 4:1. ZIP compression is in widespread use and tools are commonly available. Note that the term PKZIP is often used to refer to the compression specification; in fact, this is the trademarked name of the PKWare, Inc. product. Other vendors' products, such as the popular shareware product, WinZip, from Nico Mak Computing, Inc., may implement portions of the ZIP algorithm in native code, without requiring concurrent licensing of the PKWare product.

2.4.4 LZH, LZW and Z

LZH (Lempel-Ziv-Huffman) is an LZ77 algorithm implemented with a tree data structure, plus secondary static Huffman coding on a block basis. Its LHA implementation (LHARC) is LZ77 with hash table and binary trees, plus secondary Huffman coding. ZOO is an LHA implementation under MS-DOS and is obsolete.

LZW (Lempel-Ziv-Welch) is an LZ78 algorithm, which adds a dictionary of commonly encountered words and phrases, with the dictionary index substituted for subsequent occurrences of the word or phrase. Z is the Unix implementation of LZW. LZW is patented by Unisys, complicating its enterprise-wide usage.

LZH and LZW provide higher compression than PKZIP, but the tools are less freely available. Note that LZW is the compression used in GIF, TIFF and PostScript-formatted documents and the v.42bis modem standard; the Portable Network Graphics (PNG) specification was developed to avoid the licensing difficulties use of LZW could present (i.e., when GIF images are used).

2.4.4.1 ARC and ARJ

ARC (for ARChive) is a PKWare implementation of LZW compression. ARJ (for Archive by Jung) is a PKWare implementation of LZ77 compression with hashing, plus secondary static Huffman encoding on a block basis.

2.4.4.2 SIT and SEA

These are proprietary (StuffIt) compression schemes for Apple Macintosh systems. SEA is a self-extracting archive format.

2.4.5 Compound Document Issues/Discussion

All frameworks for assembling multimedia into compound documents in widespread use are based on proprietary standards. For this reason, mechanisms for which no freeware viewer is available can prevent free interchange of compound documents between organizations.

2.4.6 Monitored and Related Standards, Technologies, Products and Practices

Wavelet and fractal-compression algorithms are receiving some attention in the press. Wavelet techniques offer more promise, since their basis in fast Fourier transform-like technology yields reasonable results when compared with JPEG compression at reasonable compression times and platform performance requirements.

X/Open C604 is an implementation of the “pack” algorithm (Huffman coding; less than 4:1 compression). Its Single UNIX Specification Version 2 is required for the Unix 98 Brand, Commands and Utilities, Issue 5. The Open Group has proponency for this standard.

Arithmetic coding based on probabilistic schemes, produces compression with allowable losses, but is generally applicable only to images and sound, where such losses are more easily tolerated.

JBIG, based on arithmetic coding, provides excellent image compression.

Lack of broadly available tools limits the utility of these methods.

.tar Files (Tape ARchive) are a mapping of multiple files, including those of different types, to a common file structure for manipulation and distribution as a single file. No compression is inherent in the tar format; compression is often added via GZIP (.tar.gz) or LZW (.tar.z) mechanisms.

.cab Files are Microsoft CABinets used typically to distribute software and related installation files used by Windows NT and 95 Installer software.

Commercial products based on the ANSI X3.225 compression specification are essentially restricted to tape cartridges.

2.5 FONT INFORMATION

Fonts are computer graphic representations of typeset alphanumeric and symbolic characters. As such, fonts are the mechanism for display of text information.

2.5.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
TrueType for Windows NT and 95	DII COE, AITS/ITSG
Hyperlink distinguished from regular underlined text	
Higher Level of Effort	
Adobe Type I for Postscript	DII COE, AITS/ITSG
Bitmapped fonts	Proprietary implementations
Monitored	
ANSI X3.45	ANSI

PRACTICES	SOURCE
Preferred	
Times New Roman 12pt for document body text	DII COE
Courier New for software module names	DII COE
Arial or Helvetica for spreadsheets	DII COE
Font color, capitalization, etc.	DII COE
Book Antiqua or Times New Roman over Palatino for presentations	Microsoft Knowledge Base
Discouraged	
Authoring with bitmapped fonts	Proprietary implementations

2.5.2 Overview

The fonts in primary DII usage will be TrueType for Windows NT and Windows 95 platforms, because those are provided native support by the operating system. All Postscript and PDF applications should use Adobe Type I fonts, since those are provided native support, though lack of native support on MS-Windows clients can be problematic.

Bitmapped fonts (i.e., based on raster images of the font as opposed to a combination of weighted stroke vectors) are to be avoided, since they are not readily scaleable and can occupy substantial processing and storage resources. The broad variety of available TrueType and Adobe Type I fonts generally obviates the need for bitmapped fonts, even for special applications.

Translators are available (e.g., Adobe Type Manager Deluxe 4.0) to convert from either TrueType or Adobe Type I fonts to the other. Additionally, commercial services are widely available for custom conversion of bitmapped fonts to TrueType or Adobe Type I fonts. Services are also available for casting special characters, including an individual's own handwriting or handprinting, as TrueType or Adobe Type I fonts.

Note that a bug in Windows 95 necessitates the use of alternatives to the Palatino font, owing to defects in kerning of characters, particularly the forward slash (/), that severely degrade the appearance and readability of the generated text. Microsoft recommends Book Antiqua as a close approximation to Palatino. Times New Roman is a suitable alternative if Book Antiqua is not available.

Palatino
The quick brown fox jumps over the lazy dog. 0123456789

Book Antiqua
The quick brown fox jumps over the lazy dog. 0123456789

Times New Roman
The quick brown fox jumps over the lazy dog. 0123456789

2.5.3 Issues/Discussion

No single font standard covers all operating environments. Multimedia authors must be cognizant of the font support generally available to consumers of their product.

Caution should be taken to ensure that restrictive licensing does not encumber use of any fonts. Many commercially produced fonts are not free for all uses. Copyright protection must be preserved where required.

2.5.4 Monitored and Related Standards, Technologies, Products and Practices

ANSI X3.45 was originally envisioned as a standard handprinting font, composed of unique character forms, stroke weight and direction to enable tablet-interface computing devices to "read" handwriting.

2.6 COLOR DEFINITION

Like alphanumeric text and symbols, colors must be encoded for representation by computers. These codes are typically representations of quantized voltage levels to be applied to monitor or printer electronics to generate the desired color by brightening or dimming individual "dots" on the monitor screen or page. Each of these "dots" responds to electrical stimulus by producing a particular color. In the case of a monitor screen, the dots are phosphors bonded to a substrate, diodes in the case of a flat panel display, or inks in the case of a printer. Each color encoding scheme is an alternative method for combining digital representations of these quantized values. For monitors and flat panel displays, adjacent dots are combined to produce a picture element, or “pixel.” The colors of each dot combine so that the pixel appears to the eye as a distinct color different from that of each of its component dots. Similar combinations are applied to the printed page by combining inks of different colors to adjacent dots on the page.

However, baseline settings of individual monitors can be different, lot-to-lot differences in printer inks exist, and individual humans perceive color differently. These differences mean that, where color reproduction is crucial to mission success (e.g., in teledermatology or photo-interpretation), a common, standard color representation must be used for both the authoring and display platforms. CIE specifications are widely accepted as industry standards for this purpose.

So-called RGB color is composed of red, green and blue digitized (i.e., integer-valued) color components with a portion of the total bits in each quantized color devoted to each component (i.e., 8 bits or 256 discrete levels each in the case of 24-bit color). Primary (i.e., red, green and blue) and secondary (cyan, magenta and yellow) colors are reproduced when individual component values are set to the minimum and maximum permissible values (i.e., 0 or 255, respectively in the case of 24-bit color). For example, values of 255 for red and 0 for green and blue would generate pure red. Values of 255 for green and blue and 0 for red would generate pure cyan. Mixing these and intermediate levels of red, green and blue makes possible a broad range of combinations of colors, from 16 for 4-bit color to 16.7 million possible combinations for 24-bit color. HTML and similar text representations of color encode RGB values as their hexadecimal values.

Printers normally use the CMYK equivalents (for Cyan, Magenta, Yellow and black; separate black ink is used to support frequent text printing and to generate more highly saturated black images than would be possible by merely combining cyan, magenta and yellow inks. CMYK representations are normally expressed as the percentage of each color component.

So-called grayscale images result when the image contains only equal values of red, green and blue (or cyan, magenta and yellow) combined. Note that monitor displays and printers produce opposite effects when combining colors. Combining the primary colors of red, green and blue to their maximum intensity on a monitor yields the color white. Doing the same on a printer (or with cyan, magenta and yellow) yields a color very close to black.

2.6.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
Grayscale through 24-bit color	GCCS, DII COE, AITS/ITSG, JIEO
CIE (Commission Internationale de l'Éclairage or International Commission on Illumination) specifications	GCCS, DII COE, AITS/ITSG, JIEO
Higher Level of Effort	
Process, PMS, Pantone, CMYK, HSI	GCCS, DII COE, AITS/ITSG
RGB, TIFF	GCCS, DII COE, AITS/ITSG

PRACTICES	SOURCE
Preferred	
16-bit maximum color depth for images other than requiring imagery interpretation	DII COE
8-bit maximum color depth for non-imagery graphics	DII COE
No use of “reserved” colors (i.e., using only 216 out of 256 possible)	DII COE
Use of HSI representations for communicating color values	DII COE
Entropy-based (i.e., error diffusion) dithering for image palette assignments	DII COE
Nearest-color matching for graphics palette assignments	DII COE
Discouraged	
Use of 24-bit color or error diffusion dithering for all applications	Common usage
Use of RGB representations for communicating color values	Common usage

2.6.2 Overview

In order to ensure that colors in source material are accurately reflected in multimedia products when viewed, a common set of color definition standards is required. Choice of encoding scheme should be matched to the depth of color required (i.e., the number of different levels of hue, saturation and intensity). In selecting a target depth of color, consideration should be given to required color fidelity (i.e., how closely the colors in the generated image should match those in the source image). Since displayed color is based on perception by the human eye, the ability of the eye to resolve slight differences between colors should be taken into account. For example, an image with 8-bit color depth (i.e., a total of 256 possible colors) will have distinct quantization (where fewer, abrupt color transitions appear) caused by the assignment of integer values during color encoding, meaning that the image will look like a "digital" image. An image with 24-bit color will appear like a photograph, with smooth color transitions made possible by the greater range of possible integer values.

Use of CIE specifications for color-critical applications helps to ensure that both author and viewer see and perceive the exact same combinations of hue, saturation, and intensity in each displayed color. The capability to display color representations from grayscale through 24-bit color is necessary to accommodate the broad range of image products that may be presented in multimedia format. Although process, Pantone, and other color representations and transformations may be necessary for specifying colors to be transferred to printed images, software support for these specifications is limited. Direct software support for RGB and TIFF formats is similarly limited, and their specific implementations may differ between software products, complicating their use for conveying color information between organizations.

2.6.3 Issues/Discussion

2.6.3.1 Red-Green-Blue (RGB) and Cyan, Magenta, Yellow and black (CMYK) Representations

RGB and CMYK representations facilitate processing pixel codes into screen and page images, especially in 24-bit form, wherein individual 8-bit bytes can be used to discretely represent each color (See Figure). This simplifies the computer instructions needed to move color codes into and out of memory and to transfer them across the computer's data bus to the interface to the output devices. However, the human eye does not respond equally to red, green and blue shades. The eye is very sensitive to slight changes in red intensity, less so to changes in the green portion of the spectrum, and not very sensitive at all to differences between blue intensities. Therefore, common RGB encoding practice means that red intensity encoding is less than optimal, while much of the available blue color intensity discrimination is wasted. Offsetting this, however, is the fact that human discrimination of individual blue colors is extraordinary, so reducing the number of bits allocated to blue to compensate for intensity differences will still result in marked artifacts from the quantization required.

2.6.3.2 Hue, Saturation, and Intensity (HSI) Representations

One result of this facet of human color perception is that multiple combinations of red, green and blue can produce color values within a pixel or printer dot that are perceived as essentially identical (i.e., multiple points within the RGB "cube" produce identical colors).

In fact, other than for primary and secondary colors and shades of gray, no good means exists to establish on an *a priori* basis what combination of red, green and blue values will yield a particular desired color.

To resolve this dilemma, a common transform for color representation and encoding is the so-called HSI, for Hue, Saturation and Intensity. HSI is also known as HSL, with the L representing Luma, and HSB, with the B representing Brightness. HSI provides a color encoding representation that more readily matches human color perception. Altering the values of each produce individually distinct colors. Hue can be thought of as the "pure" color or shade. Saturation is the degree to which that pure color dominates the color value. Intensity is the relative amount of white or black in the color. HSI values are often portrayed as two cones joined at their bases. Hue values are portrayed as rotation about the circumference of the cones, with pure red at 0 and the other primaries and secondaries spaced at 60-degree increments around the circumference. Saturation values are portrayed as the percentage of the total radius occupied by a line drawn from the cones' center axis to the color value. Intensity is represented by the percentage of the distance from black (0%) to white (100%) along the central axis of its intersection with a plane perpendicular to the central axis and containing the color value. (See Figure)

24-bit color representations in HSI can introduce fractional values that result in poor conversion to and from RGB, as a result primarily of errors in rounding to the integer values required.

The paper at:

http://www.research.microsoft.com/research/graphics/Alvy/papers/hwb/hwb_rgb.htm

provides equations for transforming between RGB and HSI.

2.6.4 Monitored and Related Standards, Technologies, Products and Practices

Digital color encoding requires that a combination of actual colors in an image be represented as drawn from a palette rather than a continuum. For this reason, the large number of possible palette values (16.7 million) in 24-bit color has given 24-bit color the common name, "True Color." True 16-bit color, with 65,536 possible values is often termed, "High Color" or "HiColor." True Color and HiColor images are usually indistinguishable from a film-based photograph. An alternative 16-bit color implementation, Targa-16, uses 5 bits each for red, green and blue, for a total of 32,768 possible values. Images encoded to the Targa-16 specification are usually indistinguishable from a standard television image.

Higher color depths support a technique called "antialiasing", in which slight modifications to colors of adjacent pixels are made to fill in the "stair-step" appearance caused by quantization of the colors in forming the digital encoding. This technique "fools" the eye into perceiving a smooth, linear transition rather than an abrupt transition, making edges appear sharper, lines straighter, etc. Antialiasing, therefore, trades color and spatial resolution for improved edge perception.

A related technique called "dithering" is used to transform images between different color depths. Dithering represents intermediate colors through selection of colors in adjacent pixels. Dithering, therefore, trades spatial resolution for improved color perception.

2.7 VECTOR GRAPHICS

2.7.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	**SOURCE**
Lower Level of Effort	
Win32 API for Windows (WinG optional) VGA and SVGA resolution FIPS 128-1 CGM 93 vector graphics	GCCS, DII COE, AITS/ITSG
Higher Level of Effort	
ISO CGM 89, GKS, IGES, WMF, DXF	GCCS
Monitored	
DIGEST	DIGITAL GEOGRAPHICAL INFORMATION WORKING GROUP
STEP	ISO/IEC 10303

PRACTICES	**SOURCE**
Preferred	
Win32 API for Windows clients (WinG optional) FIPS 128-1 CGM 93 for Unix clients	DII COE
Discouraged	
Use of resolution outside the range 640x480 to 1024x768	Common usage

2.7.2 Overview

Vector graphic standards are used to display images using point, line, and character plotting commands. Vector graphics is distinguished from raster graphics by the absence of specific information about each pixel of the screen. For example, a line drawn for vector graphics would contain the location of the endpoints and a description of the style (width, dashed or solid, etc.).

2.7.2.1 Windows 32-bit API (Win32), Windows Game Interface (WinG) and Windows MetaFile (WMF)

These are proprietary formats developed by Microsoft for use in the MS-Windows environment. WinG is an enhancement to Win32 that provides support for accelerated drawing of vectors, shapes and other constructs. Windows MetaFile (WMF) supports both vector and raster objects. All receive native support in modern versions of the Windows operating system (i.e., Windows 95 and Windows NT 4.0).

2.7.2.2 CGM

Computer Graphics Metafile is a formal standard defined by ISO/IEC 8632 and FIPS 128-1. CGM is a computer- and operating system independent interchange format. There are elements to represent both geometric graphics content (e.g., circles, polygons) and raster graphics content (e.g., pixel arrays). An element's appearance can be affected by attributes (e.g., line cap, line join, miter limit, fill pattern). Several color models (Red, Green, Blue (RGB), Cyan, Magenta, Yellow, blacK (CMYK) color model, etc.) are supported. CGM is widely used in the publishing industry.

The elements contained in a CGM file represent a wide range of picture types. The elements are split into functional groups that delimit major structures: (1) define the representations used within the metafile; (2) control the display of the picture; (3) perform basic drawing functions; (4) control the attributes of the basic drawing actions; and (5) provide access to non-standard devices.

The file structure is defined to allow sequential access and random access to individual picture elements. Elements can also be grouped into logical or functional segments, allowing all picture elements of a certain type (all shoreline elements for instance) to be grouped together.

2.7.2.3 IGES

Initial Graphics Exchange Specification is a formal specification defined by ANSI/US PRO/IPO-100-1996, MIL-D-28000A, and FIPS 177. IGES Version 5.0 is an ANSI standard developed by the American Society for Mechanical Engineers (ASME) for the exchange of 3D data in vector file format, particularly for documents prepared in Computer-Aided Design/Computer Aided Manufacturing (CAD/CAM).

2.7.2.4 DXF

Document Exchange Format is an informal vendor specification that defines an AutoCAD 2-D graphics file format. DXF is Autodesk's format for moving AutoCAD drawings to and from the rest of the world. Many CAD systems import and export the DXF format for graphics interchange.

2.7.3 Issues/Discussion

Current DII COE Windows NT client platforms do not provide native support for CGM vector graphics.

Vector graphics are used when computer-generated graphics are to be portrayed, as opposed to images that are photographed.

A substantial body of AutoCAD DXF formatted drawings exists throughout the Department of Defense, especially for facilities.

A large quantity of clip-art is available in WMF format.

Tools for the display of CGM and IGES format graphics are not widely available on Windows platforms.

2.7.4 Monitored and Related Standards, Technologies, Products and Practices

2.7.4.1 DIGEST

The Digital Geographic Information Exchange Standard (DIGEST) permits the exchange of digital geographic data on magnetic tape. It accommodates the exchange of multiple data sets of different data structures using a single format. Particular geographic information that it supports includes vector topologically structured data, color-coded raster data, and feature identification using a feature attribute-coding catalog.

DIGEST was developed by the eleven-nation Digital Geographical Information Working Group. The format implementation is compliant with ISO 8211, Specification for a Data Descriptive File for Information Interchange.

Minimum Standards Specifications are each oriented toward particular single data structures. The Generic Standard contains the necessary file, record, field, and subfield details to exchange all data structures supported by the standard.

2.7.4.2 STEP

Standard for the Exchange of Product model data is a formal specification defined by ISO/IEC 10303. STEP provides a representation of product information along with the necessary mechanisms and definitions to enable product data to be exchanged. The exchange is among different computer systems and environments associated with the complete product life cycle including product design, manufacture, use, maintenance, and final disposition of the product. This representation is suitable for file exchange, as well as for implementing and sharing databases of archived information.

2.7.4.3 GKS

Graphical Kernel System is a formal standard defined by ISO 7942, ANSI X3.124, FIPS 120 and ISO 8805 for GKS-3 Dimensional. GKS is a machine- and language-operating system and device-independent specification of set of services for displaying and interacting with 2D and 3D pictures.

2.7.4.4 Standards Supporting Specific Mission Applications

MIL-PRF-28000, 28002 and 28003 standards are used for CALS applications. Vector Product Format (VPF) is recommended for geospatial applications.

