# Internet Standards for the Web

## Tutorial Notes
## April 1998

Larry Masinter
Xerox Palo Alto Research Center

# About the author: http://www.parc.xerox.com/masinter

## Larry Masinter

Xerox Palo Alto Research Center
mail: masinter@parc.xerox.com
fax: (650) 812-4333

**This is my External page. Xerox users can also see my internal web page.**

At PARC, I've been leading a project building an infrastructure for offering document services over the net; we'll go public sometime soon, and then I can say more.
In the Internet and Web communities, I work on building standards to improve interoperability. I'm an active participant in the Internet Engineering Task Force, which defines Internet standards. I'm a member of the Application Area Directorate.
I'm chair of the HTTP working group. After years of work, haggling, and deliberation, we're nearly done with the HTTP/1.1 draft standard; HTTP/1.1 improves the performance and reliability of the web significantly, and I'm proud to have led the group.
I was the chair of the IETF URI working group, working toward defining standards for locating, naming, and describing resources on the Internet. The URI working group closed, but there are several related efforts ongoing. With Roy Fielding, I'm a document editor for the revised URL syntax specification; I'm also document editor for and also for the URL process working group.
I'm active in the Internet Fax working group, and have authored several Internet Drafts for that group. As part of the work on Internet Fax, I've participated in the TR-29.1 US delegation and Study Group 8, Question 4 of the ITU-T (International Telecommunications Union).
I've been active in the Web Distributed Authoring and Versioning (WEBDAV) group, Internet Printing, URNs, HTML over mail, electronic commerce, content negotiation, caching and proxy, Internationalization, use of character sets in Internet protocols, etc.
I'm also active in the World Wide Web Consortium (W3C), as a member of the advisory board.
In addition to IETF and W3C work, in the **Document Management** area, I have worked with Xerox groups defining architecture and strategy for document management products, and have participated in the Document Management Alliance group.
I retain a strong interest in the **Digital Libraries** area, where I coordinated various digital library activities and Xerox's involvement in the various NSF-sponsored digital library projects. I've been on the executive committee of the Stanford Digital Library project.
I've given several talks and seminars about digital libraries.

### Some Presentations:

- (10/96) Internationalization and Multilingualism in Web Standards (also in Powerpoint).
- (5/96) Tutorial/State of the art report The State of Web Standards for Fifth International World Wide Web Conference. In Powerpoint, too.
- (3/96) Digital Libraries, Document management, and the World-Wide Web presentation for Multi-media Japan.
- (1/96) Document Images on the Internet for 1/96 SPIE conference.

- (11/95) PARC Forum, "Document Management and the Net". Most of the info is in the much smaller html version.

### Internet RFCs

- **RFC 1737** "Functional Requirements for Uniform Resource Names." K. Sollins & L. Masinter. December 1994.
- **RFC 1738** "Uniform Resource Locators (URL)." T. Berners-Lee, L. Masinter & M. McCahill. December 1994.
- **RFC 1867** "Form-based File Upload in HTML." E. Nebel & L. Masinter. November 1995. (See list of implementations and errata.)

### Internet Drafts (works in progress)

- Internet Fax related:
  - **fax-smtp-session** "SMTP Service Extension for Immediate Delivery." Larry Masinter, N. Joffe, Dan Wing. 02/98.
  - **fax-fpim** "Extended Mode of Facsimile Using Internet Mail." Larry Masinter, Dan Wing. 01/02/1998. *Update expected soon.*
  - **fax-goals** "Terminology and Goals for Internet Fax." Larry Masinter. 01/02/1998.
  - **fax-mdn-features** "Using Message Disposition Notifications to Indicate Supported Features", Larry Masinter, Dan Wing, 03/10/1998.
  - **masinter-media-features** "Media Features for Display, Print, and Fax." Larry Masinter, Koen Holtman, Andy Mutz, Dan Wing. March 1998.
- HTTP related
  - **http-v11-spec-rev** "Hypertext Transfer Protocol -- HTTP/1.1." J Mogul, Tim Berners-Lee, Larry Masinter, Paul Leach, Roy Fielding, H Nielsen, J Gettys. 02/24/1998.
- URL related
  - **masinter-url-data** "The 'data' URL scheme." Larry Masinter. 03/19/1997.
  - **hoffman-mailto-url** "The mailto URL scheme." Larry Masinter, Paul Hoffman, J. Zawinski. 01/02/1998.
  - **masinter-url-i18n** "Using UTF-8 for non-ASCII Characters in Extended URIs." Larry Masinter. 06/09/98.
  - **fielding-uri-syntax** "Uniform Resource Identifiers (URI): Generic Syntax and Semantics", Tim Berners-Lee, Larry Masinter, Roy Fielding.
  - **urlreg-guide** "Guidelines for new URL Schemes", Harald Alvestrand, Larry Masinter, Dan Zigmond, 12/22/1997.
- Other web-related
  - **masinter-form-data** "Return Values from Forms: multipart/form-data." Larry Masinter. 01/09/1998.