2.8 THREE-DIMENSIONAL AND ANIMATED GRAPHICS

2.8.1 Guidance

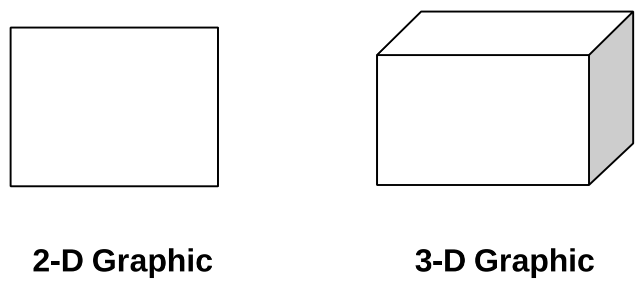
STANDARDS, TECHNOLOGIES AND PRODUCTS	**SOURCE**
Lower Level of Effort	
Win32 API for Windows (WinG optional)	GCCS, DII COE, AITS/ITSG, JTA
VGA and SVGA resolution	GCCS, DII COE, AITS/ITSG, JTA
OpenGL 3-D graphics	GCCS, DII COE, AITS/ITSG, JTA
AVI, QuickTime, layered image animation	GCCS, DII COE, AITS/ITSG, JTA
Higher Level of Effort	
ISO CGM 89, IGES, WMF, DXF	GCCS
Renderman 3-D graphics	AITS/ITSG
FLI/FLC animation	AITS/ITSG
Monitored	
VRML and DirectX 3D 3-D graphics	AITS/ITSG, Microsoft
Togglethis	Emerging technology

PRACTICES	**SOURCE**
Preferred	
Win32 API for Windows clients (WinG optional)	DII COE
OpenGL for Unix clients	DII COE
AVI for authoring large image animation with frame counts greater than 60	DII COE
Layered image animation for small images or frame counts less than 60	DII COE
Discouraged	
Use of resolution outside the range 640x480 to 1024x768	DII COE
Any animation looping other than layered image animation	DII COE

2.8.2 Overview

Three-dimensional and animated graphics are receiving increased emphasis as the basis for "virtual reality" simulations, including "fly-through" tours of facilities at the design stage, mission rehearsal, and other applications.

These graphics are distinguished from 2D graphics by the modeled geometry, transformations applied to portrayed objects, attributes associated with portrayed objects, and modeling of lighting, shading, and surface textures.



Geometry is the most obvious difference between 2D and 3D graphics, adding the Z-plane to the XY plane of 2D graphics. Additionally, the geometry modeled in 3D graphics has additional attributes or properties (e.g., how light reflects of the surface based on type of "material" portrayed). Transformations (e.g., scaling, rotating, and translating) are more complex because of the introduction of both the Z-plane and object attributes (i.e., a rotated object exposes different surfaces to a light source, with possibly different reflectance results.

Light sources are modeled as objects, also, each with their own attributes (e.g., spotlights, point-source bulbs, sun, directional, ambient – diffuse).

Shading attributes are used to determine the distribution of reflected colors across an object's surface, as defined by the lighting model. In Gouraud shading, reflected color is calculated at each of the object's vertices and then smoothly interpolated between adjacent vertices. The simpler flat shading is used when each facet of the object has the same color.

Lighting models further determine the amount of light that impinges on an object, how much of that incident light should be reflected, and in what directions. Directional reflection attributes include diffuse reflection (i.e., independent of viewing angle) and specular reflection (i.e., like a mirror, dependent on viewing angle).

2.8.2.1 OpenGL

OpenGL is a proprietary but *de facto* API for applications generating interactive 2D and 3D graphics. It is designed to be operating system, windowing-environment, and hardware independent. OpenGL was originally developed by SGI, but is now supported by at least 50 corporations, including all the major PC and workstation manufacturers and software developers (e.g., IBM, Sun, Microsoft, HP, Compaq, and DEC). OpenGL is available for PCs and workstations.

2.8.2.2 Virtual Reality Modeling Language

Virtual Reality Modeling Language (VRML, pronounced "V R M L" or "vermel"), is a subset of the Inventor File Format (ASCII) with additions to permit hyperlinks to other Web resources, similar to HREF anchors in HTML. In addition, VRML provides for specification of LOD (level of detail) to match an appropriate amount of transmitted data based on how prominent the portrayed object is, or the rendering speed of the browsing machine, or both.

VRML 1.0 is the currently coordinated standard; VRML 2.0 is under development.

VRML 1.0 provides the following simple shapes: Cube, Cone, Cylinder and Sphere and the IndexedFaceSet.

2.8.2.3 DirectX

DirectX and its Direct 3D implementation are proprietary, *de facto* standards developed largely by Microsoft. DirectX is implemented as a set of Dynamic Link Libraries (DLLs) for Windows 95 and Windows NT. DirectX provides facilities for programmers to more directly control hardware to achieve enhanced performance in 3D graphics.

2.8.3 Issues/Discussion

Most 3D graphics standards are based on proprietary technology.

3D graphics are used when computer-generated images are to be portrayed, as opposed to images that are photographed or filmed.

When selecting a particular technology for implementation, care must be taken to meet any licensing requirements.

2.8.4 Monitored and Related Standards, Technologies, Products and Practices

2.8.4.1 Renderman

Renderman is a proprietary 3D-representation language similar in construct to PostScript, developed by Pixar. Pixar’s objective was to provide a standard mechanism for sending device-independent 3D primitives to rendering systems, without requiring knowledge of the rendering algorithms applied by those systems. The general lack of readily available cross-platform Renderman creation tools and freely licensed Renderman cross-platform viewers interferes with interoperability.

2.8.4.2 FLI/FLC

AutoDesk developed the FLIC (i.e., FLI/FLC) file format for animation of its AutoCAD drawings. FLI files are an older format and are limited to a resolution of 320x200 pixels. FLC files can display animated images of up to 640x480 pixels at reasonable frame rates.

FLIC files store animation as a series of frames. Each frame contains an image and associated image attributes (e.g., palette, and label). A “compatible” FLIC file contains a ring frame at the end, so that the animation can be looped without a perceptible pause between the last frame and the (repeated) first frame. Ring frames are necessary, because decompression of the first frame is generally slower than intermediate frame updates, since the first frame contains a complete image. A segmented FLIC file optionally contains a ring frame per segment.

FLIC animation can be displayed in larger image sizes with typically better performance than MPEG-encoded images of the same animation. However, the general lack of readily available cross-platform FLIC creation tools and freely licensed FLIC cross-platform viewers interferes with interoperability.

2.8.4.3 Togglethis Interactive Character Technology

Interactive Character Technology is a relatively new animation technique based on proprietary approaches developed by Toggle Entertainment, Inc. Used primarily for entertainment thus far (e.g., games, comics), interactive characters can be used as messengers of information, web page hosts, advertising spokescharacters, actors or teachers. Interactive characters can be also have personalities ascribed to them, including scripting across a range of emotions and differing responses to circumstances encountered, and to demonstrate aspects of memory and recall.

2.9 STILL (RASTER) IMAGES

In digital multimedia, an image is visual information represented and stored as a bitmap, a spatial matrix made up of individual picture elements or pixels. Each pixel is coded with a finite number of bits, ranging from 1 to 24 or more, representing the color assigned to that pixel. Images may be captured from the real world by scanning of printed material or by video camera. Images may be static - called still images - or moving. Images may also be bi-tonal or continuous tone. In contrast to graphics, images are not revisable, though they can be retouched by paint editing tools. Image is synonymous to picture. (Fluckiger, page 581).

2.9.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
GIF 89a	GCCS, JTA, DII COE, AITS/ITSG
JPEG	GCCS, JTA, DII COE, AITS/ITSG
Group 3 and MIL-STD 188-161D for telefacsimile (FAX) images	GCCS, JTA, DII COE, AITS/ITSG
Higher Level of Effort	
TIFF R and G, Photo CD, PDF, DIB/BMP, WMF, GIF87a, Framemaker Interchange Format, JBIG, Lotus PIC, MacPaint, MS Paint, PCX, PICT, PPM, Sun raster, WPG, XBM, XPM, XWD, TGA, Postscript, EPS, IGES	GCCS, JTA, DII COE, AITS/ITSG
Group 4 FAX	GCCS, JTA, DII COE, AITS/ITSG
Other proprietary formats	Common usage
Monitored	
Interlaced PNG for applications requiring lossless compression	JTA
Wavelet-compressed images	Emerging technology
Fractal-compressed images	Emerging technology

PRACTICES	SOURCE
Preferred	
Interlaced GIF 89a for applications requiring lossless compression	DII COE
Progressive JPEG for applications permitting lossy compression	DII COE
Discouraged	
Use of resolution outside the range 640x480 to 1024x768	DII COE
Any looping of animated GIF images	DII COE

2.9.2 Overview

2.9.2.1 Raster Graphics

A raster format uses a column and row arrangement of pixels to represent an image or graphics.

2.9.2.2 GIF

Graphics Interchange Format is an industry proprietary specification published by CompuServe. GIF defines a protocol intended for the on-line transmission and interchange of raster graphic data in a way that is independent of the hardware used in the creation or display.

GIF is defined in terms of block and subblocks that contain relevant parameters and data used to reproduce a graphic. A GIF data stream is a sequence of protocol blocks and subblocks representing a collection of graphics. In general, the graphics in a data stream are assumed to be related to some degree, and to share some control information.

A data stream may originate locally, as when read from a file, or it may originate remotely, as when transmitted over a data communications line. The Format is defined with the assumption that an error-free Transport Level Protocol is used for communications; the Format makes no provisions for error-detection and error-correction.

GIF 89a is the preferred version of the GIF standard, because it permits both the definition of a transparent background color and interlacing for progressive image display, both key components of successful incorporation in HTML. Use of interlaced GIF images is preferred for inclusion of GIF images in HTML, to permit successively improved image appearance without requiring transmission of the entire image. This allows the user the freedom to decide whether to stop page display, go immediately to another hyperlink on the page, or wait for full page display. Note that the compression in GIF images is created using LZW algorithm patented by Unisys, which may complicate its enterprise-wide usage.

2.9.2.3 JPEG

Joint Photographic Experts Group is the name of the committee that wrote the standard for Digital Compression and Coding of Continuous-Tone Still Images. JPEG is a formal standard defined by ISO/IEC 10918 and ITU-T Recommendation T.83. This standard sets out requirements and implementation guidelines for continuous-tone still image encoding and decoding processes, and for the coded representation of compressed image data. These processes and representations are intended to be generic, that is, applicable to a broad range of applications for color and gray-scale images within communication and computer systems. JPEG is normally lossy, meaning the image that comes out of decompression is not identical to what went in. The algorithm achieves much of its compression by exploiting known limitations of the human eye, notably, the fact that small color details aren’t perceived as well as small details of light and dark. Thus, JPEG is intended for compressing images that will be looked at by humans and not analyzed by machine.

A useful property of JPEG is that adjusting compression parameters can vary the degree of lossiness. This means that the image-maker can trade off file size against output image quality. Images can be compressed into extremely small files if poor quality is not a problem. Less severe compression results in better quality images but larger files. JPEG can be adjusted to be lossless, but the compression ratio drops to about 2:1. Note that iterative, nested creation of JPEG images (e.g., after editing of an existing image) will result in loss of image quality with each iteration.

As with GIF images, progressive implementations of JPEG images are strongly suggested for inclusion in HTML, to permit successively improved image appearance without requiring transmission of the entire image. This allows the user the freedom to decide whether to stop page display, go immediately to another hyperlink on the page, or wait for full page display.

2.9.2.4 ITU -TSB T4 and T6 (Group 3 and Group 4 Telefacsimile)

These are formal standards for Facsimile Coding Schemes and Coding Control Functions for Group 3 and Group 4 Facsimile Apparatus published by ITU. These recommendations specify interchange of optically scanned engineering drawings and pages of technical publications, and define directives to compress raster graphics in order to reduce file size and transmission time. The compression method is bit preserving, which means it is distortion-less and that the final decoded image is identical to the original.

The T4 recommendation is applicable to Group 3 telecopy machines while the T6 recommendation is applicable to Group 4. Technically, the biggest difference is that T6 has better compression capabilities than T4. Most telefacsimile (FAX) products are based on these standards.

These widely used compression techniques for bitonal raster data can use a combination of different techniques, including 2D Read coding, 1D Huffman coding, and uncompressed mode (bitmap). The data themselves do not carry information about the number of pixels per line or the total number of lines in the image. That means that pixel and line information must be exchanged in a header (when used in computer networks) or an enveloping protocol such as T.30, when used in the Public Switched Network for telefacsimile (FAX) transmissions.

2.9.2.5 Other Formats

2.9.3 TIFF

Tagged Image File Format is a *de facto* vendor specification published by Microsoft, Aldus, and HP. TIFF is used for desktop publishing, fax, and scanner data exchange. It defines a complete format for general raster interchange, creating bitmap files. It is one of the most flexible and complicated formats. There are many versions, and no application supports all versions.

In the design of the TIFF format, a great deal of effort was taken to provide for extensibility while maintaining backward compatibility. The only demand that TIFF places on the individual operating system is that the associated storage medium supports a file structure making it almost operating-system independent. In addition, the overhead of the format is quite low for the level of sophistication it possesses, making it fast and efficient to access. For these reasons, TIFF has become a popular format among distributors of digital images and peripheral manufacturers.

With the introduction of TIFF version 6.0, direct JPEG compression was introduced. In this version, several newly defined fields provide information required by JPEG software to decompress an image such as the type of JPEG algorithm used and byte offsets to the required quantization tables.

2.9.3.1 PICT

PICT is the designation for the Apple MacPaint format, an informal vendor specification published by Apple Computer. Apple MacPaint is an older raster metafile format for Macintosh applications. Only fixed size (576x720) monochrome images are supported. Simple run-length compression is used.

Apple MacPaint is used to interchange graphics data among nearly all Macintosh applications. It is not widely used on PCs, but some file conversion programs do support importing it (such as .PCT files).

2.9.3.2 Photo CD

Photo CD is a Kodak proprietary format for storing photographs on diskette and CD-ROM. Since its use is supported within the film developing industry and is one of the primary mechanisms for converting film products to digital raster images, software products are generally available for viewing and transforming Photo CD images to other formats. When authoring multimedia products based on Photo CD images, these images should be converted first to other, more portable formats, such as JPEG. Direct use of Photo CD images, however, is not encouraged because of cross-platform interoperability shortfalls.

2.9.3.3 Device-Independent Bitmap (DIB), Windows Bitmap (BMP) and Windows Metafile (WMF)

Device-Independent Bitmap (DIB), Windows Bitmap (BMP) and Windows Metafile (WMF) are native formats to the Microsoft Windows operating environment. For this reason, software products are generally available for viewing and transforming these images to other formats. When authoring multimedia products based on these images, they should be converted first to other, more portable formats, such as JPEG. Direct use of these images, however, is not encouraged because of cross-platform interoperability shortfalls.

2.9.3.4 Other Formats

Framemaker Interchange Format, Lotus PIC, MacPaint, MS Paint, PCX, WordPerfect Graphics (WPG), and Targa-16 (TGA) are all proprietary formats whose use may present cross-platform interoperability shortfalls.

PPM, Sun raster, XBM, XPM, XWD are X-Windows environment raster formats whose use may present cross-platform interoperability shortfalls.

2.9.4 Issues/Discussion

The predominant lossless compression technique, GIF, is encumbered with patent licensing concerns over its use of LZW compression.

All photographs and other raster images may be rendered using the formats and techniques provided by these standards.

Only for the lower-level-of-effort standards are cross-platform tools generally available for displaying and manipulating the images produced.

2.9.5 Monitored and Related Standards, Technologies, Products and Practices

2.9.5.1 PNG

PNG is an industry, *de facto* PNG (Portable Network Graphics)-based graphics specification published by CompuServe. It is a fully open 24-bit lossless graphics specification for electronic graphics exchange. It is a significant enhancement to the earlier GIF 89a specification, while also eliminating the proprietary LZW software, replacing it with compression technology compliant with the PNG specification. The specification was developed as a collaborative effort between CompuServe and the Internet Group.

CompuServe believes that the new specification closely meets the future requirements for graphics interchange on the Internet, on CompuServe, and on other services, as well as for exchange of information between graphics software products. PNG makes use of a data compression technology called “deflation” used in the freeware Info-Zip programs.

2.9.5.2 IPI-IIF

Image Processing and Interchange - Image Interchange Format is a formal standard defined by ISO/IEC 12087-3. The IIF is part of the first International IPI standard, which is under elaboration by ISO/IEC JTC1/SC24. IIF comprises both a data format definition and a gateway functional specification.

Part 1 of the IPI specification provides a platform-independent architecture and a set of common image-related data types, operations, etc. Part 2 provides an Application Programmer's Interface (API) for a useful set of image processing primitives, thereby promoting program portability. Part 3 (the IIF) provides an image interchange format, richer than either CGM or any de facto standard format (like TIFF) and promotes transparent data exchange. Part 5 of the specification, Basic Image Interchange Format (BIIF), is under development. BIIF is being developed in parallel with National Imagery Transmission Format (NITF) version 2.1 to provide an international standard to be used instead of MIL-STD-2500A which is used now in DoD.

The main component of the IIF is the definition of a data format for exchanging arbitrarily structured image data. The IIF defines a format that can be used across application boundaries and that can easily be integrated into international communication services. Besides the definition of a file format, there are definitions of parsers, generators, and format converters to enhance open image communications.

Besides the data format specification, the IIF also encompasses functionality for generating and parsing image data, for compressing and decompressing, and for exchanging image data between the application program, the Programmer's Imaging Kernel System (PIKS) (which is part 2 of the IPI standard), and storage/communication devices. This functionality is located in the so-called IIF Gateway. The IIF gateway controls image data being imported and exported to and from applications, as well as to and from the PIKS.

2.9.5.3 JBIG

The Joint Bi-Level Imaging Group defined a formal standard promulgated as ISO/IEC 11544 and ITU-T Recommendation T.82. JBIG defines a method for lossless compression encoding of binary (one bit/pixel) bi-level (having only two colors such as black and white) image, but can also be used for coding grayscale and color images. The method is bit-preserving and has "progressive" capabilities which permit the design of an application with one common database that can efficiently serve output devices with widely different resolution capabilities, and which can provide superior image browsing for an application using low-rate and medium-rate communication links.

JBIG can be used on gray scale, or even color images, by simply applying the algorithm one bit-plane at a time. This works well up to about 6 bits per pixel, beyond which JPEG's lossless mode works better. The Q-coder must be used with JPEG to get this performance. Since it is lossless, JBIG can be used for storing document images and they'll be legally admissible as exact replicas of the originals. Moreover, JBIG improves compression ratios by 40% (on simple documents) to 180% (on complex images) over ITU-T Group 4, but takes two to five times as long to compress and decompress in software. However, it should be remembered that IBM has a patent on the arithmetic coding used in JBIG; this may complicate exchange of JBIG images between organizations.

2.9.5.4 Still Picture Interchange File Format (SPIFF)

SPIFF is the "official" JPEG file format. Part 3 of the JPEG standard (ISO 10918) now includes a fully defined file format for storing JPEG data. When the JPEG format was standardized, disagreements among ISO committees prevented a standard JPEG file format from being created. The *de facto* format that appeared was JFIF (JPEG File Interchange Format) from C-cube Microsystems. The JFIF format, although now quite widespread, is very limited in capability as file formats go.

SPIFF is intended to replace the JFIF file format, adding features (more colorspaces, a recognized way of including text blocks, and so forth), and providing a backwards-compatibility allowing SPIFF files to be read by most JPEG/JFIF decoders. JFIF, however, has a five-year head start on SPIFF, so the likelihood of it being completely replaced soon is not high.

2.9.5.5 Standards Supporting Specific Mission Applications

In addition to the aforementioned DICOM 3 for medical imaging and CALS image formats for CALS applications, the National Imagery Transmission Format (NITF) standard is used for imagery requiring exploitation.

2.10 AUDIO

2.10.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
ISO/IEC 11172-2 (MPEG-1), ISO/IEC 13818-2 (MPEG-2)	GCCS, JTA, DII COE, AITS/ITSG, USIGS
ADPCM, WAV, MIDI	GCCS, JTA, DII COE, AITS/ITSG, USIGS
Higher Level of Effort	
VOC, AIFF, PCM	GCCS, DII COE, AITS/ITSG
Monitored	
Internet Phone, streaming formats (e.g., Streamworks, RealPlayer)	JTA, Netscape Communications
MPEG-4 for low data rates (including video/audio mix)	GCCS, JTA, DII COE, AITS/ITSG, USIGS

PRACTICES	SOURCE
Preferred	
Support for stereo and mono audio	DII COE
11KHz, 8-bit mono sampling for voice-only streams	DII COE
MIDI for instrumental-only accompaniment	DII COE
Discouraged	
Looping except background audio	DII COE
44KHz, 16-bit sampling for voice-only streams	DII COE

2.10.2 Overview

Audio is sound (usually in air) which has been or will be converted from an electrical signal to audible sound waves by one or more speakers. Within human hearing, audible frequencies range from approximately 20Hz at the low end to a high of around 20-22,000Hz. Audio is processed in a computer by sampling the analog signal at discrete time intervals and representing those samples as digital codes. This may be referred to as digital audio.