### Some other online papers and publications:

- (10/95) Document Management and Electronic Commerce for International Conference on Electronic Commerce.
- (6/95) INET'95: Document Management, Digital Libraries and the Web. There's also a newer draft available than what went into the online proceedings.
- (12/94) Collaborative Information Retrieval: Gopher from MOO at INET'94.
- Script for a 1994 video about the Internet.

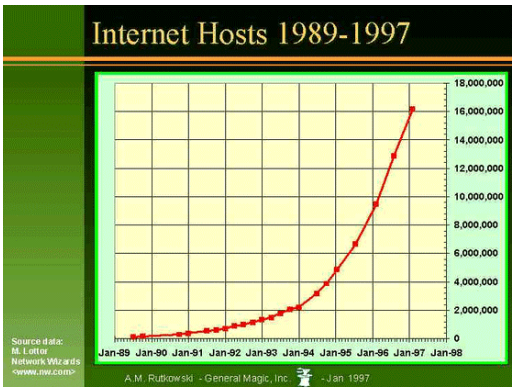
### Other activities:

- Help instigate HTTP-NG activity.
- Organizing committee for Web Internationalization and Multilinguism Symposium.
- Editorial board, Electronic Commerce World online journal.
- (7/96) ACM workshop on Strategic Directions in Computer Science, helped with Strategic Directions in Networking Research report.
- (1/95)Program committee, awards committee, Fifth World Wide Web conference.
- Program committee for the Fourth World Wide Web conference.
- I helped script and produce a video about the Internet for Xerox management.
- 1992 ACM Software System Award for Interlisp.

# Internet statistics

/

## http://www.genmagic.com/Internet/Trends/slide-3.html



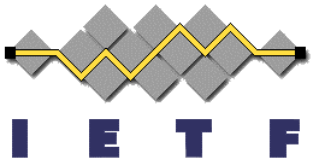

## http://www.openmarket.com/intindex

## The Internet Index

Number 21
Inspired by "Harper's Index"*
Compiled by Win Treese ( treese@OpenMarket.com)
21 January 1998
Here is Internet Index #21 and the directory of all editions.
Percentage of on-line users who watch TV and PC screens simultaneously: 40
Source: Research by Showtime, reported in Wall Street Journal, 9/11/97, p. R15
Percentage of Webmasters who like the term Webmaster": 44.6
Source: Web Week survey, 9/22/97
Circulation base of Web Week: 125,000
Source: Web Week, 1/5/98
Circulation base of BusinessWeek: 1,025,000
Source: BusinessWeek web site
Percentage increase in number of on-line auction sites listed by Yahoo, April to June, 1997: 140
Source: Inc. Technology 1997 No. 3, p.58

Good statistics on Internet size are hard to come by.

About the IETF




# The Tao of IETF -- A Guide for New Attendees of the Internet Engineering Task Force

• • •

## What is the IETF?

The Internet Engineering Task Force is a loosely self-organized group of people who make technical and other contributions to the engineering and evolution of the Internet and its technologies. It is the principal body engaged in the development of new Internet standard specifications. Its mission includes:

- Identifying, and proposing solutions to, pressing operational and technical problems in the Internet;
- Specifying the development or usage of protocols and the near-term architecture to solve such technical problems for the Internet;
- Making recommendations to the Internet Engineering Steering Group (IESG) regarding the standardization of protocols and protocol usage in the Internet;
- Facilitating technology transfer from the Internet Research Task Force (IRTF) to the wider Internet community; and
- Providing a forum for the exchange of information within the Internet community between vendors, users, researchers, agency contractors and network managers.

The IETF meeting is not a conference, although there are technical presentations. The IETF is not a traditional standards organization, although many specifications are produced that become standards. The IETF is made up of volunteers who meet three times a year to fulfill the IETF mission.

There is no membership in the IETF. Anyone may register for and attend any meeting. The closest thing there is to being an IETF member is being on the IETF or working group mailing lists (see the IETF Mailing Lists section). This is where the best information about current IETF activities and focus can be found.

• • •

## The Hierarchy

To completely understand the structure of the IETF, it is useful to understand the overall structure in which the IETF resides. There are four groups in the structure: the ISOC and its Board of Trustees, the IAB, the IESG and the IETF itself.

The Internet Society is a professional society that is concerned with the growth and evolution of the worldwide Internet, with the way in which the Internet is and can be used, and with the social, political, and technical issues which arise as a result. The ISOC Trustees are responsible for approving appointments to the IAB from among the nominees submitted by the IETF nominating committee.

The IAB is a technical advisory group of the ISOC. It is chartered to provide oversight of the architecture of the Internet and its protocols, and to serve, in the context of the Internet standards process, as a body to which the decisions of the IESG may be appealed. The IAB is responsible for approving appointments to the IESG from among the nominees submitted by the IETF nominations committee.

The IESG is responsible for technical management of IETF activities and the Internet standards process. As part of the ISOC, it administers the process according to the rules and procedures which have been ratified by the ISOC Trustees. The IESG is directly responsible for the actions associated with entry into and movement along the Internet "standards track," including final approval of specifications as Internet Standards.

The IETF is divided into eight functional areas. They are: Applications, Internet, IP: Next Generation, Network Management, Operational Requirements, Routing, Security, Transport and User Services. Each area has one or two area directors. The area directors, along with the IETF/IESG Chair, form the IESG. Fred Baker is the current IETF/IESG chair.
Each area has several working groups. A working group is a group of people who work under a charter to achieve a certain goal. That goal may be the creation of an Informational document, the creation of a protocol specification, or the resolution of problems in the Internet. Most working groups have a finite lifetime. That is, once a working group has achieved its goal, it disbands. As in the IETF, there is no official membership for a working group. Unofficially, a working group member is somebody who is on that working group's mailing list; however, anyone may attend a working group meeting (see the Be Prepared section below).
Areas may also have Birds of a Feather (BOF) sessions. They generally have the same goals as working groups, except that they have no charter and usually only meet once or twice. BOFs are often held to determine if there is enough interest to form a working group.

• • •

## Mailing Lists and Archives

As previously mentioned, the IETF announcement and discussion mailing lists are the central mailing lists for IETF activities. However, there are many other mailing lists related to IETF work. For example, every working group has its own discussion list. In addition, there are some long-term technical debates which have been moved off of the IETF list onto lists created specifically for those topics. It is highly recommended that everybody follow the discussions on the mailing lists of the working groups which they wish to attend. The more work that is done on the mailing lists, the less work that will need to be done at the meeting, leaving time for cross pollination (i.e., attending working groups outside one's primary area of interest in order to broaden one's perspective).
The mailing lists also provide a forum for those who wish to follow, or contribute to, the working groups' efforts, but cannot attend the IETF meetings.
All IETF discussion lists have a "-request" address which handles the administrative details of joining and leaving the list. It is generally frowned upon when such administrivia appears on the discussion mailing list.
Most IETF discussion lists are archived. That is, all of the messages sent to the list are automatically stored on a host for anonymous FTP access. To find out where a particular list is archived, send a message to the list's "-request" address, NOT to the list itself.

• • •

## Be Prepared

This topic cannot be stressed enough. As the IETF grows, it becomes more and more important for attendees to arrive prepared for the working group meetings they plan to attend. This doesn't apply only to newcomers--everybody should come prepared.
Being prepared means having read the documents which the working group or BOF chair has distributed. It means having followed the discussions on the working group's mailing list or having reviewed the archives. For the working group/BOF chairs, it means getting all of the documents out early (i.e., several weeks) to give everybody time to read them and announcing an agenda and sticking to it.
At the chair's discretion, some time may be devoted to bringing new working group attendees up to speed. In fact, long lived working groups have occasionally held entire sessions which were introductory in nature. As a rule, however, a working group is not the place to go for training. Observers are always welcome, but they must realize that the work effort cannot be delayed for education. Anyone wishing to attend a working group for the first time might seek out the chair prior to the meeting and ask for some introduction.
Another thing for everybody to consider is that working groups go through phases. In the initial phase (say, the first two meetings), all ideas are welcome. The idea is to gather all the possible solutions together for consideration. In the development phase, a solution is chosen and developed. Trying to reopen issues which were decided more than a couple of meetings back is considered bad form. The final phase (the last two meetings) is where the "spit and polish" are applied to the architected solution. This is not the time to suggest architectural changes or open design issues already resolved. It's a bad idea to wait until the last minute to speak out if a problem is discovered. This is especially true for people whose excuse is that they hadn't read the documents until the day before a comments period ended.
Time at the IETF meetings is a precious thing. Working groups are encouraged to meet between IETF meetings, either in person or by video or telephone conference. Doing as much work as possible over the mailing lists would also reduce the amount of work which must be done at the meeting.

## RFCs and Internet-Drafts

Originally, RFCs were just what the name implies: requests for comments. The early RFCs were messages between the ARPANET architects about how to resolve certain problems. Over the years, RFCs became more formal. It reached the point that they were being cited as standards, even when they weren't.
To help clear up some confusion, there are now two special sub-series within the RFCs: FYIs and STDs. The For Your Information RFC sub- series was created to document overviews and topics which are introductory. Frequently, FYIs are created by groups within the IETF User Services Area. The STD RFC sub-series was created to identify those RFCs which do in fact specify Internet standards.
Every RFC, including FYIs and STDs, have an RFC number by which they are indexed and by which they can be retrieved. FYIs and STDs have FYI numbers and STD numbers, respectively, in addition to RFC numbers. This makes it easier for a new Internet user, for example, to find all of the helpful, informational documents by looking for the FYIs amongst all the RFCs. If an FYI or STD is revised, its RFC number will change, but its FYI or STD number will remain constant for ease of reference.
…
For additional information, read the following documents:

- Request for Comments on Request for Comments [RFC 1111]
- F.Y.I. on F.Y.I: Introduction to the F.Y.I notes [FYI1]
- Introduction to the STD Notes [RFC 1311]
- Guidelines to Authors of Internet-Drafts [GAID]
- The Internet Activities Board [RFC 1160]
- The Internet Standards Process [RFC 2026]
- Internet Official Protocol Standards [STD1]

• • •

## IETF Area Abbreviations

- APP - Applications
- INT - Internet Services
- IPNG - IP: Next Generation
- MGT - Network Management
- OPS - Operational Requirements
- RTG - Routing
- SEC - Security
- TSV - Transport
- USV - User Services

## Acronyms

- ANSI - American National Standards Institute
- ARPA - Advanced Research Projects Agency
- ARPANET - Advanced Research Projects Agency Network
- AS - Autonomous System
- ATM - Asynchronous Transfer Mode
- BGP - Border Gateway Protocol
- BOF - Birds Of a Feather
- BSD - Berkeley Software Distribution
- BTW - By The Way
- CCIRN - Coordinating Committee for Intercontinental Research Networks
- CCITT - International Telegraph and Telephone Consultative Committee
- CIDR - Classless Inter-Domain Routing
- CIX - Commercial Information Exchange
- CNI - Coalition for Networked Information
- CREN - The Corporation for Research and Educational Networking
- DARPA - US Defense Advanced Research Projects Agency (now ARPA)
- DDN - US Defense Data Network
- DISA - US Defense Information Systems Agency
- EGP - Exterior Gateway Protocol
- FAQ - Frequently Asked Question

- FARNET - Federation of American Research NETworks
- FIX - US Federal Information Exchange
- FNC - US Federal Networking Council
- FQDN- Fully Qualified Domain Name
- FYI - For Your Information (RFC)
- GOSIP- US Government OSI Profile
- IAB - Internet Architecture Board
- IANA - Internet Assigned Numbers Authority
- I-D - Internet-Draft
- IEN - Internet Experiment Note
- IESG - Internet Engineering Steering Group
- IETF - Internet Engineering Task Force
- IGP - Interior Gateway Protocol
- IMHO - In My Humble Opinion
- IMR - Internet Monthly Report
- InterNIC - Internet Network Information Center
- IPng - IP: Next Generation
- IR - Internet Registry
- IRSG - Internet Research Steering Group
- IRTF - Internet Research Task Force
- ISO - International Organization for Standardization
- ISOC - Internet Society
- ISODE - ISO Development Environment
- ITU - International Telecommunication Union
- MIB - Management Information Base
- MIME - Multipurpose Internet Mail Extensions
- NIC - Network Information Center
- NIS - Network Information Services
- NIST - National Institute of Standards and Technology
- NOC - Network Operations Center
- NREN - National Research and Education Network
- NSF - National Science Foundation
- OSI - Open Systems Interconnection
- PEM - Privacy Enhanced Mail
- PTT - Postal, Telegraph and Telephone
- RARE - Reseaux Associes pour la Recherche Europeenne (no longer exists)
- RFC - Request For Comments
- RIPE - Reseaux IP Europeenne
- SIG - Special Interest Group
- STD - Standard (RFC)
- TERENA - Trans European Research & Education Networking Association
- TLA - Three Letter Acronym
- TTFN - Ta-Ta For Now
- UTC - Universal Time Coordinated
- WG - Working Group
- WRT - With Respect To
- WYSIWYG - What You See is What You Get

## Acknowledgments

The IETF Secretariat would like to acknowledge the time and efforts of Gary Malkin, who prepared the original RFC from which this material has been excerpted and who coordinated the changes to the first revision. Without his help, this document might still be "in progress."

## How to get RFCs

```
                    Where and how to get new RFCs
                    =============================
RFCs may be obtained via EMAIL or FTP from many RFC Repositories.  The
Primary Repositories will have the RFC available when it is first
announced, as will many Secondary Repositories.  Some Secondary
Repositories may take a few days to make available the most recent
RFCs.
Many of these repositories also now have World Wide Web servers. Try
the following URL as a starting point:
              http://www.isi.edu/rfc-editor/

Primary Repositories:
RFCs can be obtained via FTP from NIS.NSF.NET, NISC.JVNC.NET,
FTP.ISI.EDU, WUARCHIVE.WUSTL.EDU, SRC.DOC.IC.AC.UK, FTP.NCREN.NET,
FTP.SESQUI.NET, FTP.NIC.IT, or FTP.IMAG.FR.
1.  NIS.NSF.NET
To obtain RFCs from NIS.NSF.NET via FTP, login with username
"anonymous" and password "name@host.domain"; then connect to the
directory of RFCs with cd /internet/documents/rfc.  The file name is
of the form rfcnnnn.txt (where "nnnn" refers to the RFC number).
For sites without FTP capability, electronic mail query is available
from NIS.NSF.NET.  Address the request to NIS-INFO@NIS.NSF.NET and
leave the subject field of the message blank.  The first text line of
the message must be "send rfcnnnn.txt" with nnnn the RFC number.
contact:  rfc-mgr@merit.edu
…
```