Audio fidelity, the degree to which the audio signal reproduces the original sound, depends largely on five factors:

- Signal-to-noise ratio and frequency response of the microphone used to capture the (analog) sound
- Signal-to-noise ratio and frequency response of the speaker(s) used to output the (analog) sound
- Sampling rate, the number of samples collected per second, expressed in kHz since each sample represents one state change in the sound signal
- Sample size, the number of digital bits used to represent each sample
- Number of channels, ranging from one (monaural sound or "mono") to two (stereo) to four (quadraphonic) or more.

Sampling rates in common use include approximately 11kHz ("telephone-quality"), 22kHz ("tape-quality") and 44kHz ("CD-quality"). Normally, the sampling rate is matched to the type of sound being sampled, with 11kHz used for human speech with 44kHz reserved for complex music. Use of higher-than-standard sampling rates may afford slightly better fidelity by "oversampling" resulting in a better fit to the analog waveform.

Sample sizes in common use are 8 bits (256 different values per sample) and 16 bits (65,536 different values per sample). The larger the sample size, the more precisely the fit to the analog waveform. In addition, larger sample sizes afford the opportunity to include multiple tracks of sound data. They also afford the opportunity to improve sound quality through signal processing, much as image processing can be used to improve image quality when larger color depths are recorded.

2.10.2.1 Motion Picture Experts Group (MPEG) Formats

Since these formats were designed for mixed audio and video, though they may be used quite efficiently for audio alone, a more complete description is provided below, in Section 2.11. Although the original MPEG-1 specification supported audio encoding, current practice generally limits encoding of audio to MPEG-1 and MPEG-2 Audio Layer 3 (MP3). MP3 encoding, the highest-compression audio layer defined within the MPEG-2 standard, typically yields 10:1 to 12:1 compression while retaining the original CD-quality audio.

2.10.2.2 MIDI

The Musical Instrument Device Interface (MIDI) is not a digitization of analog sound, but rather is like sheet music, in that it describes the notes, voices and duration of tones to be produced by sound generation chipsets on the audio board of a computer. For this reason, it is generally impossible to transform a digital audio stream (e.g., WAV or AIFF) to a

MIDI instruction set. The opposite, however, is true – the “tune” represented by a MIDI stream can be recorded as a digital audio stream (e.g., WAV or AIFF), though MIDI files are much smaller and are therefore recommended for instrumental sequences where practical.

2.10.2.3 AU

AU (AUdio) is the NeXt/Sun/DEC Unix standard format for digitized analog sound. Also seen as .snd files. Supports stereo as well as mono audio.

2.10.2.4 AIFF

Apple developed the Audio Interchange File Format (AIFF) for storing sampled audio and musical instrument information. AIFF has also been exploited on SGI and Commodore Amiga platforms. A compression-based extension, called AIFC or AIFF-C, provides more compact storage. Supports stereo as well as mono audio.

2.10.2.5 RIFF and WAV

Resource Interchange File Format (RIFF) is a Microsoft extension of the Windows Device-Independent Bitmap format, structured into “chunks” and “sub-chunks”. Each chunk has a four-character tag identifying its associated file or media type. The RIFF file itself is defined as a chunk containing other chunks. RIFF files can include WAV chunks, AVI chunks, and MIDI chunks.

WAVe (WAV) is a common PC-based Pulse Code Modulation (PCM) digitized analog sound format developed by Microsoft and IBM. WAV format supports stereo as well as mono audio. WAV is a subset of the Microsoft-developed RIFF file format for encoding multimedia data. WAV files can and do, therefore, contain non-audio information, such as copyright information and author/artist identification.

2.10.3 Audio Compression Algorithms

Digitized analog audio signals are difficult to compress because of the broad variation in data values. For 8-bit digital audio, encoding of the changes between successive sound samples, including silence detection, gives reasonable compression; 16-bit digital audio is much more difficult to compress, because of the broader variations possible.

2.10.3.1 ITU-T G.721 and G.723 (Adaptive Delta Pulse Code Modulation, ADPCM)

ADPCM provides a public standard for adaptive compression of digital audio signals. ADPCM, for example, converts 16-bit linear pulse-code modulation (PCM) samples and converts them to 4-bit samples (i.e., 4:1 compression).

2.10.3.2 ITU-T G.711 (Mu-Law and A-Law Pulse Code Modulation, PCM)

In PCM, sound is sampled according to a set sampling frequency (e.g., 8KHz), and each sample is assigned a certain number of bits (e.g., 8, 16). The G.711 standard specifies 8 bits and 8KHz sampling, for a bandwidth required of 64Kbps (i.e., 8 times 8K). Mu-law and A-law encoding are actually logarithmic quantization algorithms rather than compression methods. However, a 12- to 16-bit to 8-bit reduction is normally part of the encoding. Mu-law encoding presumes that more signals lie near the low end than the high, so greater quantization is provided at low signal levels. A-law encoding is used outside the United States and Japan, and provides slightly lower signal-to-quantizing performance. However, Mu-law encoding has slightly higher noise levels, because an all-zero value is not used.

2.10.4 Issues/Discussion

The format differences between the US telephony standard (Mu-law) and the European telephony standard (A-law) complicate interchange between international users.

Encoded digital audio can be used to add music or voice-over annotation to any compound document.

Use of MPEG-encoded audio will avoid most portability problems.

2.10.5 Monitored and Related Standards, Technologies, Products and Practices

2.10.5.1 Streaming Audio

Streaming audio (RealPlayer, Streamworks) is currently implemented only in proprietary formats, but affords better performance by providing progressive output of sound at the user’s workstation, rather than requiring the entire file to be downloaded before playing begins. As with streaming video technologies, implementations in which the streaming data server selects the IP port addresses (i.e., as opposed to selection by the client) may conflict with firewall implementations that guard usage of IP ports by hosts outside the protected domain.

2.10.5.2 VOC

VOC (Creative VOiCe) format is an essentially obsolete proprietary format developed by Creative Labs for their SoundBlaster line of sound cards for PCs. Its use has been largely supplanted by WAV-format sound. VOC is limited to 8-bits and a single channel with silence detection.

2.11 MOTION VIDEO AND VIDEO/AUDIO MIX

Video is an audio/visual playback and recording technology used in TV. It also refers to computer screens and terminals. There are dozens of computer/video display standards. Full motion video refers to compressed video with picture quality that is generally acceptable to users although not of broadcast quality. Freeze frame refers to the capture of a single field or a full frame from a sequence in full motion. It is also used to describe transmission of a single frame of video depicting a still or frozen graphic image.

2.11.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
SMPTE 259M, ISO/IEC 11172-2 (MPEG-1), ISO/IEC 13818-2 (MPEG-2) digital video	GCCS, DII COE, JTA, AITS/ITSG, USIGS
QuickTime, AVI	GCCS, DII COE, JTA, AITS/ITSG, USIGS
EIA-608 closed captioning until replaced by digital standards	USIGS
Higher Level of Effort	
Indeo/DVI	JTA, AITS/ITSG, USIGS
525/60 (NTSC and RS-170 component) and 625/50 (PAL/SECAM) and S-Video analog video	USIGS
ITU-R BT601-4 Component (4:2:2; CCIR 601) for uncompressed digital video	USIGS
Monitored	
MPEG-4 for low bit rates, MPEG-7	JTA
HDTV and ATSC advanced video	JTA
Streaming video (e.g., VDO, VivoActive, RealPlayer, VXtreme, Streamworks, QuickTime, NetShow)	Netscape, Microsoft
Shockwave	Macromedia

PRACTICES	SOURCE
Preferred	
Compatibility with Kodak DCS and Sony DVBK-1000 digital cameras and commercial frame grabber boards (e.g., Sun VideoPix)	GCCS
Discouraged	
Looping without user intervention except background audio	DII COE, USIGS
VHS, U-Matic, D2 and D3 videotape for acquisition, archiving, or processing	

2.11.2 Overview

2.11.2.1 MPEG-1

Moving Picture Experts Group is the name of the ISO committee (JTC1/SC29/WG11) working on the standard for "Coding of Moving Pictures and Associated Audio for Digital Storage Media Up to About 1.5 Mbits/s." MPEG-1 is a formal standard defined by ISO/IEC 11172. MPEG-1 is an open international standard for video compression that has been optimized for CD-ROM data transfer rates. The standard defines a bit-stream representation for synchronized digital video and audio, compressed to fit into a bandwidth of 1.5 Mbit/sec. This corresponds to the data retrieval speed from CD-ROM and DAT, a major application of MPEG for the storage of audiovisual information on this media. MPEG is also gaining ground on the Internet as an interchange standard for video clips.

The compressed data contains three types of frames: I (intra) frames are coded as still images; P (predicted) frames are deltas from the most recent past I or P frame; and B (bi-directional) frames are interpolations between I and P frames. I frames are sent once every 10 or 12 frames. Reconstructing a B frame for display requires the preceding and following I and/or P frames, so these are sent out of time order.

Substantial computing power is required to encode MPEG data in real time. MPEG-2 is optimal for a variety of data rates ranging from three to 10 megabits per second and higher. Hardware or software implementations with hardware acceleration are several times faster than software-only implementations, which require higher baselines in computing platform power for minimally acceptable performance. For occasional viewing, software implementations are likely to be acceptable; for authoring, hardware implementations are usually required for practical throughput. For further information on these issues, see Section 6.3.

MPEG-1 compression techniques are geared to asymmetric applications where the decompression process is extremely faster than the compression process. Applications making heavy use of this technology include electronic publishing, video games, and delivery of movies. It is also used extensively for downloading and uploading video on the World Wide Web. The shell format is interoperable across platforms and considered to be platform independent.

Note that, because MPEG compression is lossy, successive edits and saves of MPEG streams suffer from the same type of iteratively increasing loss characteristics as JPEG images. Furthermore, because frame interpolation is one compression technique used by MPEG encoding, the image is not a true representation of events as they occurred. This latter characteristic may complicate MPEG use in situations where absolute image fidelity is required (e.g., medical, law enforcement and scientific applications).

2.11.2.2 MPEG-2

Generic Coding of Moving Pictures and Associated Audio is a formal standard defined by ISO 13818. MPEG-2 is required for generic compression coding of moving pictures and associated audio. MPEG-2 systems will form the infrastructure for delivery of audio and video in many industry sectors, including television distribution, visual telecommunications and a host of computer and multimedia applications.

MPEG-2 was initially intended for the recording and transmission of studio-quality motion video at bit rates in the order of 4 to 6 Mbps. The objective of MPEG-2 is to improve the quality of the image while maintaining the bit rate significantly below 10 Mbps. At the same time, more audio capabilities, such as multiple channels, have been included. The initial MPEG-2 standard has also been extended to support several HDTV formats. For information on the MP3 audio standard (Audio Layer 3 of the MPEG-2 standard), see Section 2.10.2.1.

2.11.2.3 Society of Motion Picture and Television Engineers (SMPTE) Specifications

These specifications are oriented toward the capture of high-quality analog and digital video, typically for audiovisual rather than computer-based multimedia applications, so complete discussion of their features is beyond the scope of this handbook. However, the specifications provided by the SMPTE afford the opportunity to bridge between analog video, uncompressed digital video and compressed digital video (e.g., MPEG-2).

2.11.2.4 Consultative Committee for International Radio (CCIR) and Other Specifications

These specifications include CCIR 601, which accommodates interlaced video in both NTSC (525-line, 60Hz vertical refresh) and PAL/SECAM (625-line, 50Hz vertical refresh) video streams. CCIR 601 uses 4:2:2 subsampling to encode NTSC and PAL/SECOM video, with the interlacing providing half the image information in each frame (e.g., 243 lines for encoded NTSC video).

EIA-608 provides for encoding of additional program information on scan lines not normally visible onscreen (e.g., Line 21).

2.11.3 Issues/Discussion

Implementations of SMPTE standards are not yet complete, especially in computer-based systems.

MPEG encoding is extremely computationally intensive and therefore processor-intensive (e.g., 90+% of a 200MHz Pentium), and is usually left to dedicated hardware devices, since off-the-shelf desktop workstations normally cannot handle the production workload.

MPEG implementations are typically fairly small frame sizes (e.g., 320x200) in order to gain acceptable performance.

Mixed audio and video, or video by itself, conforming to the guidance presented above should be used whenever motion picture applications include images that are photographed or filmed. Images that are generated directly by a computer should be conveyed as 3D or animated graphics.

Hardware-based encoding, using dedicated circuitry, of MPEG video and audio is platform-specific, but generally necessary for acceptable performance.

2.11.4 Monitored and Related Standards, Technologies, Products and Practices

2.11.4.1 Streaming Digital Video

Proprietary streaming systems that provide progressive implementations of video and audio data, to ensure that initial imagery is displayed without requiring download of the entire (usually large) file, are becoming increasingly popular. These include the VDO, VivoActive, RealPlayer, VXtreme, and Streamworks server and client authoring, delivery and display subsystems. In addition, the QuickTime format offers streaming implementations; Shockwave implementations generally do not use streaming, but require downloading the entire image file. Because each offers a freeware viewer, their popularity and widespread implementation has proceeded in the absence of defined standards. Unfortunately, this means that each browser must include viewers for each of these formats. As with streaming audio, care must be taken to ensure that the implementation is compatible with currently installed firewall implementations.

2.11.4.2 MPEG-4

Work is underway on MPEG-4 (ISO/IEC 14496 Working Draft) which is intended to be an extremely powerful standard for videoconferencing quality at extremely low bit rates. Its fields of application should be desktop videoconferencing and videophony. A typical context where MPEG-4 is used consists of a sender and a receiver. At the sender, audiovisual information related to an audiovisual scene is compressed, error protected if necessary, and multiplexed in one or more coded binary streams that are transmitted. At the receiver, these streams are demultiplexed, error corrected, decompressed, composited, and presented to the end user. The end user is given an opportunity to interact with the presentation. Interaction information can be processed locally, or transmitted to the sender. MPEG-4 Systems, the first part of the standard is expected to be completed in late 1998.

2.11.4.3 MPEG-7

MPEG has started a new work item called “Multimedia Content Description Interface” (MPEG-7). MPEG-7 will specify a standardized description of various types of multimedia information. This description shall be associated with the content itself, to allow fast and efficient searching for material that is of a user’s interest. This material includes still pictures, graphics, audio, moving video and information about how these elements are combined in a multimedia presentation (scenarios and composition information).

Potential applications of MPEG-7 include education, journalism (e.g. searching speeches of a certain politician), investigative services (human characteristics recognition, forensics), medical applications, architecture, and shopping (e.g. searching for certain types of clothes). MPEG-7 is intended to provide complementary functionality to the other MPEG standards. It will standardize multimedia content descriptions and represent information about the content, not the content itself. MPEG-7 is not expected to become an approved standard until late in 2000.

2.11.4.4 Analog Video

National Television Standards Committee (NTSC; RS-170A) video represents U.S. 525-line horizontal, 30Hz vertical interlace-refresh broadcast video. RS-170A has color (“chrominance”) information composited with brightness and shading information (“luminance”). Note that this compositing results in a loss of picture quality above that associated with the 525-line resolution. RS-170A is nominally a 4:3 aspect ratio image. “Standard” graphic image sizes (e.g., 320x200, 640x480) are based on historical implementation of NTSC monitors as computer monitors, with their inherent 4:3 aspect ratio.

PAL and SECAM standards represent European 625-line, 50Hz vertical refresh video.

NTSC, PAL and SECAM standard video and S-Video signals require higher levels of integration effort because their composited signals require additional processing for time synchronization (e.g., for overlay keying and matting), color separation for digitizing, and additional processing/display controls to offset their lower resolution and lack of scalability.

3. MULTIMEDIA AUTHORING

Assembly of information from multiple media sources into a single multimedia framework is termed multimedia authoring (also known as multimedia programming). This section provides guidance on selection of platforms, authoring languages and tools for multimedia authoring. In addition, guidance for selection of appropriate interchange media is provided. Note that multimedia authoring, although a part of videoconferencing and related collaborative applications, is considered in this section to be distinct from teleconferencing applications, since the user community is distinct in each case.

3.1 AUTHORING PLATFORMS

3.1.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
DII COE and JIEO 8300-compliant platforms (DoD Minimum Desktop Configuration)	GCCS, DII COE, JTA, AITS/ITSG, JIEO
JIEO 8303 and 8306-compliant platforms (Multimedia Extensions to DoD Minimum Desktop Configuration)	JIEO 8303/8306
HTML 3.2, CGI and frames capability for servers	GCCS, DII COE, JTA, AITS/ITSG
TLS – Secure Sockets Layer (SSL) 3.0	GCCS, DII COE, JTA, AITS/ITSG
Higher Level of Effort	
Secure Sockets Layer (SSL) implementations other than TLS	GCCS, DII COE, AITS/ITSG
MPC	JIEO 8303/8306
Monitored	
HDTV and ATSC advanced video	JTA
Streaming video (e.g., VDO, VivoActive, RealPlayer, VXtreme, Streamworks, Shockwave, QuickTime)	Netscape Communications

PRACTICES	SOURCE
Preferred	
Compatibility with Kodak DCS and Sony DVBK-1000 digital cameras and commercial frame grabber boards (e.g., Sun VideoPix)	GCCS
SMPTE-standard videotape for acquisition, archiving, or processing	USIGS
VHS for distribution videotapes	USIGS
Discouraged	
VHS, U-Matic, D2 and D3 videotape for acquisition, archiving, or processing	USIGS

3.1.2 Overview

Any multimedia-authoring platform for DoD must be capable of generating, encoding and compressing as required the components of the multimedia data stream. At a minimum, these platforms should ensure the proper generation of appropriate HTML code, should be capable of running a DII COE COErnel and Common Support Applications, and should conform to JIEO standards for a minimum desktop configuration for multimedia purposes, as may be amended to accommodate changes in technology.

DII COE-compliant platforms provide a more stable, better defined base of operating system, file structure, and applications to ease the task of integrating other applications and technologies. Such platforms are, therefore, assessed as requiring lower levels of effort for integration, compared to dedicated multimedia hardware, such as MPC platforms.

Since most viewing platforms are capable of at least HTML 3.2 browsing, with support for frames and CGI interactions with httpd daemons (i.e., web servers), HTML 3.2 should be considered as a least-common-denominator for production of multimedia data streams, until such time as HTML 4.0 comes into more popular usage and is fully supported by browser platforms in common usage. For secure implementations, the use of TLS - SSL 3.0 will ensure the required compatibility with other DoD implementations, whereas other SSL implementations may not.

Streaming video servers (which also stream audio data) afford generally better performance than applications which require the entire video file to be downloaded before viewing. Since streaming video permits the user to begin viewing/listening to the data stream once minimal startup information is passed, the user perceives better performance. This performance can be enhanced further through proper caching of the input data stream and matching the transmission packet size to available bandwidth. However, many streaming applications (e.g., VXtreme) are incompatible with most current firewall implementations, since the server rather than the client selects the IP port over which to transmit the data stream. The firewall sees this server selection process as an unapproved attempt to bypass security controls and disallows the access, interrupting the flow of the data stream. Intensive coordination and technology exploration is underway between firewall developers and streaming video product developers to ensure that such products will work in future firewall implementations.

Like streaming products, platforms supporting HDTV and other advanced video systems are not yet sufficiently mainstream to ensure universal accessibility to multimedia presentations authored with those technologies as their foundation.

Multimedia platforms should be compatible with appropriate video capture devices, for lower level-of-effort integration. However, use of VHS and similar lower-quality videotape media should be limited to endpoint distribution, since acquisition of video data directly onto these media will result in lower quality products when they are reproduced for distribution.

3.1.3 Issues/Discussion

Specific standards for multimedia platforms have not been successful (e.g., MPC); because in general, dedicated authoring stations are not cost-effective. Instead, authoring tools operate as applications on a general-purpose office-automation workstation.

Compatibility between authoring platform and presentation platform is ensured by adhering to mandated standards for multimedia data components (e.g., JPEG, GIF, and HTML).

Multimedia authoring platforms will be used primarily by training course and Web page developers. Other uses include preparation of publications, including those with embedded video and audio materials, preparation of simulation materials, and other uses requiring mixtures of input data formats.

In selecting multimedia authoring platforms, buyers should remain cognizant of Federal and Department of Defense requirements that all computing devices conform to so-called "Green" or "Energy-Star" specifications for power consumption, automatic shutdown, and disposal limitations. Frequently, higher-end or custom integrated platforms fail to comply with these regulations, mostly those limiting allowable power consumption.

3.2 AUTHORING TOOLS

3.2.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
Acrobat, Word 6, PowerPoint 4, Alchemy	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Netscape browser with HTTP and CGI links to servers via TCP/IP	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Netscape Composer, Asymmetric Toolbook II, Microsoft FrontPage, MacroMedia Director and Authorware, Word 97 SR-1, PowerPoint 97 SR-1, Folio, Apple QuickTime, Photoshop	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Higher Level of Effort	
MS Visual Basic, Office 95 or 97 w/o SR-1, Envoy	JTA, GCCS, DII COE
Monitored	
HDTV and ATSC advanced video	JTA
Streaming video (e.g., VDO, VivoActive, RealPlayer, VXtreme, Streamworks, streaming Shockwave and QuickTime, NetShow)	Netscape Communications
Java, JavaScript and ActiveX, HyTime, XML	GCCS, DII COE, AITS/ITSG, JTA

PRACTICES	SOURCE
Preferred	
Frames spawned within the current browser window (vice creation of an additional browser copy) and HTML 3.2(-) for clients or alerts for users with less capable browsers	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Perl scripts for CGI	GCCS, DII COE, JTA, AITS/ITSG, JIEO
DII COE guidance on page design and organization (e.g., HTML minimum contents, image size, viewers)	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Use of formats for which freeware or DII COE viewers are available	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Use of META tags to improve search engine "hits"	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Discouraged	
FRAME tags which spawn an external copy of the browser and browser-specific tag implementations	DII COE

3.2.2 Overview

Any DoD multimedia-authoring tool must be capable of generating, encoding and compressing as required the components of the multimedia data stream. At a minimum, a suite of tools should be available for encoding Microsoft Office, PDF and HTML multimedia documents.

Authoring tools which are part of the DII COE suite or are supported under the DII Enterprise License program will require lower levels of effort to integrate and use. Earlier editions of these products and others not supported DoD-wide will require higher levels of effort both to integrate and use, since training materials will not be generally available, users will not be likely to have prior experience with the product, and software errors corrected in more modern releases (e.g., Microsoft Office SR-1) will still remain.

For the reasons cited under the discussion of authoring platforms, use of advanced video (e.g., HDTV) and streaming video should be restricted to prototyping efforts only, until these technologies are in more widespread use.

Browsing within the existing window, including launching new frames, is preferred to keep memory usage and screen “real estate” consumption to a minimum. Similarly, following DII COE guidance on HTML page design and organization will ease integration of multimedia products with other DII COE applications, including minimizing required user training and user interactions needed to resize or reposition multimedia browser windows.

Use of Perl scripts to support CGI interactions with the httpd daemon (i.e., web server) will minimize training and maintenance burdens, since Perl is in common usage.

3.2.3 Issues/Discussion

A range of browser versions and types, including Netscape Navigator and Communicator, Microsoft Internet Explorer, and the National Center for Supercomputer Applications (NCSA) Mosaic browser are deployed throughout DoD. Each product and version may implement a different set of extensions to the basic HTML standard. This means that not all implementations of the HTML 3.2 standard will be viewed properly on all client platforms.

Multimedia authoring tools will be used primarily by training course and Web page developers.

Browser-specific implementations of HTML are to be avoided when constructing Web pages.

No multimedia data stream should be constructed for general consumption unless a freeware viewer is available for all data within that stream for all relevant client platforms. Failure to do so will add cost to the viewer end, complexity in integrating products, and limit the potential viewing audience to those owning the viewer product.

3.3 INTERCHANGE MEDIA

3.3.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
HTTP and CGI links to servers via TCP/IP CD-ROM and CD-R for non-network distribution Videotape, DAT	GCCS, DII COE, AITS/ITSG, JIEO
Higher Level of Effort	
Audio CD and videodisc	JTA, GCCS, DII COE
Monitored	
DVD, including DVD-R, DVD-ROM, DVD-RAM	JTA
Web casting, Mbone, IRC, streaming video (e.g., VDO, VivoActive, RealPlayer, VXtreme, Streamworks, Shockwave, QuickTime)	Netscape Communications
Java, JavaScript and ActiveX, HyTime, XML, SMSL	GCCS, DII COE

PRACTICES	SOURCE
Preferred	
Use of mechanisms and formats for which freeware or DII COE viewers are available	GCCS, DII COE, JTA, AITS/ITSG, JIEO
SMPTE-standard videotape for acquisition, archiving, or processing	USIGS
VHS for distribution videotapes	USIGS
Discouraged	
Presentations requiring multiple CD-ROMs	DII COE

3.3.2 Overview

Since multimedia distribution mechanisms are usually hardware-specific (n.b., including mechanisms for alleviating TCP/IP latency), these mechanisms are standardized only within vendor consortia.

Two distinct types of interchange media are to be used, depending on the application.

Where network distribution is required, some mechanism for ensuring minimum latency is needed, since the TCP/IP protocol is predicated on time-insensitivity in the data being transmitted. IP multicasting and IP switching techniques and UDP broadcasting, which eliminate some of the overhead associated with TCP/IP, are often used to alleviate network latency.

Wherever practicable, distribution to the using organization via physical media such as CD-ROMs can obviate the complexity associated with wide-area distribution of multimedia data streams.

When selecting physical media, the degree to which player hardware is distributed throughout the enterprise must be considered.

3.3.3 Monitored and Related Standards, Technologies, Products and Practices

3.3.3.1 DVD

The development of Digital Video Disc (DVD) technology and MPEG-2 digital video standards is making feature-length, high-quality digital video a reality. Sometimes also called Digital Versatile Disc, officially, DVD does not stand for anything; it is simply the name adopted for a family of related optical disc formats. The first DVDs will hold 4.7GB, seven times the capacity of a CD-ROM; later models are expected to hold 17GB, the equivalent of 30 CDs. Like CD-ROMs, DVDs record data by indentations on the disc surface, but the newer technology allows the indentations to be closer together and in two layers. The MPEG committee had compromised video quality in MPEG-1 in order to fit on the single-speed CD-ROM format, but the advent of DVD makes possible the development of broadcast-quality digital imaging. DVD standards include Dolby AC-3 surround-sound technology. Different formats used by different vendors, including several formats for creating DVD media, still complicate the selection of appropriate DVD-writer hardware and software.

4. MULTIMEDIA PRESENTATIONS

Multimedia presentations encompass delivery of text, graphics, audio and video. The need to synchronize different parts of the presentation requires the use in some applications of embedded time codes. The requirement for interoperability with other applications (including noninterference) and accurate representation of visual information requires the application of color definition standards. Ensuring that the delivery of visual information and sound to the user requires the selection of appropriate monitors and speakers. This section provides guidance on selection of standards for these applications.

4.1 PRESENTATION PLATFORMS

4.1.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort	
Use of Netscape browser for HTML	GCCS, DII COE, JTA, AITS/ITSG, USIGS
Use of DII COE Office Automation components	
SMPTE video time base	
Global Positioning System (GPS) time reference	
SVGA (VESA) and VGA-compatible (i.e., 1024 x 768 pixels of 64K colors each, minimum)	
15" diagonal minimum, 17" diagonal preferred, 19" or better ideal for desktop applications	
9" diagonal active-matrix color minimum for notebooks and laptop computers	
Stereo speakers with 20-20,000Hz frequency response	
Higher Level of Effort	
Audio CD and videodisc	JTA, GCCS, DII COE
Proprietary or inclusive platforms (e.g., MPC)	
NTSC, PAL and SMPTE studio monitors	
Projection systems with inadequate brightness	
13" diagonal or smaller monitors	
Mono audio	
Monitored	
DVD, including DVD-R, DVD-ROM, DVD-RAM	JTA
Streaming technology (e.g., VDO, VivoActive, RealPlayer, VXtreme, Streamworks, Shockwave, QuickTime)	Netscape Communications, GCCS
HDTV and ATSC	JTA

PRACTICES	SOURCE
Preferred	
Use of freeware viewers and plug-ins	GCCS, DII COE, JTA, AITS/ITSG, JIEO
Plug-ins over external viewers	

PRACTICES	**SOURCE**
Discouraged	
Reliance on internal speakers	DII COE, JIEO

4.1.2 Overview

Proprietary streaming systems that provide progressive implementations of video and audio data, to ensure that initial imagery is displayed without requiring download of the entire (usually large) file, are becoming increasingly popular. These include the VDO, VivoActive, RealPlayer, VXtreme, Streamworks, and Shockwave and QuickTime server and client authoring, delivery and display subsystems. Because each offers a freeware viewer, their popularity and widespread implementation has proceeded in the absence of defined standards. Unfortunately, this means that each browser must include viewers for each of these formats.

4.2 PRESENTATION TOOLS

The tools used for presentation of multimedia should conform to the previously defined guidance for components of that multimedia. This means that the tools used must be capable of creating, saving, managing, restoring, and editing the media to be distributed. If the media include audio or video to be linked with other media, as with a narrated presentation, for example, then the tool must be capable of creating and editing some method of time synchronization between the various media.

4.3 PRINTER DATA INTERCHANGE

Data to be printed must be formatted in such a way to convey appropriate information to the print device. With the advent of page (i.e., as opposed to line or character) printers, the need for including page description information (e.g., margins, portrait or landscape orientation) became important. Since most new printers acquired will be page printers, the mandated standards focus on those applications.

Note that TAFIM, JTA and DII COE I&RTS guidance is to obtain print services from the parent operating system, rather than generating print data from within the application. This means that applications should not generate native printer instructions (e.g., HP-PCL) directly, but should pass the data to be printed to an operating system printer driver for that conversion. Failure to do so means that, for example, if one printer is off-line, then the output cannot be mapped to another. Alternatively, using operating system services means that both color and black-and-white versions of a presentation can be created using two different printers but one application.

For this reason, these standards are designed to apply to selection of appropriate print devices, rather than software applications.

4.3.1 Guidance

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort HP-PCL or PostScript where PCL not adequate HPGL4 for plotters, etc.	GCCS, DII COE, AITS/ITSG
Higher Level of Effort EPS, PDF	GCCS

PRACTICES	SOURCE
Preferred Access to printer through OS only 8.5x11 or ISO A4 paper size	DII COE
Discouraged Direct generation of printer interchange data by application	JTA, DII COE, Proprietary Implementations

4.3.2 Overview

4.3.2.1 PostScript, PDF and Encapsulated PostScript (EPS)

PostScript and PDF were described earlier. In its use as a printer data transfer format, PostScript's chief advantages are output scaleability, the ability to redirect output to a different printer without regenerating the entire print stream, and the ability to view the image to be printed using software independently from the application or operating environment in which the print stream was generated.

EPS provides a means of translating raster and vector graphics images into pseudo-PostScript language for printing purposes. It should not be chosen for graphics use other than printing, since there are more robust, faster, smaller constructs available.

4.3.2.2 HP-PCL and HPGL4

Hewlett Packard Printer Control Language (HP-PCL or PCL) is a *de facto* standard developed by, and under the control of, Hewlett Packard for their laser page printers. It has been adapted for the full range of page printers, including inkjet, dye-transfer, and other printer engines from a broad range of manufacturers and software developers.

HP-PCL is token-based, rather than being fully descriptive. This limits its application to all circumstances. For example, some PDF constructs can only be printed via PostScript, since no token exists in PCL to represent that construct. However, the smaller sizes of PCL print images yield lower overall storage and bandwidth requirements.

HPGL4 (HP Graphics Language v4) is a proprietary but widely adopted standard instruction set for color plotters and related output devices.

4.3.3 Issues/Discussion

Attempts to interest printer manufacturers and software developers in supporting an open, non-proprietary standard have failed.

A key shortcoming of PCL is that should the print stream fail to be printed, the entire stream must be regenerated using a driver specific to any alternative printer chosen as a substitute.

PostScript is fully descriptive, rather than being token-based like HP-PCL. However, this means that more storage space and higher bandwidth are required to accommodate PostScript print images.

5. VIDEO AND AUDIOGRAPHIC TELECONFERENCING

The form of multimedia presentation in most widespread use is teleconferencing. There are several types of teleconferences and multiple ways to conduct those conferences. A conference could be a document conference, an audio conference, a videoconference, or a combination of the three. It could be an interactive point-to-point conference, an interactive multipoint conference, or simply a broadcast multipoint conference. A variety of collaborative capabilities may be available such as:

(1) File Sharing/Transfer
(2) Graphic Document Sharing/Annotation
(3) Shared Whiteboard
(4) Application Sharing
(5) Video Mail
(6) Text E-mail
(7) Slide Presentation
(8) Computer Chat

A document conference could be the exchange of text and graphics, but not real-time audio and video. A chat window could be used to communicate among the parties in a textual manner. Sometimes an out-of-band audio connection may be made (i.e., telephone) to help with communications.

An audio conference could be a conference call. In this context, however, it refers to the use of a computer system for the exchange of real-time audio information so that the participants can talk to each other. Also, some audioconferencing systems include the ability to have a simultaneous document conference, using the audio to facilitate the document conference. This may be the preferred type of conference when there is limited network bandwidth for real-time video or the participants do not need to see each other. (Research shows that the addition of facial expressions and body language improves the effectiveness of communication at least 50% over audio alone.)

A videoconference also contains audio and, unless it is solely an oral discussion, normally requires document collaboration capabilities as well (e.g., transmission of presentation slides).

In designing and selecting videoconferencing and audio teleconferencing systems and employing each for the conduct of meetings, the relative utility of each must be considered.

5.1 GUIDANCE

STANDARDS, TECHNOLOGIES AND PRODUCTS	SOURCE
Lower Level of Effort VTC001 Profile with H.320 or H.324 for low bit rate ITU-T G.700 series or MPEG-2 audio T.120 compatible	GCCS, JTA, AITS/ITSG
Higher Level of Effort EPS, PDF, IRC, CU-SeeMe, proprietary solutions	JTA
Monitored H.321, H.323, H.310, H.322	JTA, AITS/ITSG

PRACTICES	SOURCE
Preferred VTC001 and H.324 for room-sized applications VTC001 and H.320 for room-sized applications H.323 for desktop applications once firewall problems are solved Automatic panning to and zooming on person speaking for connections below 384Kbps	JTA
Discouraged Low bit rate connections for room-sized applications	JTA

5.2 OVERVIEW

A videoconferencing system needs certain hardware components in addition to a computer to provide the encoding and decoding of the video and audio signals.

1. Camera — Provides the video input of local user and is usually mounted in a stationary position on the monitor.

2. Monitor — The video is displayed on desktop computer's monitor or on a large screen NTSC or PAL monitor (a converter box may be needed).

3. Microphone — The microphone may come as part of the camera, as a separate microphone, or as part of a headset. It is used for picking up local user speech. The manufacturer's choice of microphone, noise suppression and echo cancellation techniques used is what makes the system sound good or bad.

4. Speaker — The speakers may be separate or in a headset. A headset may be preferred to minimize feedback into microphone. Multimedia speakers, for outputting the audio, are different from stereo speakers. They are magnetically shielded to not interfere with the monitor and are self-amplified (no external power source needed), producing higher quality sound. They are able to operate in the 20Hz to 20KHz range.

5. Video Board — The video board compresses the video input from the camera and feeds it to the network interface. It also accepts incoming video, decompresses it and sends it to the monitor. Video boards are not all the same. The number of video inputs may vary. If playing videotape through conferencing system is important to your application, you need to look at the time base correction capabilities of the video board. Tape decks have inherent frame-to-frame time-sync instability that modern monitors easily correct. Some codecs are not as forgiving.

6. Network Interface — The network connection can be may through a high-speed modem, or LAN adapter.

5.2.1 Conferencing and Network Capabilities

1. POTS - Plain Old Telephone System modems 14.4/28.8 KBPS

   The initial transfer of a detailed photograph (7.5 MB) using JPEG 10:1 compression would take about 7 minutes at 14.4 KBPS and 3.5 minutes at 28.8 KBPS.

2. ISDN - Integrated Services Digital Network (64/128Kbps)

   Bandwidths include Basic Rate Interface (BRI) at 128 KBPS and Primary Rate (PRI) at 1.544 Mbps. While ISDN can support isochronous data transmission (the regular timed delivery of data), it is not sufficient to provide TV quality video. ISDN has 5-10x performance improvement over POTS. (The 7.5 MB file for the detailed photograph can be sent over the network using JPEG 10:1 compression in 47 seconds.)

3. 10 Mb/s Ethernet

   The 10 Mb/s Ethernet could support up to 4 users for education and entertainment-type applications. Ten to 15 users could be supported for video conferencing-type application. (The 7.5 MB file for the detailed photograph can be sent over the network using JPEG 10:1 compression in .6 seconds, assuming no collisions occur and the full bus bandwidth can be used.)

4. 100 Mb/s Ethernet or FDDI

   The 100 Mb/s Ethernet or FDDI network could support 10 to 40 users for education and entertainment-type applications. One hundred to 150 users could be supported for video conferencing-type application. (The 7.5 MB file for the detailed photograph can be sent over the network using JPEG 10:1 compression in .06 seconds, assuming no collisions occur and the full bus bandwidth can be used.)

5. 2 Gbps ATM Hub

   A 2 Gbps Asynchronous Transfer Mode (ATM) hub could support 333 MPEG-2 users or 1,333 MPEG-1 users for education and entertainment-type applications. Up to 5,200 users could be supported for H.261 video conferencing-type application. (The 7.5 MB file for the detailed photograph can be sent over the network using JPEG 10:1 compression in .003 seconds, assuming no collisions occur and the full bus bandwidth can be used. At these rates there may be other bottlenecks that slow down the file transfer. Will a desktop system have a 2 Gbps interface? If it did, can the system bus handle it?)

Depending on location, POTS only gives 19.2 KBPS from a 28.8 KBPS connection. This provides only 15 frames per second (fps) video, which would then have a strobe effect. The latency in the phone system is 125 msec. People find it difficult to communicate if audio latency is greater than 200 msec and inconsistent latency tends to introduce audible pops and clicks in the audio stream. Typical video latency is in 200-500 msec range. Asynchronous LANs do not support video conferencing well. The video needs synchronous type of networks for best performance. Isochronous networks for video transmission is needed, i.e., regular timed delivery of data.

5.2.2 Protocols for Multimedia Conferencing

5.2.2.1 T.120 Series

T.120 is a document/data conferencing standard developed by the ITU-T. It describes ways of sharing and annotating data (not containing moving video images) between desktop computers, video conferencing, and audiographic systems. It is not just one standard, but a suite of international standards that describe specifications for distributing data during a multipoint, multimedia conference supported by a variety of transmission media. T.120 defines protocols and APIs for cross-platform whiteboarding, pointing, annotation, binary file transfer, and even application sharing, all in a multipoint environment using heterogeneous networks. It will operate with H.320/H.261-compatible desktop systems.

T.120 follows the 7-layer OSI model. The lower layers specify an application-independent mechanism for providing multipoint data communication services (T-121, t.122, T.123, T.124, T.125). If an application supports only the lower layer, the same product is needed for each conference member. The upper layers define protocols for specific conferencing applications, such as shared whiteboarding and multipoint file transfer (T.126 and T.127). Applications that also support the upper layers should be able to conference with other products supporting the similar features.

T.121 is the Generic Application Template (GAT). It defines a generic Application Resource Manager (ARM) which will manage GCC (T.124) and MCS (T.122/T.125) resources on behalf of the application protocol-specific functionality defined as an Application Service Element (ASE). Services it provides include enrolling an application in GCC and attaching it to MCS domains. Resources managed on behalf of the application include tokens, channels, and capabilities. It also responds to GCC indications and can invoke peer applications on other nodes in the conference.

T.122 and T.125 form the Multipoint Communications Services (MCS). T.122, the Multipoint Communication Service for Audiographic and Audiovisual Conferencing, defines the multipoint services available to the developer. T.125, Multipoint Communication Service Protocol Specifications, specifies the data transmission protocol. Together, these specifications provide connection-oriented services and ensure that data received from multiple source nodes is received in the same sequence. The range of conferencing options supported includes one-to-one, one-to-many, and many-to-many.

T.123 is the Protocol Stack for Audiographics and Audiovisual Teleconference Applications specification. It provides reliable point-to-point sequenced data delivery connections for:

- Public-switched telephone network (PSTN) using standards V-series modems at the physical layer;
- Packet-switched digital network (PSDN) for X.25 protocol.
- ISDN and Circuit-Switched Digital Networks (CSDN) where audio, video, and T.120 data are multiplexed according to H.221.