## Standard ‘boilerplate’ for Internet Drafts:

```
This document is an Internet-Draft. Internet-Drafts are working
documents of the Internet Engineering Task Force (IETF), its areas,
and its working groups.  Note that other groups may also distribute
working documents as Internet-Drafts.
Internet-Drafts are draft documents valid for a maximum of six months
and may be updated, replaced, or obsoleted by other documents at any
time.  It is inappropriate to use Internet-Drafts as reference
material or to cite them other than as "work in progress."
To learn the current status of any Internet-Draft, please check the
"1id-abstracts.txt" listing contained in the Internet-Drafts Shadow
Directories on ftp.is.co.za (Africa), ftp.nordu.net (Europe),
munnari.oz.au (Pacific Rim), ds.internic.net (US East Coast), or
ftp.isi.edu (US West Coast).
```

World Wide Web Consortium: http://www.w3.org

# About The World Wide Web Consortium

## *Leading the Web to its Full Potential...*

[Membership - Services - Process - Further information]
The W3C was founded in October 1994 to lead the World Wide Web to its full potential by developing common protocols that promote its evolution and ensure its interoperability. We are an international industry consortium, jointly hosted by the Massachusetts Institute of Technology Laboratory for Computer Science [MIT/LCS] in the United States; the Institut National de Recherche en Informatique et en Automatique [INRIA] in Europe; and the Keio University Shonan Fujisawa Campus in Japan. Services provided by the Consortium include: a repository of information about the World Wide Web for developers and users; reference code implementations to embody and promote standards; and various prototype and sample applications to demonstrate use of new technology. Initially, the W3C was established in collaboration with CERN, where the Web originated, with support from DARPA and the European Commission. For details on the joint initiative and the contributions of CERN, INRIA, and MIT, please see the statement on the joint World Wide Web Initiative.
The Consortium is led by Tim Berners-Lee, Director and creator of the World Wide Web, and Jean-François Abramatic, Chairman. W3C is funded by Member organizations, and is vendor neutral, working with the global community to produce specifications and reference software that is made freely available throughout the world.

## W3C Membership

**Membership is open to any organization which signs a membership agreement**.
If your organization would like to become a member of the W3C, please see

- **How to Join W3C**
- **The World Wide Web Consortium: Prospectus.**

We are sorry that the W3C cannot take individual membership: those interested in W3C activities are encouraged to subscribe to the World Wide Web Journal, the official Journal of W3C.
To learn about W3C Member organizations and visit their World Wide Web sites, see the Member list.

## W3C Services

For members, the Consortium provides a place to meet, to discus and to reach agreement on common specifications. W3C staff help by

- Organizing meetings of various forms;
- Facilitating discussion and arrival at consensus;
- Providing expert editing and writing skills;
- Helping to plan a architecturally consistent and timely way forward.

The Consortium provides a number of public services:

- A repository of information about the World Wide Web for developers and users, especially specifications about the Web;
- A sample code implementation to embody and promote standards
- Various prototype and sample applications to demonstrate use of new technology

All products of the Consortium are available during development and on initial release to Members. One month after formal internal release, all software produced by or officially contributed to the W3C is available for general public use, commercial or otherwise.

## W3C Process

The Consortium attempts to find common specifications for the Web so that through dramatic and rapid evolution, many organizations can work in their own fields to exploit and build on top of the global information space which is the web. The technologies involved in the web are changing very rapidly, and so the Consortium must have both efficiency and flexibility in its process, to be able to respond to the needs of the community in a timely manner. At the same time, it must be clear that the Consortium is neutral forum, and no member has a priori a greater say than another.

# Further information

- W3C Backgrounder
- Help wanted: come and work with the W3C Team at MIT or INRIA
- People of the W3C
- Frequently Asked Questions about W3C
- How to contact W3C
- Press Information

---

*TimBL*
*Webmaster*
*$Date: 1998/02/05 17:31:59 $*

## Microsoft on Standards

### From http://www.microsoft.com/standards/default.asp

Last updated: April 6, 1998
**New!** CDF spec update and Namespaces in XML spec
There are a number of interesting sites and specifications on the Web maintained by various Internet organizations. We've listed key resources below. We've also expanded this area to provide standards documents that are relevant


to Microsoft® products, including drafts of specifications for new technologies such as CDF.
Microsoft is committed to working with the World Wide Web Consortium (W3C) to implement W3C-approved HTML standards, and has confirmed its pledge to work through W3C and other standards bodies on enhancements to HTML and other key Web technologies.
Before you explore the links below, you'll want to read Standard Bearers: How the W3C Sets Benchmarks for the Web, which explains who sets standards for the World Wide Web, what the best known standards are, and where to find more detailed information about them.