T.124 is the Generic Conference Control (GCC) for Audiographic and Audiovisual Terminals specification. It provides conferencing capabilities, outlining the services necessary to set up and manage a multipoint meeting. It establishes and terminates conferences, negotiate capabilities between applications, and provide conference administration (who is conducting the conference and how to pass the “baton” and control the camera). It also provides security, such as passcode protection, limiting those who can enter the conference.

T.126 is the Still-Image Protocol Specification. This is a higher layer application protocol, which provides the ability in a multipoint document conference to view and annotate images, share applications, and exchange fax images.

T.127 is the Multipoint Binary File Transfer Protocol Specification. This is a higher layer application protocol, which provides the ability to initiate simultaneous multipoint file transfers. Transfers can be to all participants or to a selected subset.

T.128 is the Audiovisual Control for Multipoint Multimedia. It provides a framework for control and management of interactive audio and visual services within a multipoint multimedia communication environment. It also provides a toolkit function for management, routing, identification and processing of audio and visual streams, together with remote device control and source selection. This is going to be replaced by T.130 - T133 to specify the way T.120 data channels are used for control of real-time media streams.

Areas not covered in the T.120 standard that vendors may implement include directory services, security services, format of how data is presented on the user’s screen, and how calls are set up and torn down.

There are a number of industry standards for video conferencing. The difference is usually the transmission media being used for the conference. The T.120 data/document conferencing standards are incorporated into the video conferencing standard suites.

5.2.2.2 H.320 - Narrow-band Video Conferencing System

H.320 is the Narrow-band Visual Telephone System and Terminal Equipment specification developed by the ITU-T. It describes multimedia conferencing over the Integrated Services Digital Network (ISDN). It is a suite of standards for video, audio, and document conferencing applications that specifies signaling, coding, channelization, compression, transport, and display parameters. It is configured to support video conferencing over ISDN, calling for the constant bit rate, isochronous transmission and dedicated bandwidth. Conferences can be held over two switched 56 KBPS or fractional T-1 lines.

(1) Video Specifications

H.261 is the Video CODEC for Audiovisual Services at Px64Kbps specification. It defines the syntax and semantics of video bitstream for data rates of Px64Kbps, where P = 1 to 30. For videophone applications, which are less demanding of image quality, data rates of P = 1 to 2 (64/128 KBPS) are used. Video conferencing requires a higher quality (P>= 6, or >=384 KBPS). H.261 deals with coding, compression, and decoding of video streams over digital circuits composed of one or more 64 KBPS channels (ISDN). It defines two picture resolutions: Quarter Common Intermediate Format (QCIF) and Common

Intermediate Format (CIF). QCIF (176 pixels/line x 144 lines) is used mainly for desktop and videophone applications. It provides lower quality than CIF and is used in applications where P<3. CIF is used for room system applications and is lower quality than broadcast TV (352 pixels/line x 288 lines, which is about 1/2 NTSC resolution). It optimizes bandwidth by trading off picture quality against motion, i.e., a quickly changing picture will have lower quality than a relatively static picture.

(2) Audio Specification

G.711 is the Pulse Code Modulation (PCM) of Voice Frequencies (Audio Compression) specification. It defines the audio compression onto 64 KBPS channel using only 3 kHz of the channel.

G.722 is the 7 kHz Audio Coding Within 64 KBPS (Audio Compression) specification. It defines the audio (50 Hz to 7 kHz) compression onto 64 KBPS channel and requires 48 KBPS of the channel.

G.728 is the Coding of Speech At 16 KBPS Using Low-Delay Code Excited Linear Prediction (Audio Compression) specification. It defines the audio (50 Hz to 3.4 kHz) compression onto a 16 KBPS channel using only 3 kHz of the channel.

(3) Data Specification

T.120 as described above.

(4) Multiplexing Specification

H.221 is the Framing Structure for 64 to 1920 KBPS Channel in Audiovisual Teleservices specification. It defines the end-to-end synchronization scheme for voice and video and specifies what information is in a bitstream, so each codec can keep track of video frames. It also describes the order in which the bits are multiplexed before being transmitted and how to label the bits of transmitted information (audio, visual, or control).

(5) Signaling Specification

H.320 is the Frame-synchronous Control and Indication Signals for Audiovisual Systems specification. This specification describes how commands between codecs are exchanged during a video conferencing session. It also sets up a table to determine if certain signals are mandatory or optional in the transmission. Certain diagnostic and indication messages are required to be sent to the user.

H.242 is the System for Establishing Communications Between Audiovisual Terminals Using Digital Channels up to 2 Mbits/s specification. It specifies the call setup, disconnect, and transfer.

(6) Multipoint Specification

H.243 is the System for Establishing Communications Between Three or More Audiovisual Terminals Using Digital Channels up to 2 Mbits/s specification. It defines communications between the multipoint control unit and the codec.

H.331 is the Half Duplex Multiplexing specification. It defines for the codec how to string together data for transmission over half-duplex channels, such as satellites.

(7) Encryption Specifications

H.233 is the Confidentiality System for Audiovisual Services specification. It defines the method for identifying and negotiating encrypted data.

5.2.2.3 H.324 - Low Bit-Rate Video Conferencing System

H.324 is the "Visual Telephone Terminals over GSTN" specification from the ITU. This is a suite of standards for video conferencing on analog phone lines (POTS) or cellular links. Video would have 20 KBPS bandwidth and the voice would have 8 KBPS bandwidth.

(1) Video Specification

H.263 is the Video Coding For Low Bitrate Communications specification. It is a video codec algorithm that includes H.261 compatibility. It contains four negotiable options to improve performance and supports five video resolutions (QCIF (in H.261), CIF (in H.261), SQCIF (1/2x resolution of QCIF), 4CIF (4x resolution of CIF), 16CIF (16x resolution of CIF)).

(2) Audio Specification

G.723 is the Dual Rate Speech Coder for Multimedia Telecommunications Transmitting at 6.4 and 5.3 KBPS specification. It is the audio compression algorithm.

AV.25Y (Audio codec algorithm)(no further information)

(3) Data Specification - T.120

(4) Multiplexing Specification

H.223 is the Multiplexing Protocol for Low Bitrate Multimedia Communications specification.

(5) Signaling Specification

H.245 is the Control Protocol for Multimedia Communications specification. It provides end-to-end signaling for proper operation. It multiplexes video, audio, data, and control streams into a single bit stream for transmission, and demultiplexes a received bit stream into the various multimedia streams. It also performs logical framing, sequence numbering, error detection, and error correction by means of retransmission, as appropriate to each media type.

(6) Multipoint Specification - (none specified)

(7) Encryption Specification - H.333

5.2.2.4 H.323 - LAN-based Video Conferencing System

H.323 is the Visual Telephone Over Non-Guaranteed Quality of Service LANs specification. It is a suite of standards detailing the modes of operation on LAN video conferencing.

(1) Video Specification - H.263

(2) Audio Specification - G.711, G.723, G.722, G.728

(3) Data Specification - T.120

(4) Multiplexing Specification - H.22z (?)

(5) Signaling Specification - H.230, H.245

(6) Multipoint specification - none

(7) Encryption Specification - (not known)

5.2.2.5 H.321 - Video Conferencing over ATM

H.321 is the Visual Telephone over ATM specification. It is the same as H.320, but on a different transmission media.

(1) Video Specification - H.261

(2) Audio Specification - G.711, G.722, G.728

(3) Data Specification - T.120

(4) Multiplexing Specification - H.221

(5) Signaling Specification - H.230, H.242

(6) Multipoint specification - H.243

(7) Encryption Specification - H.233

5.2.2.6 H.322 - Video Conferencing over ISO-Ethernet

H.322 is the Visual Telephone over Guaranteed Quality of Service LANs specification. It is the same as H.320, but on a different transmission medium.

(1) Video Specification - H.261

(2) Audio Specification - G.711, G.722, G.728

(3) Data Specification - T.120

(4) Multiplexing Specification - H.221

(5) Signaling Specification - H.230, H.242

(6) Multipoint specification - H.243

(7) Encryption Specification - H.233

5.3 CONFERENCING ISSUES/DISCUSSION

5.3.1 Audio Teleconferencing Considerations

The utility of audio teleconferencing is that participation in meetings is possible without requiring that all participants be physically in the same room. Generally, this requires reasonable audio fidelity and the capability to pause transmission (known as muting) at the discretion of participants for sidebar discussions.

Reasonable audio fidelity is provided by good signal-to-noise ratios, muting and filtering to control feedback, and measures (i.e., meeting protocols or full duplex transmission) to preclude transmission by one party from overriding transmission by another.

The advantages of audio teleconferencing are that multipoint participation is generally easier to achieve than with videoconferencing. With videoconferencing, a requirement to display simultaneously the image of all participants would be difficult to achieve and would be limited by the amount of monitor screen space and resolution available. Audio teleconferences are easier to arrange, as well, normally requiring distribution of only a single call-in telephone number and access to speakerphones for any rooms with multiple participants. The latter are becoming increasingly common even in small offices, owing to the need for hands-free telephone conversations as a routine business transaction and the need to avoid repetitive-motion injuries associated with cradling a handset between head and shoulder.

Major drawbacks of audio teleconferencing are three: information presented visually (e.g., viewgraphs and whiteboards) must be communicated through separate means, remote participants cannot be identified when speaking unless all participants recognize each other's voices, and "body language" of remote location participants cannot be seen.

5.3.2 Video Teleconferencing Considerations

ITU H.320 is an umbrella for a series of audio and video standards ensuring that videoconferencing systems can talk to one another over digital phone lines, primarily ISDN lines, but also the fractional T-1 and Switched-56 lines used by some companies, universities, and government agencies. Essentially, H.320 means that your videoconferencing system can link to any other system connected to a digital phone network, regardless of brand or type.

Room-type conferencing systems have followed H.320 specs for years. Desktop and point-to-point videoconferencing systems have only recently embraced H.320 after some attempts to push proprietary standards or an alternative standard for PC-based conferencing. Today, H.320 compliance has become a given for most products and a must-have feature for many buyers.

Recently, the ITU has added several spin-offs to the H.320 standard. Notably, a new standard called H.324 extends the basic audio and video compression and transmission specifications of H.320 to analog or POTS phone lines.

Though H.324 can't boost regular phone lines' modest bandwidth, its efficient, scaleable video-compression algorithms may nonetheless improve video frame rates. Its primary contribution, however, will be to fill the standards vacuum that currently limits systems using POTS products such as Creative Labs' ShareVision to talk only with identical systems.

Two other pending H.320 extensions should help standardize videoconferencing over LANs. H.322 is basically an enhanced version of H.320 for networks designed to handle time-dependent (isochronous) data traffic such as real-time video.

This standard will be employed in so-called isoEthernet LANs. Another version, dubbed H.323, will work with traditional packet-switching LANs, including Ethernet and Token Ring networks.

An H.321 standard is also pending. It involves next-generation networks using asynchronous transfer mode (ATM) and broadband ISDN technologies. In theory, at least, all these H.32x standards will bridge the gaps between current videoconferencing products and delivery media.

The increased utility of videoconferencing over audio teleconferencing is that participants can discern who is speaking at any given instant as well as observing the “body language” of other participants. Also, information presented visually does not require other means of transmission.

However, achieving this improvement requires four capabilities: a frame rate at or above ten frames per second, sufficient color and spatial resolution to distinguish facial features, adequate room lighting and a sufficiently large and bright screen image so that all remote location participants can be identified, and the same audio capabilities as audio teleconferencing.

Unfortunately, the full range of these capabilities is lacking in many COTS videoconferencing systems. For example, projection systems lack the degree of brightness seen in cathode ray or active-matrix LCD monitors. This limits the ability of projection systems to serve as effective videoconferencing components, since they generally do not furnish sufficient screen brightness to offset the ambient room lighting necessary to ensure that all participants can be seen. The specific lighting requirements depend on the low-light capabilities of the video cameras used, so specific rules of thumb are impractical. Additionally, data rates below 384Kbps are usually inadequate to provide an appropriate combination of frame rate, spatial resolution, image size, and color fidelity to allow participants to identify speakers and discern “body language”.

- At least ten frames per second (10fps) and Full Common Intermediate Format (FCIF) service is needed to see facial features, to distinguish who is speaking, expressions of delight or dismay, and other body language. FCIF is defined in H.261 as 352x288 pixels of luminance, and 176x144 pixels of chrominance. Note that both H.320 and VTC001 permit QCIF, which is no better than one-quarter of this already limited resolution.

- A Class 3 H.320/VTC001 system will yield 15-30fps at FCIF at 384Kbps, presuming both endpoints are Class 3 systems. Note that if a Class 3 is connected to a Class 1 or a Class 2 quality system, then the lower capability prevails (e.g., QCIF at 7.5fps or videophone quality – suitable for seeing a face, but not for distinguishing between pleasure or pain).

- A Class 2 pair could theoretically deliver 15fps at FCIF at 128Kbps (i.e., an ISDN pair), normally adequate for body language/facial expression/participant identification, were it not for the fact that too many implementations will allocate one of the ISDN channels completely to the audio portion, leaving only 64Kbps for the video. At that bandwidth, the video reverts to 7.5fps, with the attendant quality losses.

This is the rationale for establishing 384Kbps as an "enterprise-wide" lower limit of practicality for communications circuits supporting videoconferencing.

These factors, coupled with the additional setup time often required of videoconferences and their generally lower reliability than audio teleconferences, mean that videoconferencing should not used when its marginal value does not sufficiently exceed that of audio teleconferencing. The advantages of videoconferencing can prove to be expensive to realize when compared to audio teleconferencing.

5.4 MONITORED AND RELATED STANDARDS, TECHNOLOGIES, PRODUCTS AND PRACTICES

5.4.1 DoD Joint Technical Architecture

The original Joint Technical Architecture (v1.0) mandated the following standards for video:

MPEG-1 provides for a wide range of video resolutions and data rates but is optimized for single and double-speed CD-ROM data rates (1.2 and 2.4 Mbits/s). With 30 frames per second video at a display resolution of 352 x 240 pixels, the quality of compressed and decompressed video at this data rate is often described as similar to VHS recording. MPEG-1 is frequently used in applications with limited bandwidth, such as CD-ROM playback or Integrated Services Digital Network (ISDN) videoconferencing. The following standards are mandated:

- ISO/IEC 11172-1: 1993 Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s -- Part 1: Systems

- ISO/IEC 11172-1: 1993/Cor. 1:1995 Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s -- Part 1: Systems Technical Corrigendum 1

- ISO/IEC 11172-2: 1993 Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s -- Part 2 Video.

MPEG-2 is designed for the encoding, compression, and storage of studio-quality motion video and multiple CD-quality audio channels at bit rates of 4 to 6 Mbits/s. MPEG-2 has also been extended to cover HDTV. The following standards are mandated:

- ISO 13818-1: 1996 - Generic Coding of Moving Pictures and Associated Audio Information - Part 1: Systems
- ISO 13818-2: 1996 - Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video.

Video Teleconferencing (VTC) Standards

VTC terminals operating at data rates of 56-1920 KBPS shall comply with the Industry Profile for Video Teleconferencing, VTC001. The purpose of the profile is to provide interoperability between VTC terminal equipment, both in point-to-point and multipoint configurations. This profile is based on the ITU-T H.320 and T.120 series of recommendations. VTC terminals operating at low bit rates (9.6-28.8 KBPS) shall comply with ITU-T H.324. The following standards are mandated:

- VTC001, Industry Profile for Video Teleconferencing, Revision 1, April 25, 1995
- ITU-T H.324, Terminal for Low Bit Rate Multimedia Communications, March 19, 1996.

5.4.2 Related ITU-T Standards

H.310 is an alternative specification to H.321, for broadband ATM. Products built to this standard are not yet widely marketed.

6. OTHER ISSUES AND CONSIDERATIONS

6.1 ARCHITECTURAL IMPLICATIONS OF MULTIMEDIA PROCESSING AND COMMUNICATIONS

A broad range of applications arise from the flexibility multimedia affords through its capability for mixing media types. Training, presentations, reference materials, video publishing, multimedia database applications, mapping/charting/geodesy and imagery (MCG&I) integration, and real-time or asynchronous workgroup collaboration are facilitated through multimedia.

The nature of multimedia data and the dynamics of multimedia interactions mean that network, processing and storage requirements are difficult to predict. User actions are not necessarily sequential. The timing (i.e., pacing) of a multimedia presentation (and hence its bandwidth demand) is not necessarily proportional to file size.

6.1.1 User Interaction Protocols

The demands placed on processing and communication systems by multimedia applications are directly related to the general protocol used to convey the multimedia information. These protocols may be grouped into two general categories, interactive and message-oriented.

Interactive multimedia generates demand for processing and communications resources, because during interactive sessions, users can alter the type, timing and sequence of information or communication requirements. This is in contrast to non-interactive applications, where the user cannot deviate from a predetermined path. Interactive multimedia presentations are analogous to (and may actually be used to support) conferences and meetings.

Message-oriented, non-interactive multimedia generates a different kind of demand for processing and communications resources. In this protocol, users pass shared workspace and information to the next worker in a coordinated workgroup flow of multimedia information. This is in contrast to an interactive protocol, where workers converge on the same, shared workspace at the same time. Message-oriented protocols are analogous to inter- or intraoffice document coordination.

6.1.2 Performance Implications

Media data streams and multimedia protocols require different levels of network and computing performance.

With interactive multimedia, delays adversely affect quality and usefulness of the information conveyed. This is because the degree of interactivity decreases as delays increase.

With message-oriented multimedia, the opportunity to avoid real-time demands lessens bandwidth requirements. However, use of hyperlinks in compound documents to external sources still require robust communications to support retrieval of that information, especially by a large number of users. Judicious application of portable storage systems (e.g., CD-ROM, DAT, DVD) can provide document sharing without imposing network loads, since the transport of information from one user to another is physical rather than electronic.

In general, combined video and audio data streams require more processing and communications bandwidth than video by itself. Video requires more bandwidth, in general, than still images, rendering of 3-D graphics or audio, any of which require generally more bandwidth than 2-D graphics. All of these require far more bandwidth than that needed for conveying text alone.

Generally, higher bandwidth multimedia data streams are more tolerant of transmission quality flaws. Text, on the other hand, is relatively intolerant of quality flaws. Video and imagery are relatively more tolerant than other streams, though this can be application dependent.

6.2 SECURITY

Security must be included in any consideration of multimedia authoring, distribution, and display. Often, commercial applications do not provide any security whatsoever. There are documented instances of unauthorized interception of Government videoconferences. In addition, mission applications may have special requirements, such as maintenance of patient privacy in telemedicine applications. Protection to a level of C2 assurance may or may not be adequate.

6.2.1 Information Integrity

High bandwidth demands and large file sizes complicate data integrity measures. Appropriate means must be provided to ensure that the required data integrity is maintained from the authoring platform to the display platform.

6.2.2 Confidentiality Protection

Although encryption processing can throttle delivery throughput, some mechanism for confidentiality protection is required for those applications requiring it. Note that the “password” protection offered by some compression methods (e.g., ZIP), is not considered adequate protection, nor is the compression itself.

6.2.3 Non-repudiation of Origin and Receipt

Signature facsimiles (i.e., a raster or vector facsimile of an individual's signature) can be transmitted easily as part of a multimedia data stream. Note that an accompanying, robust mechanism is additionally required for verifying that the individual to whom the signature supposedly belongs actually applied the signature. A public-key cryptographic digital signature, such as that provided by the RSA, PGP or DSS mechanisms, is appropriate for this purpose. Use of digital signatures, however, presumes that the public portion of the key comes from an integrity-protected source and that there is a robust mechanism for associating the individual with his or her digital signature (e.g., a token-based mechanism such as Fortezza).

6.2.4 System Availability

Multimedia data streams are relatively fragile, so for mission-critical applications, mechanisms must be included to ensure system availability and preclude denial-of-service attacks. Additionally, multimedia desktop platforms can be unreliable, owing to the complexity of their integration. Acquiring burned-in, proven, integrated solutions from single or a limited number of vendors can alleviate this latter difficulty. Also, confirming that any hardware or software used has passed operating system vendor compatibility checks (e.g., Microsoft's Hardware Compatibility List) offers additional assurances.

6.2.4.1 Audit

Audit capabilities are required to ensure that the required security functionality is being delivered as a matter of course. This includes assurance both that the security measures are providing adequate protection, and that they have been implemented as designed. Audit is a key issue with applet technologies (Java, JavaScript, and ActiveX), especially where no capability is provided to generate or write audit results to a protected log.