(Please note that the links marked with below point to servers that are not under Microsoft's control. See Microsoft's official statement egarding other servers.)

## Netscape on standards

**From http://search.netscape.com/comprod/columns/intranet/open_standards.html)**

**June 11, 1997** - This open letter to Netscape's enterprise customers, industry partners, and developers outlines Netscape's approach to open Internet standards. At Netscape, support for open standards is at the core of our company's philosophy. We believe that open standards give our customers the widest range of choices and provide a high degree of interoperability between products from different vendors. In fact, to emphasize how strongly we feel about the importance of embracing Internet standards, we have established the Netscape Open Standards Guarantee. The guarantee is simple:
We will adopt the major open standards recommended by the Internet Engineering Task Force (IETF) and other appropriate, vendor-neutral standards bodies, such as the World Wide Web Consortium (W3C), Object Management Group (OMG), ECMA, and others. We will also assume a leadership role in proposing new standards when there is real customer demand and where such standards do not yet exist.
Because interoperability is so critical for the continued growth of the Internet and corporate intranets, the Netscape Open Standards Guarantee applies to all of our products. For example, look at Dynamic HTML, which is actually an umbrella term for a number of different, yet interrelated, proposals at various stages in the W3C review cycle. Web designers have been telling us for the past year that they need a more precise way to position HTML objects on a page, so we proposed an extension to HTML that helps developers create more compelling Web pages. While Netscape and Microsoft originally offered different visions of how to solve these design problems, we have been working through the W3C over the past six months to bring these different proposals into alignment. Since the first preview release of Netscape Communicator in December 1996, Netscape has adopted the W3C recommendation for Cascading Style Sheets (CSS1), an element of Dynamic HTML that Microsoft has also pledged to support.

## Sun on Java standards

**From http://java.sun.com/aboutJava/standardization/index.html**

### How we develop our Java Platform specifications

The PAS guidelines require that potential submitters must develop and maintain their specifications using a cooperative process that allows for broad participation by interested parties. The resulting specifications must be readily available to the public and must not impose any unreasonable restrictions to prevent people from producing their own implementations of the technology that is described.
Evolutionary in nature, our process builds upon the specification development processes already in use by leading-edge industry consortia. Like the consortia processes, no single company (not even JavaSoft) can dominate our

development process. Unlike most consortia, however, we don't charge fees to participate - anyone with access to the Internet can review and contribute to our Java Platform specifications.
We are in the unique position of being a single company that develops and evolves its specifications using an open and inclusive process that produces high-quality specifications in "internet-time". The resulting specifications have typically received broad industry endorsement and implementations have been quickly produced and used industry-wide.

## Oracle on Standards

From http://www.oracle.com/corporate/press/html/standards.html

### ORACLE CHALLENGES INDUSTRY TO MAINTAIN OPEN INTERNET STANDARDS, SYSTEMS

"Open" Web Strategy Includes Browser/Server Agnosticism, Browser Extensibility, Published APIs, Standards Compliance

BOSTON, Oct. 30, 1995-With today's introduction of the Oracle WebSystem for building client/server Web applications, Oracle Corp. is sending an unmistakable message to the computer industry about the information superhighway: Keep it open.
Recent concerns expressed by users, industry analysts and the press warn that some popular I-way products would stray from the Internet's open foundations, as software vendors attempt to secure market share by creating proprietary and incompatible standards. To counter that trend, Oracle WebSystem is designed as an open environment for developing Web applications. This strategy includes interoperability with third-party browser and server software, unlimited browser extensibility by downloading application objects on the fly from enhanced Web sites, compliance with existing standards, and submission of two new Oracle specifications to the leading Internet standards bodies.
"The Internet's commercial and social potential will only be realized when Web browsers work with Web servers to provide open, universal access to massive amounts of information," said Marc Benioff, senior vice president of Oracle's Web/Workgroup Systems Division. "Open systems have been the foundation for both the Internet and Oracle's success."

## From an AltaVista search

### 14. Major New Players Join Effort to Finalize Open Standards for Internet Commerce

3) || (browserName == "Microsoft Internet Explorer" && browserVer >= 4)) version = "n3"; else version = "n2"; if (version == "n3") { hotdealson = new...