6.3 APPLICATIONS AND EMERGING TECHNOLOGIES

6.3.1 Multimedia Support by Defense Information Infrastructure Common Operating Environment Infrastructure and Common Support Applications – "Multimedia for the Warfighter"

By most accounts, the war of the future will focus on technology and information. The warrior who controls the information is usually the winner. Multimedia will be used in everything from interactive training programs to digitizing information on the battlefield.

The Coast Guard has developed several training courses for its helicopter and aircraft technicians using interactive lessons on CD-ROM. Students can run through procedures on test sets, antennas, control heads, and other instruments throughout the aircraft without actually crawling through the entire aircraft.

The Defense Advanced Research Projects Agency (DARPA) has initiated a project for research and development of visualization technologies for warfighters that will draw heavily on multimedia technology. The objective of the Warfighter Visualization program is to revolutionize and accelerate the availability of technology and systems that allow warfighters to see, feel, hear, and interact with time-critical information while operating in real-world military operations. One of the goals of this DARPA project is to develop enabling technologies for combining an individual warfighter's real view of the world with geospatially registered computer-generated information using graphics, virtual icons and/or location-based text information that is based upon the user's current activities or mission time.

DoD has established a basic technical architecture to which all present and future simulation projects must conform. The new DoD High Level Architecture lays out the basic rules of behavior simulation that programs should follow and defines the interfaces through which these programs, built independently, can interact with each other. DoD wants to be able to simulate a joint exercise by tying together a program developed by the Navy with another developed by the Marines.

6.3.2 Portable Information Carrier (PIC)

The following information is from MIL-HDBK-0501, DoD Portable Information Carrier Handbook, 15 April 1997. PIC is a portable device that contains one or more technologies used to store information related to an individual and may be used in one or more applications. The PIC may contain both updatable and static technologies such as an integrated circuit chip, magnetic stripe, bar code, digitized photo, and other printed information. The major focus of the PIC is the International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) 7816 (all parts) Integrated Circuit Card (ICC).

A PIC is intended for information exchange with outside application(s). It is critical to realize the PIC is not intended to be exclusively a data storage device. The card delivers information and may modify its content based on the information exchange and the applications using the data. This data exchange is accomplished through the PIC Integration Interface (PI^2.) The key difference between the PIC and other types of information carrying cards is the ability to read from, add to, and completely erase the ICC data on the card.

The PIC is to be a multiservice, multiapplication card. Each application and each service may have an exclusive section of the ICC on which to implement service or mission specific applications, but the overall card is meant to be used by all DoD services and components. Due to the limited storage space available, data management that eliminates redundancy and conserves storage capacity is central to effective implementation.

Only Unclassified and Sensitive But Unclassified (SBU) content will be placed on the PIC as currently envisioned. Future advancements in the technology and encryption capabilities may justify future changes in this assumption. Applications have specific read/write permissions. A "unique identifier" is required for each card issued. A DoD PIC management infrastructure and a configuration control process will be in place to manage the process of standardizing security requirements. The PIC and its supporting infrastructure will provide the following security services: confidentiality, availability, and integrity of data stored on the chip. All applications will complete a security certification and accreditation process. Readers/writers will provide the level of security consistent with the PIC security guidance indicated in the DoD Handbook.

6.3.3 Mass Storage

The way of looking at data storage has evolved with the development of personal computers, local area networks, and client-server architectures. Originally confined to mainframe computers and directly accessible by only a few MIS personnel, data today can be readily available with a few keyboard strokes or clicks on a mouse. Advances in technology have made the one-gigabyte disk drive the minimum starting point for a personal computer.

Multimedia applications require enormous amounts of storage space for images, video clips, and audio segments. The Small Computer System Interface (SCSI) is one of the critical parts of getting information from the computer to the user. The new UltraSCSI standard of 40 Megabytes per Second bus rate for transferring data will help cope with the demands of multi-user and multiprocessor systems. Tied to the UltraSCSI standard for data transfer is the evolving Fiber Channel-Arbitrated Loop (FC-AL) LAN to WAN standard that will advance mass storage capabilities by linking local area to wide area networks. FC-AL is a storage/ connectivity/ data access method that will provide faster data rates.

Another new technology called Solid State Disk (SSD) from Quantum, Inc. now provides users with a lightning fast alternative to magnetic disk drives. By massing powerful dynamic RAM (DRAM) chips into a compact chassis, SSD drives can perform data storage and retrieval tasks up to 30 times faster than a magnetic disk at the application level. The concept behind SSD is to load the application’s most heavily used files, or “hot files,” into DRAM, where the CPU can access them at electronic rather than mechanical speeds.

6.3.4 MMX

Intel has released Pentium processors containing its new MMX (multimedia extensions) feature, but significant performance improvements may come only from applications customized to the new instruction set for multimedia and communications. The first such packages are likely to be gaming and graphics applications, especially those involving image manipulation and 3-D rendering. MMX represents the first addition to Intel's base instruction set since the 1985 introduction of the 80386 processor.

MMX incorporates 57 new instructions that were developed with multimedia in mind. The instructions assist in processing video, audio, and graphics more efficiently. These instructions allow software developers to code for faster video frame rates, making for smoother, more realistic video. And they give them the ability to create more lifelike audio, thanks to improved noise reduction and music synthesis for compressed audio files.

Intel will initially ship 166MHz and 200MHz MMX desktop Pentiums and 150MHz and 166MHz MMX Pentiums for portables. Intel claims a 20% performance boost even with applications not optimized for MMX thanks to 32KB of internal cache. Testing with beta releases of image manipulation software optimized for MMX yielded performance gains as high as 133%.

Initial tests by NewMedia magazine showed that applications written to take advantage of Pentium MMX chips run from 15 to 40 percent faster than they do on standard Pentiums. Applications such as Adobe Photoshop 4.0 and Macromedia Director 6 showed significant performance gains with MMX. Entertainment software is expected to top the list of early MMX-enabled applications.

6.3.5 Media Streaming

Media streaming is the ability of transmitting a media file and viewing it in real time. The problem with the TCP/IP standard today is that while packet delays seldom exceed 100 to 200 msec for the User Datagram Protocol (UDP), it does not guarantee delivery of packets. The Transmission Control Protocol (TCP) guarantees that packets are delivered in the proper order, but the overhead causes delay. Today, most streaming media products use proprietary methods. Error mitigation for video is to freeze the frame until new data arrives; for audio, white noise is injected for the lost data.

Standard solutions have been proposed to correct the problem with the TCP/IP protocol and have been submitted to the Internet Engineering Task Force (IETF). The Real-Time Streaming Protocol (RTSP) is a protocol for transmitting audio and video across networks in real time. The audio and video can be accessed across a network and operated like using a tape recorder (stopping and starting on command). RTSP works with the Real-Time Transport Protocol (RTP) and can handle all application layer data flows via the Internet using UDP. It provides synchronization between source and destination, as well as a mechanism for handling lost packets and reducing delay variations. It also clarifies the use of translators, which can convert an RTP data stream from one format or bandwidth to another, and mixers, which can pull multiple streams into a single stream to conserve bandwidth. The Real-Time Transport Control Protocol (RTCP) ensures a channel of communication between server and client so that the application can make adjustments as needed to facilitate delivery. The Resource reSerVation Protocol (RSVP) provides a way for applications to request a certain bandwidth for a specified time for a stream of data (called a flow) from the routers along the data pathway. Once all the routers grant the request, the stream can move through the net at the rate that ensures quality playback on the user desktop. A time-stamp in the RTP/RTCP packets enables synchronization of time-dependent data.

6.3.6 Webcasting or “Push” Technologies

IP multicasting shortcuts the TCP/IP assumption that all packets routed need not arrive in a timely fashion. In implementing that assumption, TCP/IP routers examine and route each individual packet. The overhead associated with this routing can markedly delay throughput of multimedia data streams to an unacceptable level.

IP multicasting avoids part of this problem by only sending one copy of a data stream to a list of subscribers, rather than sending one copy to each subscriber. This reduces the overall burden on the network and improves throughput.

Its chief disadvantage is that it is a broadcast technique, with no opportunity to recover or retransmit any information not received. As such, multicasting techniques are at the low end of the assured-delivery spectrum.

6.3.7 Network Computers or “Thin Clients”

A technology being explored, but not yet in widespread use, is the deployment of austere workstations whose primary user interface is a browser application drawing all its content from networked servers, rather than maintaining substantial client-side content at the workstation itself.

6.3.8 Distance Learning

The House Appropriations Committee, in HAC Report 104-617, expressed support for the DoD’s increased use of distance learning techniques as a way to disseminate training materials and reduce the costs of training. In order to realize all of the benefits of distance learning, DoD personnel must utilize compatible systems that are described in standards and related documentation.

The Distance Learning Association and distance learning leaders in DoD, education and private sector activities commonly define distance learning as “structured learning that takes place without the physical presence of the instructor.” It is not limited to any specific presentation medium. Distance learning almost always involves the use of more than one medium (e.g., audio and videotapes; audio, video, and graphics software programs; and audio/video teleconferencing). The media being considered and used for distance learning include: 1) paper-based correspondence; 2) audio and video tapes; 3) optical media (analog laser discs, digital video discs); 4) computer-based instruction/training; 5) video teletraining; 6) embedded training; and 7) networked computers linked to Internet-based, electronic classrooms and campuses.

6.4 ACCOMMODATION FOR DISABLED OR PARTIALLY BUSY USERS

Choosing appropriate multimedia accommodation for disabled or partially busy users (e.g., hands-busy, eyes-busy, ears-busy) is normally done on a case-by-case basis. Individuals with disabilities have different capabilities and varying degrees of disabling conditions. Partially busy users have varying task requirements. Accommodation needs must be evaluated in light of an individual's job responsibilities and technical environment (e.g., workstation configuration). For these users and their supervisors, selecting accommodations can be challenging.

A needs assessment should be conducted to identify accommodations suitable to an individual’s specific situation. It is very important to select the right team of people to conduct a needs assessment. The affected users should always be involved. Others who may make valuable contributions include personnel from human resources, computer support services, occupational safety, procurement, facilities management, and state and community organizations, and the Computer-Electronics Accommodation Program (CAP) within the Office of the Assistant Secretary of Defense (Health Affairs).

Unified web site accessibility guidelines, prepared by the University of Wisconsin under funding from the National Institute on Disability and Rehabilitation Research, Office of Special Education and Rehabilitation Services, U.S. Department of Education, are available at http://trace.wisc.edu/docs/html_guidelines/version8.htm.

A needs assessment should address the following questions.

Job
What are the essential job requirements?
What skills (physical and cognitive) are required for job performance?

Individual
What are the individual's functional limitations to performing the job requirements?
Are there alternative approaches to accomplishing the essential job requirements?

Solutions

What accommodations, compatible with workstation, computer, and telecommunications systems, are available to meet the individual's needs?
Which alternative approach or accommodation is most suitable for the individual in

the specific work environment?

6.5 INTELLECTUAL PROPERTY RIGHTS

The need to protect copyrights and intellectual property rights is a key issue in the study of multimedia. As more kinds of works are presented digitally - such as digital music, digital video, and multimedia works that mix these forms - all the associated industries are trying to find solutions. A single document may have copyrighted text, pictures, and music extracted from different sources. The problem is to find a way to protect against unauthorized use of copyrighted material and where appropriate to compensate the owner of the copyright.

An introduction to this material, which is specifically targeted to multimedia developers, is the following: An Intellectual Property Law Primer for Multimedia and Web Developers by J. Dianne Brinson and Mark F. Radcliff.

The primer is based on the Multimedia Law and Business Handbook, which is designed to provide accurate information on the legal issues in multimedia. The primer is provided with the understanding that the authors are not engaged in rendering legal services. If you have a legal problem, you should seek the advice of experienced counsel. The primer is available at the Web site: http://www.eff.org/pub

6.5.1 Digital Watermarking

Much of the following information was obtained from an article entitled “Digital Watermarking” Jian Zhao, published in BYTE Magazine, Jan 97.

In contrast to a traditional watermark on a sheet of paper, a digital watermark can be detected only by appropriate software. Rather than ensuring the authenticity or integrity of documents, as a digital signature or a digital seal does, a digital watermark focuses on identifying the origin, author, owner, usage rights, distributor, or authorized user of an image, video clip, or audio clip. The watermark accomplishes this even if the image or clip has been processed and distorted (via analog-to-digital conversion, low-pass filtering, resampling, lossy compression, cropping, or rotation).

Although the digital watermark must exist in the material, it has to be invisible -- or, in the case of music and sound files, inaudible. Otherwise it could interfere with the quality of the document and also be more easily removed after illicit copying. (For more information on digital copyright protection, see "Handcuff Digital Thieves," BYTE international edition, April 1996.) A watermark must be able to withstand any processing and tampering. It must be present throughout the entire document in case parts are changed or removed.

Most common watermarking methods for graphics and audio signals work in the spatial, time, or frequency domains. The advantage of frequency-domain watermarking is that the watermark is spread throughout the entire video or audio clip and hence is resistant to cropping or cutting. However, a standard frequency filter, or a lossy compression algorithm, which usually filters out the less significant frequencies, could damage the watermark.

Watermarks can also be embedded in an image's luminance and color bands, or in the contour and texture of an image. Common watermarking methods use the luminance band since it contains the most significant information of a color image.

6.5.2 Watermarking Tools

There are two categories of watermarking tools available. The first is based on a method known as Fingerprinted Binary Information (FBI), as exemplified by a product from a UK company: HighWater FBI. The other, based on techniques developed at NEC Research and the University Catholique de Louvain (Belgium), identifies documents by hidden numbers (fingerprints). Other approaches, such as SysCoP (System for Copyright Protection), developed at the Fraunhofer Institute for Computer Graphics; Digimarc, from Digimarc (Portland, OR); and Argent, from DICE, can encode additional identification information.

Direct-sequence and frequency-hopping spread-spectrum techniques are the major watermark embedding methods used in existing tools. Both modify the noise value of the target documents. The direct-sequence technique adds noise to every element of the document, whereas the frequency-hopping method selects a pseudorandom subset of the data to be watermarked. Digimarc and FBI, for example, use direct-sequence methods to superimpose a watermark over an image by modulating a noise pattern of the same size onto the image. Other systems use secret keys to determine which lines or words of a text will be slightly shifted vertically or horizontally.

Hiding secret messages in the least-significant bits of some pseudorandom frequencies or pixels of an image can be considered a simple example of frequency hopping. Because frequency hopping modifies only a subset of pixels or other elements of a document, it tends to be much faster than direct-sequence methods. It is, however, less robust and more vulnerable to attack.

6.5.3 Watermark Extraction

A watermark must be extractable even from degraded documents that might have been photocopied, scanned, or manipulated by imaging programs. If a degraded document does not have the same format, resolution, or physical size as the original, it has to be normalized to the original format before the watermark can be extracted. Typical normalization processes include format conversion, resampling, enlarging a cropped part to full size, and scaling of the signal level.

Watermark extraction includes two main steps: (1 selecting the locations where the watermark has been inserted (only in frequency hopping); and (2 retrieving the watermark from those locations.

By applying multiplexing techniques as in data communications, the SysCoP system, for example, can embed several identification codes and extract them separately. This feature is extremely important for identifying ownership and other intellectual property rights in works composed of many copyrighted assets, such as a complex multimedia presentation.

Though commercial use of digital watermarking has begun -- Adobe is using Digimarc's technology in the new Photoshop 4, and IBM will use its own technology in its InfoMarket electronic commerce products -- several barriers still exist that prevent this technology from becoming effective and widespread. The technology for digital copyright watermarking is still in its early development. The major technical challenge is to develop a foolproof protection system while keeping watermarks hidden. Absolute robustness is impossible, but there is room for improvement as this technology gains more acceptance and wider use. Today, none of these systems can claim that its watermarks will survive all imaging or signal-processing operations. But like encryption, this technology will be useful as long as it makes tampering with or removing watermarks a time-consuming and costly task.

6.6 DISTRIBUTION, ARCHIVING AND STORAGE MEDIA

6.6.1 Distribution and Archiving

The distribution of multimedia products, because of the large size of the encoded video and audio data embedded in most such products, requires substantial attention to detail in the architecting and design of distribution and archiving systems. This attention is necessary to ensure the delivery of appropriate performance while controlling costs. The performance capabilities of the actual presentation device can be immaterial to the overall delivery of multimedia information if the distribution means is of substantially inferior performance.

Because human perception plays an important role in the delivery of multimedia information, recognition of that role (and of the limitations of human perception) in designing multimedia systems can provide great benefit in yielding acceptable performance at reasonable cost.

The key architectural building blocks in the distribution and archiving of multimedia information can be separated into networking and storage, with attention paid to the form of compression to achieve reductions in the capacity required of each. Note that portable storage media (e.g., DAT, DVD, and CD-ROM) may be used in the distribution of multimedia information with or without networks.

6.6.2 Multimedia Network Characteristics

Data (not audio or motion video), audio/voice, and video have different demands on the network and require different levels of accuracy on the data streams. Data, such as textual information, is not tolerant to data lost, but it can handle information being delayed or received in a bursty fashion. On the other hand, audio and video can lose data and still be useable, but it may not handle delay of the data stream well (the audio becomes choppy and difficult to understand; the video becomes jerky and may freeze the image at times). The following table shows some characteristics and demands for the various media data types.

Table 6-1. Media Data Characteristics and Demands

DATA TYPE	BANDWIDTH NEEDS	DATA STREAM	ERROR TOLERANCE	DELAY TOLERANT
DATA	moderate – useable at dial-up rates (33.6 KBPS)	bursty	none – retransmission of lost data (may introduce unpredictable delays)	yes
AUDIO/ VOICE	low to high – 8-176.4 KBPS (uncompressed)	steady	moderate – can tolerate minor error rate without losing information	no – delays can cause problems
VIDEO	High 27.7 Mbps (uncompressed, 24-bit color) 9.2 Mbps (uncompressed, 8-bit color) 550 KBPS (MPEG-1, 50:1 compressed, 24-bit color) 184 KBPS (MPEG-1, 50:1 compressed, 8-bit color)	steady	moderate – can tolerate minor error rate without losing information	no – delays can cause problems
	Note: 30 fps, 640 X 480 pixels/frame. Education and entertainment applications require data rate support between 1.5-6 Mbps. Video conferencing requires data rate support between .4-.7 Mbps (for .7 Mbps, roughly 10:1 compression of a 176 x 144 pixels/frame, 24-bit color and 10 fps; less bandwidth would have slower frame rate and/or smaller picture.)			

6.6.3 Archiving and Storage Requirements

Storage requirements need to be considered when sizing a system. Video, audio, and images all have the potential of using a great deal of storage space. The following are storage estimates for different media types.

1. Text (ASCII, EBCDIC)

   a. KB/page
   b. Typically 30 KB/file (with graphics 200 KB)
2. Image (bit-mapped graphics, still photos and faxes)
   a. Simple: 64 KB/image (uncompressed)
   b. Detailed: 7.5 MB/image (uncompressed)
   c. Diagnostic X-ray: 7 – 45MB/image and up (lossless compression)
3. Audio (noncoded stream of digitized audio or voice)
   a. Voice/phone (8 KHz/8-bit (mono)): 8 KBPS
   b. Audio CD DA (44.1 KHz/16-bit (stereo)): 176.4 KBPS
4. Animation (synched image and audio stream at 15-19 fps)
   a. 16 fps, 16-bit color, and 320 x 240 pixels/frame: 2.5 Mbps (uncompressed)
5. Video (TV analog or digital image with synched streams at 24-30 fps)
   a. 30 fps, 24-bit color, 640 x 480 pixels/frame : 27.7 Mbps (uncompressed) ~550 KBPS (MPEG-1, 50:1 compression)

6.6.3.1 Data Compression

Data compression is needed to conserve storage space and network bandwidth. Some media types (such as text, and medical images used for diagnosis) need to use a compression algorithm that restores the information to its original form (lossless compression). Other media types (such as video) can tolerate some loss of data by the compression algorithm (lossy compression) because the user can not perceive a difference from the original data or the perceived difference is acceptable (seeing a jerky picture while in a video conference may be acceptable, but a smooth picture is expected when watching a movie). Compression and decompression of data may be a symmetric or asymmetric process. A symmetric process (such as videoconferencing) would need the same processing power and provide the same conversion rate to compress, as well as decompress the data. An asymmetric process (such as an MPEG movie) requires more processing power and time to compress the data (to the required bandwidth and viewing quality) than to decompress the data (to view the picture). The following are common data compression standards:

1. H.261 — H.261 is an ITU standard for video conferencing. It supports a number of rates (112 KBPS, 384 KBPS, 768 KBPS, and 1.544 Mb/s).