# MIME Media types (RFC 2046)

The Content-Type field is used to specify the nature of the data in the body of a MIME entity, by giving media type and subtype identifiers, and by providing auxiliary information that may be required for certain media types. After the type and subtype names, the remainder of the header field is simply a set of parameters, specified in an attribute/value notation. The ordering of parameters is not significant.
In general, the top-level media type is used to declare the general type of data, while the subtype specifies a specific format for that type of data. Thus, a media type of "image/xyz" is enough to tell a user agent that the data is an image, even if the user agent has no knowledge of the specific image format "xyz". Such information can be used, for example, to decide whether or not to show a user the raw data from an unrecognized subtype -- such an action might be reasonable for unrecognized subtypes of "text", but not for unrecognized subtypes of "image" or "audio". For this reason, registered subtypes of "text", "image", "audio", and "video" should not contain embedded information that is really of a different type. Such compound formats should be represented using the "multipart" or "application" types.
Parameters are modifiers of the media subtype, and as such do not fundamentally affect the nature of the content. The set of meaningful parameters depends on the media type and subtype. Most parameters are associated with a single specific subtype. However, a given top-level media type may define parameters which are applicable to any subtype of that type. Parameters may be required by their defining media type or subtype or they may be optional. MIME implementations must also ignore any parameters whose names they do not recognize.
MIME's Content-Type header field and media type mechanism has been carefully designed to be extensible, and it is expected that the set of media type/subtype pairs and their associated parameters will grow significantly over time. Several other MIME facilities, such as transfer encodings and "message/external-body" access types, are likely to have new values defined over time. In order to ensure that the set of such values is developed in an orderly, well-specified, and public manner, MIME sets up a registration process which uses the Internet Assigned Numbers Authority (IANA) as a central registry for MIME's various areas of extensibility. The registration process for these areas is described in a companion document, RFC 2048.
The initial seven standard top-level media type are defined and described in the remainder of this document.

## Definition of a Top-Level Media Type

The definition of a top-level media type consists of:

(1) a name and a description of the type, including criteria for whether a particular type would qualify under that type,

(2) the names and definitions of parameters, if any, which are defined for all subtypes of that type (including whether such parameters are required or optional),

(3) how a user agent and/or gateway should handle unknown subtypes of this type,

(4) general considerations on gatewaying entities of this top-level type, if any, and

(5) any restrictions on content-transfer-encodings for entities of this top-level type.

## Overview Of The Initial Top-Level Media Types

The five discrete top-level media types are:

(1) text -- textual information. The subtype "plain" in particular indicates plain text containing no formatting commands or directives of any sort. Plain text is intended to be displayed "as-is". No special software is required to get the full meaning of the text, aside from support for the indicated character set. Other subtypes are to be used for enriched text in forms where application software may enhance the appearance of the text, but such software must not be required in order to get the general idea of the content. Possible subtypes of "text" thus include any word processor format that can be read without resorting to software that understands the format. In particular, formats that employ embeddded binary formatting information are not considered

directly readable. A very simple and portable subtype, “richtext”, was defined in RFC 1341, with a further revision in RFC 1896 under the name “enriched”.

(2) image -- image data. “Image” requires a display device (such as a graphical display, a graphics printer, or a FAX machine) to view the information. An initial subtype is defined for the widely-used image format JPEG. . subtypes are defined for two widely-used image formats, jpeg and gif.

(3) audio -- audio data. “Audio” requires an audio output device (such as a speaker or a telephone) to “display” the contents. An initial subtype “basic” is defined in this document.

(4) video -- video data. “Video” requires the capability to display moving images, typically including specialized hardware and software. An initial subtype “mpeg” is defined in this document.

(5) application -- some other kind of data, typically either uninterpreted binary data or information to be processed by an application. The subtype “octet-stream” is to be used in the case of uninterpreted binary data, in which case the simplest recommended action is to offer to write the information into a file for the user. The “PostScript” subtype is also defined for the transport of PostScript material. Other expected uses for “application” include spreadsheets, data for mail-based scheduling systems, and languages for “active” (computational) messaging, and word processing formats that are not directly readable. Note that security considerations may exist for some types of application data, most notably “application/PostScript” and any form of active messaging. These issues are discussed later in this document.

The two composite top-level media types are:

(1) multipart -- data consisting of multiple entities of independent data types. Four subtypes are initially defined, including the basic “mixed” subtype specifying a generic mixed set of parts, “alternative” for representing the same data in multiple formats, “parallel” for parts intended to be viewed simultaneously, and “digest” for multipart entities in which each part has a default type of “message/rfc822”.

(2) message -- an encapsulated message. A body of media type “message” is itself all or a portion of some kind of message object. Such objects may or may not in turn contain other entities. The “rfc822” subtype is used when the encapsulated content is itself an RFC 822 message. The “partial” subtype is defined for partial RFC 822 messages, to permit the fragmented transmission of bodies that are thought to be too large to be passed through transport facilities in one piece. Another subtype, “external-body”, is defined for specifying large bodies by reference to an external data source.

It should be noted that the list of media type values given here may be augmented in time, via the mechanisms described above, and that the set of subtypes is expected to grow substantially.