2. JPEG — The Joint Photographic Experts Group (JPEG) standards are for the compression of photo-quality still images. Its compression range is 20-200:1. The actual compression ratio used depends on whether lossless or lossy compression algorithm is used. The useful compression range is 25-50:1. As the compression becomes greater than 50, artifacts (distortions) become more intrusive. A 10:1 compression ratio produces results difficult to distinguish from the original. In the lossless mode, JPEG usually produces compression ratios in the order of 2:1.

3. MPEG — The Moving Pictures Experts Group (MPEG) standards are for full-motion images. The MPEG-1 standard uses a lossy algorithm to compress the data and high compression ratios (150:1) are possible. The standard includes an algorithm for compressing audio in a range 5-10:1. MPEG-1 provides reasonable (broadcast and VHS-like) quality for data rates of 1.544 Mb/s. The MPEG-2 standard is for entertainment quality video and the compression algorithm supports data rates of 6 Mbps. The MPEG standards support various rates, matching the encoded signal to the source requirements and minimize the bandwidth used. For example, a 24-frame-per-second movie converted to NTSC quality needs 3 Mb/s bandwidth. A 24-fps movie converted to PAL quality needs 4 Mb/s bandwidth. If a fast moving show (such as a soccer game) is encoded, it needs 5-6 Mbps bandwidth to be viewed.

6.6.4 Estimating Storage and Bandwidth Requirements

Text storage requirements can be estimated using a nominal 30-40 characters per line and 50-60 lines per page times the number of pages of text divided by the approximate compression ratio, if compression is used.

Required still image storage in KB can be calculated according to the following formula:

> Image storage (KB) = ((Frame width in pixels) x (Frame height in pixels) x (Color depth in bits / 8) / (1024 x Average compression factor)

Required audio storage in KB can be calculated according to the following formula:

> Audio storage (KB) = ((Sampling size in bits / 8) x (Sampling rate in kHz x 1000)

x (Number of channels) x (Number of seconds of audio)) / 1024

Required video storage in KB can be calculated according to the following formula:

Video storage (KB) = ((Frame width in pixels) x (Frame height in pixels) x (Color depth in bits / 8) x (Frame rate in frames per second) x (Number of seconds of video)) / (1024 x Average compression factor)

The total amount of multimedia storage space in MB is then the sum of the video, audio, image, graphics, text, and data storage requirements divided by 1024. These factors must be adjusted to account for multiple copies needed for redundancy or distribution.

If network transmission is used for distribution, the bandwidth required is determined from the following formula:

Bandwidth (Mbps) = (Total storage required in MB * 10) x (Fraction which must be transmitted in one second) x (Transmission overhead for the circuit)

Transmission overhead on a TCP/IP circuit can range from a factor of 50% up to a factor of 800% for noisy satellite links.

Finally, in selecting storage media, the required remaining lifetime of the information must be considered. Contrary to popular belief, optical media do not have limitless shelf life; some literature reports that digital audio tape (DAT) media may last up to 30 years, while other reports claim that CD-ROMs degrade after 6-10 years.

APPENDIX A

REFERENCES

An Intellectual Property Law Primer For Multimedia And Web Developers, J. D. Brinson and M. F. Radcliffe, 1996

Digital Watermarking, Jian Zhao, BYTE, Jan 97

Federal Comp Week, Aug. 26, 1996, page 30

Fed Comp Week, Oct. 7, 1996, page 10

Government Computer News (GCN), Sept 23, 1996

Handcuff Digital Thieves, BYTE international edition, April 1996

Information Technology Standards Guidance (ITSG), version 3.1, April 7, 1997

Joint Integration and Engineering Organization (JIEO) Report 8300, July 19, 1996

Joint Integration and Engineering Organization (JIEO) Report 8303, Multimedia Extensions to the DoD Minimum Desktop Configuration

Joint Integration and Engineering Organization (JIEO) Report 8306, Guidelines for Selecting Multimedia Computing Equipment, March 1997

Joint Technical Architecture, version 2.0, July 18, 1997

Multimedia Technology Standards Assessment, version 2, August 1995

NewMedia, Scott Kahn, April 14, 1997

OMB Memorandum M-97-16, Information Technology Architectures, Franklin D. Raines, June 18, 1997

Technical Architecture Framework for Information Management (TAFIM), version 3.0, April 30, 1996

The Big Picture, Chris O'Malley, Computer Shopper, April 1996

Turbocharge your software, PC World: Jan 1997

Global Command and Control System (GCCS) Multimedia and Collaboration Services Software Requirements Specification (SRS)

GCCS Strategic Technical Architecture (Final)

Common Desktop Environment (CDE) Integration Specification v1.0

DII COE Office Automation SRS

GCCS Version Description Document (VDD) v2.2

GCCS Presentation Manager Functional Requirements v3.2

Joint Mapping Tool Kit SRS v2.0

Enhanced Linked Virtual Information System (ELVIS) SRS

Defense Information Infrastructure (DII COE) Baseline Specification v3.1

DII COE Integration and Runtime Specification (I&RTS) v3.0

DII COE User Interface Specification v2.0

Style Specifications for Web Applications in the DII

United States Imagery and Geospatial System (USIGS) Video Standards Profile

UCS-2 Encoding Form, Unicode Consortium

The Unicode Standard, Version 2.0 Addison-Wesley, 1996.

APPENDIX B

ACRONYMS

The acronyms used in this handbook are defined as follows:

AITS	-	Adopted Information Technology Standards
ANSI	-	American National Standards Institute
API	-	Application Programmer's Interface
ATM	-	Asynchronous Transfer Mode
BOPS	-	Billions of Operations Per Second
CALS	-	Continuous Acquisition Life Cycle Support
CD-ROM	-	Compact Disk-Read Only Memory
CGM	-	Computer Graphics Metafile
CMYK	-	Cyan Magenta Yellow black color model
C4I	-	Command, Control, Communications, Computers and Intelligence
C4ISR	-	Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance
CCITT	-	International Telephone and Telegraph Consultative Committee (see ITU)
CODEC	-	Compression/Decompression
COTS	-	Commercial off-the-shelf (also COTS products)
DAT	-	Digital Audio Tape
DII COE	-	Defense Information Infrastructure Common Operating Environment
DISA/CFS	-	Defense Information Systems Agency/Center for Standards

DISN	-	Defense Information Systems Network
DVD	-	Digital Video Disc; also known as Digital Versatile Disc
DXF	-	Document Exchange Format
E-mail	-	Electronic mail
FDDI	-	Fiber Distributed Data Interface
FPS	-	Frames Per Second
FTS 2000	-	Federal Telephone System 2000
GIF	-	Graphics Interchange Format
GKS	-	Graphical Kernel System
HDTV	-	High Definition Television
HTML	-	HyperText Markup Language
HTTP	-	HyperText Transfer Protocol
Hz	-	Hertz (cycles per second)
ICC	-	Integrated Circuit Card
IEC	-	International Electrotechnical Commission
IEEE	-	Institute for Electrical and Electronics Engineers
IETF	-	Internet Engineering Task Force
IGES	-	Initial Graphics Exchange Specification
IMA	-	Interactive Multimedia Association
IPI-IIF	-	Image Processing and Interchange-Image Interchange Format
ISDN	-	Integrated Services Digital Network
ISO	-	International Organization for Standardization
ITSG	-	Information Technology Standards Guidance

ITU	-	International Telecommunication Union (formerly CCITT)
ITU-R	-	ITU-Radiocommunication Sector
ITU-T	-	ITU-Telecommunication Standardization Sector
JBIG	-	Joint Bi-Level Imaging Group
JFIF	-	JPEG File Interchange Format
JPEG	-	Joint Photographic Experts Group
JTA	-	Joint Technical Architecture
JTC 1	-	Joint Technical Committee 1
KBPS	-	Kilobits per second
KBPS	-	Kilobytes per second
LAN	-	Local Area Network
LZW	-	Lempel-Ziv-Welch compression scheme
MCS	-	Multipoint Communications Services
MHz	-	(Megahertz) One million cycles per second
MIME	-	Multipurpose Internet Mail Extensions
MIS	-	Management Information System
MMX	-	MultiMedia Extensions
MPC	-	Multimedia Personal Computer
MPEG	-	Moving Pictures Expert Group
MPPP	-	Multilink Point-to-Point Protocol
NTSC	-	National Television Standard Committee signaling (U.S.)
PAL	-	Phase-ALternating signaling (International)
PDF	-	Portable Document Format

PHIGS	-	Programmers' Hierarchical Interactive Graphics System
PIC	-	Portable Information Carrier
PIKS	-	Programmer's Imaging Kernel System
PNG	-	Portable Network Graphics
POTS	-	Plain Old Telephone System
PPP	-	Point-to-Point Protocol
RAID	-	Redundant Array of Independent Disks
RAM	-	Random Access Memory
RGB	-	Red Green Blue color model
RTF	-	Rich Text Format
SBU	-	Sensitive But Unclassified
SC	-	SubCommittee
SCSI	-	Small Computer System Interface
SG	-	Study Group
SGML	-	Standard Generalized Markup Language
SLIP	-	Serial Line Internet Protocol
SMDL	-	Standard Music Description Language
SPDL	-	Standard Page Description Language
SPIFF	-	Still Picture Interchange File Format
STEP	-	STandard for the Exchange of Product model data
TAFIM	-	Technical Architecture Framework for Information Management
TIFF	-	Tagged Image File Format
VTC	-	Video TeleConferencing

WAN	-	Wide Area Network
WWW	-	World Wide Web

APPENDIX C

GLOSSARY

Algorithm: A procedure for solving a mathematical problem in a finite number of steps that frequently involves repetition of an operation.

Analog gateway: A means of connecting dissimilar codecs. Incoming digital signal from one type of codec is decoded by a similar codec and converted to analog. The analog signal is then passed to the dissimilar codec, coded, and decoded at the far end. Analog gateways achieve interoperability in a non-standard environment, but have the disadvantages of degrading video and audio quality and often reducing functionality.

Application Program Interface (API): An interface used by application programmers to gain access to an operating system, a device, another program, or a network. It typically consists of a collection of procedure calls with parameters.

Arithmetic encoding: A patented, proprietary mathematical entropy encoding process for compression of data. May provide better compression than Huffman encoding.

Asynchronous Transfer Mode (ATM): A digital transmission/switching scheme that is used for high-speed communications of video, voice, and data on a high bandwidth network that can range from T1 (1.5 Mbps) to OC12 (612 Mbps). ATM transmits data in 53-byte cells (5 header bytes and 48 bytes of payload). Since bandwidth can be allocated, ATM seems applicable to multimedia.

Audio conferencing: Conferencing that employs voice communications, usually accomplished using standard telephone lines. When more than one person is in a single location, speakerphones or special audio conferencing terminal equipment is employed. When more than two locations are involved, audio conferencing network equipment or services may be employed to improve the quality of the conference.

Audiographic: Documents and still image data (graphics, text, drawings, photos) with accompanying audio.

Audiovisual: Audio and/or video capability.

Authoring: A process implemented by a "developer" where a multimedia application is "programmed" into a series of icons, menus, hypermedia, etc.

Authoring system: Software that helps developers create multimedia programs or presentations without requiring the painstaking skills involved in traditional programming.

"B" frame: Bi-directional or B frames (sometimes called B pictures) refer to part of the MPEG video compression process whereby both past and future pictures/frames are used as references. B frames typically result in the most compression.

Bandwidth: Amount of transmission capacity. The complete range of frequencies over which a circuit or electronic system can function with minimal signal loss, typically less than 3 dB. Bandwidth can be dedicated to one user or shared among many users.

Baud: A unit of transmission speed equal to the number of discrete signal events (or symbols) per second. Baud is equivalent to bits per second in cases where each signal event represents exactly 1 bit.

Bit (Binary digit): The smallest unit of information in a binary number system; a bit can have a zero or a one value.

Bitmap: A pixel-by-pixel description of an image. Each pixel is a separate element. Also referred to as a raster image.

Bridge: An electronic device that mixes the signals from three or more locations (e.g., for audio or audiographic teleconferencing).

Broadband: Refers to networks capable of delivering high bandwidth, i.e., many bits per second.

Broadcast: A one-to-many transmission technology.

Byte: A unit of information consisting of eight bits, used mainly in referring to data transfer and data storage.

CD ROM (Compact-disc read-only-memory): A laser-encoded 5 ¼-inch optical disc typically used to store text, images, audio, video, and program which can run on computers with CD-ROM drives.

CODEC (Coder/Decoder): Hardware or software responsible for encoding, decoding, compression, and decompression of audio and video signals.

COErnel: Shorthand for "DII COE kernel" – the core components of the DII COE, including the operating system and extensions, a common desktop, software installation tools, security extensions, and printer services.

Collaboration: Working together in a joint effort (modified from the American Heritage Dictionary).

Colorspace: A system for measuring and describing color, e.g., CMYK (cyan, magenta, yellow, and black) is a common colorspace for printing.

Common Intermediate Format (CIF): A compromise television display format adopted by the ITU that is relatively easy to derive from both PAL and NTSC video signals.

Compatibility: Ability of components to interoperate without regard to hardware, data file structure, network, or other attributes.

Compression: A technique used to reduce the amount of digital information by representing redundant data with symbols and/or equations for the purpose of cost-effective digital transmission.

Conference: (As referred to in Multipoint Multimedia Communications) A group of geographically dispersed nodes that are joined together and that are capable of exchanging audiographic and audiovisual information across various communication networks.

Conferencing: 1. A generic term that refers to an application's ability to allow users to share the same file or bulletin board in real-time (synchronous), or at different times in a store-and-forward (asynchronous) manner. 2. Programs and meetings which may be for the purpose of presenting and exchanging information, comparing views, learning, planning, and decision-making. Conferences can be held in one location or conducted simultaneously at multiple locations and linked together by communications systems with the ability to contain images, annotations, or pointers.

Data compression: See Compression.

Data transmission rate (DTR): 1. The data transfer speed within the computer or between a peripheral and computer. 2. The data transmission speed in a network. (Can be expressed in bits per second (BPS)) (Same as Data Transfer Rate).

Defense Information Infrastructure Common Operating Environment (DII COE): An architecture, an approach, a collection of reusable software, a software infrastucture, , and a set of guidelines and standards designed to provide a framework for integration and to assist with interoperability while maximizing the use of COTS products.

Desktop: Refers to computers found on your desktop, such as the Personal Computer (PC) and Macintosh.

Desktop video conferencing: Video communication between users of desktop terminals (or PCs) that involves video and audio of the other party.

Digital: Information contained in the form of 0’s and 1’s.

Distance learning: Learning situations in which the information and the learner(s) are geographically separated.

Echo Cancellation: The process of reducing or eliminating sound originated at one end of a link, picked up by the far end microphone, and retransmitted to the originating end from being output to the speakers.

Encryption: The process of “encoding” data so it cannot be interpreted by anyone or any machine that does not have the “code.”

File Sharing/Transfer: The ability in a document conference to simultaneously transfer files to one or more conference participants.

Frames per second (fps): Frequency with which video frames appear on a monitor.

Glyph: Visual depiction of a character.

Graphic Document Sharing/Annotation: The ability in a document conference to simultaneously transfer, to one or more conference participants, a graphics document (such as scanning an image or faxing information) and possibly permitting the participants to annotate the document. (This is a higher resolution workspace compared to shared whiteboard).

Graphics: Electronically generated images in either raster or vector format.

Groupware: 1. A generic term for real-time data sharing programs that support team projects. 2. Tools designed to manage broad and unstructured activities using sophisticated messaging systems, e-mail, discussion databases, and other collaborative tools. 3. An information system that lets multiple users work on a single application simultaneously.

H.221: The ITU-T standard relating to the data formatting and framing protocol for voice and video in video conferencing.

H.261: The ITU-T Px64 standard relating to the video compression algorithm.

H.320: The ITU-T recommendation for a suite of international standards which describe specifications for distributing data during a narrowband videoconference supported by a variety of transmission media. H.320 specifies H.261 for video compression, H.221, H.230 and H.242 for communications, control and indication, respectively, G.711, G.722 and G.728 for audio depending on bandwidth used, and other specifications.

H.321: The ISDN implementation of the H.320 recommendation.

H.322: The local area network/Internet version of H.320, with quality specifications included.

H.323: The local area network/Internet version of H.320, without any quality specifications included.

H.324: The “plain old telephone system” (POTS) implementation of the H.320 recommendation.

HDTV: High Definition Television.

Hypertext: A document format where text and pictures can act as links to other places in the same or different documents.

HyperText Markup Language (HTML): 1. A document format where text and pictures can act as links to other places in the same or different documents. 2. An application conforming to ISO 8879--Standard Generalized Markup Language. It provides a simple way of structuring hypertext documents which refer to one another and which collectively create an enormous “web”.

HyperText Transfer Protocol (HTTP): The protocol used to access data on the World Wide Web.

Image: Collectively, the representations of objects reproduced electronically or by optical means on film, electronic display devices, or other media. (JCS Pub 1-02)

Imagery: Visual information represented and stored as a bitmap, a spatial matrix made up of individual picture elements (pixels in 2D images, voxels in 3D images)

Imaging: Creating a film or electronic image of any picture or paper form. It is accomplished by scanning or photographing an object and turning it into a matrix of dots (raster graphics), the meaning of which is unknown to the computer, only to the human viewer. Scanned images of text may be encoded into computer data (ASCII or EBCDIC) with page recognition software (OCR). (Techweb Technology Encyclopedia)

Interactive: The ability of a user to make decisions or selections that (can) alter the type and sequence of information or communication; the “give and take” to impact a decision.

Interoperability: The ability of hardware, software, systems, or services to function together without regard to manufacturer or service provider.

International Telecommunication Union (ITU): (Formerly known as CCITT.) An international body that sets worldwide telecommunications standards, such as the Px64 standard for video conferencing.

Joint Photographic Expert Group (JPEG): ISO standard for the compression of still pictures. Also can be used for compression of video for desktop multimedia.

Live Video: Video in which the motion observed at the point of display occurs either simultaneously or with no more than a few (e.g., seven) seconds' delay following the actual motion at the source; near-real-time or "broadcast" video.

MBone: "Virtual Internet backbone for Multicast IP," or Multicast backbone; a collection of sites on the Internet that support the IP multicasting protocol (one-to-many) and allow for live audio and videoconferencing.

MetaMail: A public domain UNIX multimedia mail package.

MHEG: Multimedia/Hypermedia Information Coding Expert Group; ISO standard (ISO 13522) and ITU (T.171) standard for interchange of multimedia documents which permits describes the logical structure of documents. It supports continuous media such as sound and motion video, and allows the description of time-dependencies between elements of the document. *The SGML mark-up language is used to mark up the logical structure; the HyTime mark-up language is used to describe the various temporal relationships of time-based information. The objective is to provide a structuring format for distribution over networks or on digital storage medium (such as CD-ROMs) of multimedia documents that can be interactively consulted but cannot be modified.*

MIME (Multipurpose Internet Mail Extension): An Internet standard (RFC 1341) for the support of multimedia-mail in which formats of message bodies are defined to permit multipart textual, audio, motion video, and image bodies.

M-JPEG: Motion JPEG is a variation of MPEG which does not use interframe coding and is thus easier to edit in a nonlinear editing system than full MPEG. (See Motion JPEG.)

MIDI (Musical Instrument Digital Interface): Interface standard used to connect musical instruments to computers.

Modem: Abbreviation for MOdulator/DEModulator. A modem converts digital signals to be transmitted in analog form, and demodulate signals produced by compatible devices. Typically used to connect a computer to a telephone line.

Motion JPEG: Though designed for still images, the JPEG encoding and compression standard is also used to compress motion video. Each image is encoded and compressed independently, resulting in a higher bit rate because redundancies between images are not reduced.

MPEG (Moving Picture Experts Group): The ISO standards body responsible for the MPEG-1 and MPEG-2 international video compression standards.

MPEG-1: ISO standard for the encoding and compression of digital motion video and associated audio, intended for storage of VCR-quality audio-video sequences on CD-ROMs with optimal functioning at 1.2 Mbps with Standard Interchange Format (SIF) images.

MPEG-2: ISO standard for the compression of motion video (to an entertainment quality picture) for transmission at 5 to 6 Mb/s.

Multicast: Single packets copied by the network and sent to a specific subset of network addresses. (In IP multicasting, messages are sent to all, unless users are explicitly excluded).

Multimedia: 1. Multiple Media. The combination of text, graphics, still images, sound, or video integrated as a single entity. 2. Two or more media types (audio, video, imagery, and text) in which the data streams are electronically manipulated, integrated, and reconstructed to maintain their time relationship (synchronization).

Multipoint: A communication system which allows three or more sites to both transmit and receive voice, video, and data signals to and from those specific points.

MUX (Multiplexer): Combines several different signals (e.g., video, audio, and data) onto a single communication channel for transmission. Demultiplexing separates each signal at the receiving end.

NTSC: Video signal conforming to the National Television Standards Committee RS-170A standard, with color (“chrominance”) information composited with brightness and shading information (“luminance”). Current U.S. broadcast standard; used in the U.S. and Japan. Nominally a 4:3 aspect ratio image of 525 scan lines with a 30Hz interlaced vertical refresh.

PAL: Current European broadcast standard.

Pixel: Picture element. In a digitized image, the smallest addressable point.

Point-to-Point Protocol (PPP): A protocol that encapsulates common network-layer protocols in specialized packets to pass through certain network commands.

Polymedia: In computing, this term refers to a more advanced and highly sophisticated multimedia system used in defense, industrial, or medical fields. Polymedia computer systems use supercomputers to format data to provide the user with immediate calculations for real-time animation or simulation. (Prentice Hall)

Portability: The ease with which application software and data can be transferred from one application platform to another.

Px64: The ITU-T's international video standard, which provides a standard algorithm for video compression and decompression. Formally known as H.261, it was adopted in December 1990.

Quality of Service (QOS): A methodology for delivering similar "content" at different "levels" of quality and/or service. This is determined by both the content provider and the user.

Recorded Video: Video in which the motion observed at the point of display occurs well after corresponding motion at the source, with the intervening time buffered by storage of the video on magnetic or optical media; "tape-delay" video.

Rendering: The manipulation of visual objects in real-time.

Small Computer System Interface (SCSI): A type of controller system that enables users to connect devices to their computer.

SECAM: French standard for analog video format.

Serial Line Internet Protocol (SLIP): A protocol that allows IP packets to be exchanged over a serial link or dial-up server. PPP is replacing SLIP.

SGML: Standard Generalized Markup Language. An ISO standard for defining the formatting in a text document.

Shared Whiteboard: The ability to view and annotate still images transmitted among participants in a document conference. This also can allow users to simulate a traditional "chalk-talk" business meeting over the communications channel.

Standards: Uniform specifications to permit interoperability that is manufacturer independent.

Still frame: A term for a captured frame of motion video or, less often, a still graphic image of any type. Also used to describe conferencing systems restricted to this mode of transmission and display.

Still Image: Non-moving visual information such as graphs, drawings, pictures, or video frames.

Streaming: The displaying of data (usually audio or video) as it arrives rather than waiting for all the data to arrive before displaying it.

Synchronous: Occurring at the same time; moving or operating at the same rate (from the American Heritage Dictionary).

Synchrony: A synchronous occurrence, movement, or arrangement (from the American Heritage Dictionary).

T.120: The ITU suite of international standards that describe specifications for distributing data during a multipoint, multimedia conference supported by a variety of transmission media.

Teleconferencing: The use of telephony communications for real-time conferencing among groups of people at two or more locations.

Transmission speed: Data rate for video conferencing, usually expressed in kilobits per second.

Video: A sequence of still images captured over a fixed time, at a fixed frame rate, and transmitted and displayed in synchronous order giving the appearance of live motion (e.g., video broadcasts, synthetic video, finished Intel products).

Video conferencing: A form of conferencing that employs voice and video communications, usually accomplished using digital channels that operate at 56 KBPS to 1.544 Mbps.

Video on demand: A proposed video service that would enable a consumer to select a video (or movie) and play it using VCR-like controls (play, stop, rewind, fast forward, and pause).

Video server: A computer that serves digital video streams to one or more client devices. Used extensively for video-on-demand.

Virtual reality: The art/science of creating worlds that are totally rendered by computers in which the user can immerse himself and take 360-degree perspectives.

Voice-activated switching: Used in multiway video conferencing so that all participating sites automatically see the site that is currently speaking.

Waveform audio file format (WAV): A file format for sampled audio used extensively on Windows-based PCs.

World Wide Web (WWW): The graphical section of the Internet whereby users have a simple interface for navigating and hyperlinking to other information around the Internet world. Mosaic and Netscape are two examples of products that help to enable the WWW.

APPENDIX D

OVERVIEW OF RELATED DOD GUIDANCE

The Department of Defense has recognized that information systems must be interoperable in order to be effective. If interoperability is to be achieved DoD-wide, mandatory and preferred systems and components must be identified and specific guidance must be provided to users. This is especially important considering the large number of both “standard” and “proprietary” systems in the marketplace.

The DoD has historically published several documents (see Appendix A) that identify information technology (including multimedia) standards and provide guidance for using those standards. The following paragraphs describe these documents and their relationships.

Documents such as the Joint Technical Architecture (JTA) and Technical Architecture Framework for Information Management (TAFIM) are both DoD guidance documents currently in effect within DoD. In using these documents, however, the guidance contained therein should be balanced against that provided by this handbook. This is important, because some of the information in those documents may now be outdated, documents being updated may not have revisions completed or fully coordinated, or both.

Specifically, the following documents should be reviewed as appropriate:

- ***Technical Architecture Framework for Information Management (TAFIM)* –** Provides general guidance and documents the processes and framework for defining the JTA (and other technical architectures). The TAFIM applies to many DoD mission/domain areas and lists all adopted information technology standards that promote interoperability, portability, and scaleability.

- ***Joint Technical Architecture (JTA)* –** Identifies the minimum set of standards for multimedia applications, providing listings of mandatory and emerging standards. The JTA is used to determine the specific standards to be implemented within new or upgraded Command, Control, Communications, Computers, and Intelligence (C4I) systems. Note that significant revisions are being made to the JTA, and that these revisions are still being coordinated within the Department.

  The JTA focuses on C4I requirements as related to interoperability by identifying the minimum set of standards for service areas (one standard per function where possible). For the C4I service area domain, the JTA set of standards supersedes those listed in the TAFIM. In general, the JTA shall be used to determine the specific standards to be implemented within new or upgraded C4I systems. However, there are three key considerations in using the JTA (as excerpted from the JTA):

1. “First, the mandatory standards in the JTA must be implemented by systems that have a need for the corresponding services. That is, a standard is mandatory in the sense that if a service/interface is going to be implemented, it shall be implemented in accordance with the associated standard. If a required service can be obtained by implementing more than one standard (e.g., satellite standards), the appropriate standard should be selected based on system requirements.

2. Second, the JTA is a "forward-looking" document. It guides the acquisition and development of new and emerging C4I functionality and provides a baseline towards which existing systems will move. It is not a catalog of all information technology standards used within today's DoD systems. It represents those standards (for interfaces/services) that should be used now and in the future. If legacy standards are needed to interface with existing systems, they can be implemented on a case-by-case basis in addition to the mandated standard.

3. Third, specification of any other standards (outside of those identified in the JTA) must be additive, complementary, and non-conflictive with JTA mandated standards. Refer to the for adopted standards in areas not addressed by the JTA.”

- ***Adopted Information Technology Standards (AITS; Volume 7 of the TAFIM) and Information Technology Standards Guidance (ITSG)*** – Support the JTA by tracking activity in new standards that may be included in the JTA at a future date. Contain specific information on mandated, informational, and emerging standards. Part 12 of the ITSG contains information concerning Multimedia Services. The ITSG provides information about those standards not found in the JTA or the AITS. The ITSG also provides more information about standards addressed in the AITS.

- ***MIL-HDBK-1379-3, Development of Interactive Multimedia Instruction*** **–** Provides background information for the planning, design, development, implementation, evaluation, and management of Interactive Multimedia Instruction (IMI) products. IMI products include electronic products used in the delivery of instruction or supporting the delivery of instruction. It provides a very valuable resource for insight into what is involved in a multimedia product and it takes into account the information contained in JIEO Report 8300 by reference. The handbook also provides acquisition guidance unique to IMI products.

- ***JIEO Report 8300, Minimum Desktop Personal Computer Configuration*** **–**

Identifies the minimum desktop personal computer (PC) configuration for computers that will be used in DoD-wide systems and infrastructure. The report specifies a minimum configuration required for desktop office automation and other terminals to serve as a DoD baseline PC. That baseline will be able to support all of the Defense Information Infrastructure (DII) systems, and includes support for new technologies such as multimedia. The procurement specifications for PC systems to support the DII may need to exceed the requirements in the 8300 Report. However, the specifications should not be diluted during acquisition, as the report specifies a *minimum* that is needed to assure that the systems can operate within the DII.

- ***JIEO Report 8303, Multimedia Extensions to the DoD Minimum Desktop Configuration* –** Provides an addendum to JIEO Report 8300. The report recognizes that multimedia platforms are not limited to PCs, but identifies that PCs are the most common platform architecture that is available to the DoD community for purchase. A set of PC configuration specifications for the needs of both the developer and the user are shown in an Annex to Report 8303. The specifications in 8303 identify the recommended characteristics for the CPU, Hard disk drive, Device bus, Mouse, Audio Card, CD-ROM drive, and Video display controller.

- ***JIEO Report 8306, Guidelines for Selecting Multimedia Computing Equipment* –** Provides guidance for developing a specification for a computer to be obtained through an acquisition. The resulting specification is expected to be at a level of granularity that will match off-the-shelf hardware and software from the general-purpose consumer market.

Table D-1 provides a high-level summary of the specific standards and guidance from these and other source documents (e.g., GCCS and DII COE Software Requirements Specifications) enumerated in Appendix A, References.

Table D-1. Summary of Standards and Guidance by Source Documentation

Guidance	**Source(s)**
Text and Document Encoding and Fonts	
MIME (RFC 1521)	GCCS
Secure-MIME (S/MIME)	GCCS
ASCII and Unicode text	DII COE, AITS/ITSG
TrueType fonts	DII COE, AITS/ITSG
Preferred font color	DII COE
Preferred typeface and point size, capitalization	DII COE
Hyperlink distinguished from regular underlined text	DII COE
UCS and ISO 8-bit	AITS/ITSG
Adobe Type I fonts	AITS/ITSG

Guidance	Source(s)
Graphics	
DIB/BMP	GCCS, JTA, AITS/ITSG
EPS	GCCS
Framemaker Interchange Format	GCCS
GIF87a	GCCS, AITS/ITSG
GIF89a	GCCS, JTA, AITS/ITSG
IGES	GCCS
JPEG	GCCS, JTA, AITS/ITSG
Lotus PIC	GCCS
MacPaint	GCCS
MS Paint	GCCS
PCX	GCCS, JTA, AITS/ITSG
PPM	GCCS
Sun	GCCS
TIFF R and G	GCCS, JTA, AITS/ITSG
WPG	GCCS
XBM	GCCS
XPM	GCCS
XWD raster graphics	GCCS
ISO CGM 89 and 93 (FIPS 128-1)	GCCS, JTA, AITS/ITSG
Group 3 and 4 FAX	GCCS, JTA, AITS/ITSG
VGA and SVGA resolution	DII COE
Win32 API for 2-D graphics (WinG optional)	DII COE
OpenGL for 3-D graphics	DII COE
Interlaced GIFs and progressive JPEGs for still images	DII COE
PNG	JTA
MIL-PRF-28000, 28002 and 28003 for CALS	JTA, AITS/ITSG
OpenGL	JTA
VRML	JTA, AITS/ITSG
MIL-STD 188-161D FAX	JTA
WMF	AITS/ITSG
PICT	AITS/ITSG
DXF	AITS/ITSG
IGES	AITS/ITSG
Renderman	AITS/ITSG
JBIG	AITS/ITSG

Guidance	Source(s)
Photo CD	AITS/ITSG
TGA	AITS/ITSG
FLI/FLC	AITS/ITSG
Document Interchange	
ASCII	GCCS, JTA
HTML (including HTML 2.0 and 3.2)	GCCS, JTA, AITS/ITSG
RTF	GCCS, JTA, AITS/ITSG
Postscript	GCCS, JTA, AITS/ITSG
PowerPoint 3 and 4	GCCS, JTA
SGML	GCCS, JTA, AITS/ITSG
PDF	GCCS, JTA, AITS/ITSG
ISO HyTime	GCCS, JTA, AITS/ITSG
WP 5.2 and 6.0	GCCS, JTA
Word 2 and 6	GCCS, JTA
Lotus WKS/1/3	GCCS, JTA
Excel 5	GCCS, JTA
DIF	GCCS
ACR/NEMA DICOM 3	GCCS, JTA, AITS/ITSG
dBase 4	JTA
GZIP and ZIP compression	JTA
MIL-STD-1840 and MIL-PRF-28001 for CALS	JTA
Freelance 2	JTA
Color Definition	
Grayscale through 24-bit color	GCCS
Process	GCCS
PMS	GCCS
CMYK	GCCS, AITS/ITSG
Pantone color selection	GCCS, AITS/ITSG
CIE	AITS/ITSG
RGB	AITS/ITSG
HIS	AITS/ITSG
TIFF	AITS/ITSG
24-bit	AITS/ITSG
Printer Data Interchange	
Postscript	GCCS, AITS/ITSG
PCL	GCCS

Guidance	Source(s)
HPGL4	GCCS
Through OS only	DII COE
To 8.5x11 and ISO A4 paper	DII COE
EPS	AITS/ITSG
PDF	AITS/ITSG
SPDL	AITS/ITSG
HPDL	AITS/ITSG
Video and Audio	
ISO/IEC 11172-2 (MPEG-1)	GCCS, JTA, AITS/ITSG
ISO/IEC 13818-2 (MPEG-2)	GCCS, JTA, AITS/ITSG, USIGS
SND	GCCS
WAV	GCCS, JTA, AITS/ITSG
ADPCM	GCCS, AITS/ITSG
AIFF	GCCS, AITS/ITSG
MIDI	GCCS
VOC	GCCS, AITS/ITSG
Stereo and mono audio	GCCS, USIGS
QuickTime	GCCS, JTA, AITS/ITSG
AVI	GCCS, JTA, AITS/ITSG
No looping except background audio	DII COE
MPEG-4 for low bit rates	JTA
HDTV and ATSC	JTA, USIGS
Internet Phone	JTA
Streamworks	JTA
Indeo/DVI	JTA, AITS/ITSG
MIDI	AITS/ITSG
CCIR 601 (ITU-R BT601-4) 4:2:2 component, uncompressed digital video	JTA, USIGS
Analog and Other Video Sources and Products	
NTSC and RS-170A (ANSI/SMPTE 170M, 525/60)	GCCS, JTA, AITS/ITSG, USIGS
PAL (625/50)	GCCS, AITS/ITSG
SECAM (625/50)	GCCS, AITS/ITSG
S-Video	GCCS

Guidance	Source(s)
Compatibility with Kodak DCS and Sony DVBK-1000 digital cameras and commercial frame grabber boards (e.g., Sun VideoPix)	GCCS
SMPTE VITC video time base	JTA, AITS/ITSG, USIGS
SMPTE 259M compressed video	USIGS
SMPTE videotape standards, with VHS, U-Matic, D2 and D3 prohibited for acquisition, archiving, or processing	USIGS
EIA-608 closed captioning until replaced by digital standards	USIGS
Web Access and Multimedia Presentations	
HTTP and CGI links to servers via TCP/IP	GCCS, JTA, AITS/ITSG
Netscape browser	GCCS
Java	GCCS, DII COE, JTA, AITS/ITSG
JavaScript	GCCS, DII COE, JTA, AITS/ITSG
TLS – Secure Sockets Layer (SSL) 3.0	GCCS, JTA, AITS/ITSG
Web casting	GCCS
Accomplished via HTML browser	GCCS
HTML 3.2, CGI and frames for servers	DII COE
Prefer no frames and HTML 3.2(-) for clients or alerts for users with less capable browsers	DII COE
Page design and organization (e.g., HTML minimum contents, image size, viewers)	DII COE
IRC	DII COE
Multimedia Authoring	
Netscape Composer	GCCS
Asymmetric Toolbook II	GCCS
Microsoft FrontPage	GCCS
MacroMedia Director	GCCS
MacroMedia Authorware	JTA
MS Visual Basic	JTA
Teleconferencing	
VTC001 Profile	GCCS, JTA, AITS/ITSG
T.120 and H.32x compatible	GCCS, JTA, AITS/ITSG
Shared whiteboards (including maps and images), audio, chatter, video and applications	GCCS
Synchronous point-to-point and multipoint	GCCS, JTA, AITS/ITSG
ITU-T G.700 series or MPEG-2 audio	GCCS
H.324 for low bit rate	JTA, AITS/ITSG

Guidance	Source(s)
MCG&I Display	
Arc Digitized Raster Graphics (ADRG) and Compressed ADRG (CADRG)	GCCS
Arc Digitized Raster Imagery (ADRI)	GCCS
Compressed Aeronautical Chart (CAC)	GCCS
Meteorological Image Format (MIF)	GCCS
Raster Product Format (RPF)	GCCS, JTA, AITS/ITSG
Vector Product Format (VPF)	GCCS, JTA, AITS/ITSG
NITFS v1.1 and v2.0	GCCS, JTA, AITS/ITSG, USIGS
JMTK for image preprocessing and geographic map display	DII COE
Platforms and Distribution	
SVGA (VESA) and VGA-compatible	AITS/ITSG
NTSC, PAL and SMPTE studio monitors	AITS/ITSG
HDTV	AITS/ITSG
MPC	AITS/ITSG
CD-ROM	AITS/ITSG
DVD	AITS/ITSG

APPENDIX E

SELECTION CRITERIA CROSS-REFERENCE

Criteria	Relevant Sections
Input Source(s)	2
Text	2.1, 2.3
Vector or Animated Graphics	2.5, 2.6
Still Images and Raster Graphics	2.6, 2.7
Audio	2.8
Video	2.9
Compound Documents	2.2, 3
Intended Application	2, 3, 4
Teleconferencing	2.8, 2.9, 4
Training	2.8, 2.9, 4
Required Performance/Throughput	6.3.6
Protocol Used	
Interactive on-line	
Interactive off-line	
Message-oriented on-line	
Message-oriented off-line	

Communications/Distribution Path	6.6.1
Hardcopy	2.4
Other physical media	
Very high bandwidth (> T-1)	6.6.2, 6.6.4
High bandwidth (384Kbps – T-1)	6.6.2, 6.6.4
Medium bandwidth (56Kbps – 384Kbps)	6.6.2, 6.6.4
Low bandwidth (Less than 56Kbps)	6.6.2, 6.6.4

APPENDIX F

COST-BENEFIT ANALYSIS CHECKLIST

1. Identify and characterize the features associated with the system to be acquired.

   Mandatory
   Won't work without these features

   Critical
   Will still work without these key features

   Desirable
   Nice-to-have, but subject to resource constraints

   Bonus
   Included but not worth paying to get
   May duplicate other capabilities already available

   Unacceptable
   Included but counterproductive

2. Ensure that the technologies to be incorporated are broadly available within the intended target audience domain.

   Is special hardware required for playback?

   Are special user needs addressed?

3. Assess the technical feasibility and appropriateness of the total system solution.

   Technology and Product Availability
   When integrable products can be expected?
   How much integration is required, and are there precedents already fielded with identical integration?

   Information and Communication Requirements
   Real-world availability of required information/source data and supporting communications paths and bandwidth

Environmental Suitability
Will the system fit within space, temperature, humidity and other related constraints?

Support Requirements
What are the training, consumables resupply, maintenance, and power requirements and implications?

Security Implications
How will any required identification and authentication of authors and users, auditing, and object reuse functionality and assurance be supplied and implemented?

Cost "Reasonableness"
Cost growth risk estimates
Commercial comparisons, if available

4. Assess the relative merit of any incremental improvements over lower-cost or easier-to-implement solutions.

   Will the average user distinguish the value of those incremental improvements (i.e., in color and spatial resolution, frame rate, etc.)?

   Will significant savings be gained in authoring multimedia data streams?

   Will any required long-term preservation of multimedia data be enhanced?

5. Ensure that all investment costs will be recovered from savings after operating costs are deducted, within four years. This assessment must consider total cost of ownership, including all support costs (e.g., training, maintenance, software upgrades).