# Internet Assigned Numbers Authority http://www.isi.edu/div7/iana/

## Overview

The Internet Assigned Numbers Authority (IANA) is the central coordinator for the assignment of unique parameter values for Internet protocols.
The IANA is chartered by the Internet Society (ISOC) and the Federal Network Council (FNC) to act as the clearinghouse to assign and coordinate the use of numerous Internet protocol parameters.
The Internet protocol suite, as defined by the Internet Engineering Task Force (IETF) and its steering group (the IESG), contains numerous parameters, such as internet addresses, domain names, autonomous system numbers (used in some routing protocols), protocol numbers, port numbers, management information base object identifiers, including private enterprise numbers, and many others.
The common use of the Internet protocols by the Internet community requires that the particular values used in these parameter fields be assigned uniquely. It is the task of the IANA to make those unique assignments as requested and to maintain a registry of the currently assigned values.
Please send requests for parameter assignments (protocols, ports, etc) by mail to iana@iana.org.
Please send requests for SNMP network management private enterprise number assignments by mail to iana-mib@iana.org.
The most recent summary of these assigned parameter values is "Assigned Numbers" which is STD-2 and RFC-1700 published in October 1994. The assignments may also be found online.
The IANA is located at and operated by the Information Sciences Institute (ISI) of the University of Southern California (USC).

### Assignments from ftp://ftp.isi.edu/in-notes/iana/assignments/media-types

```
MEDIA TYPES

[RFC1521] specifies that Content Types, Content Subtypes, Character
Sets, Access Types, and Conversion values for MIME mail will be
assigned and listed by the IANA.


Content Types and Subtypes
--------------------------

Type            Subtype         Description                 Reference
----            -------         -----------                 ---------
text            plain                                 [RFC1521,Borenstein]
                richtext                              [RFC1521,Borenstein]
                enriched                                        [RFC1896]
                tab-separated-values                      [Paul Lindner]
                html                                            [RFC1866]
                sgml                                            [RFC1874]
                vnd.latex-z                                       [Lubos]
                vnd.fmi.flexstor                                 [Hurtta]
                uri-list                                         [Daniel]
                vnd.abc                                           [Allen]
                rfc822-headers                                  [RFC1892]
                vnd.in3d.3dml                                    [Powers]
                prs.lines.tag                                     [Lines]
                vnd.in3d.spot                                    [Powers]
                css                                             [RFC2318]

multipart       mixed                                 [RFC1521,Borenstein]
                alternative                           [RFC1521,Borenstein]
                digest                                [RFC1521,Borenstein]
                parallel                              [RFC1521,Borenstein]
                appledouble                 [MacMime,Patrik Faltstrom]
                header-set                                [Dave Crocker]
                form-data                                       [RFC1867]
                related                                         [RFC2112]
```

```
                report                                          [RFC1892]
                voice-message                                   [RFC1911]
                signed                                          [RFC1847]
                encrypted                                       [RFC1847]
                byteranges                                      [RFC2068]

message         rfc822                                [RFC1521,Borenstein]
                partial                               [RFC1521,Borenstein]
                external-body                         [RFC1521,Borenstein]
                news                            [RFC 1036, Henry Spencer]
                http                                            [RFC2068]
                delivery-status                                 [RFC1894]
                disposition-notification                        [RFC2298]

application     octet-stream                          [RFC1521,Borenstein]
                postscript                            [RFC1521,Borenstein]
                oda                                   [RFC1521,Borenstein]
                atomicmail                      [atomicmail,Borenstein]
                andrew-inset                   [andrew-inset,Borenstein]
                slate                             [slate,terry crowley]
                wita                 [Wang Info Transfer,Larry Campbell]
                dec-dx               [Digital Doc Trans, Larry Campbell]
                dca-rft          [IBM Doc Content Arch, Larry Campbell]
                activemessage                                 [Ehud Shapiro]
                rtf                                           [Paul Lindner]
                applefile                    [MacMime,Patrik Faltstrom]
                mac-binhex40                 [MacMime,Patrik Faltstrom]
                news-message-id                [RFC1036, Henry Spencer]
                news-transmission              [RFC1036, Henry Spencer]
                wordperfect5.1                                [Paul Lindner]
                pdf                                           [Paul Lindner]
                zip                                           [Paul Lindner]
                macwriteii                                    [Paul Lindner]
                msword                                        [Paul Lindner]
                remote-printing                               [RFC1486,Rose]
                mathematica                                    [Van Nostern]
                cybercash                                         [Eastlake]
                commonground                                        [Glazer]
                iges                                                [Parks]
                riscos                                              [Smith]
                eshop                                                [Katz]
                x400-bp                                          [RFC1494]
                sgml                                             [RFC1874]
                cals-1840                                        [RFC1895]
                pgp-encrypted                                    [RFC2015]
                pgp-signature                                    [RFC2015]
                pgp-keys                                         [RFC2015]
                vnd.framemaker                                     [Wexler]
                vnd.mif                                            [Wexler]
                vnd.ms-excel                                         [Gill]
                vnd.ms-powerpoint                                    [Gill]
                vnd.ms-project                                       [Gill]
                vnd.ms-works                                         [Gill]
                vnd.ms-tnef                                          [Gill]
                vnd.svd                                            [Becker]
                vnd.music-niff                                     [Butler]
                vnd.ms-artgalry                                   [Slawson]
                vnd.truedoc                                          [Chase]
                vnd.koan                                             [Cole]
                vnd.street-stream                                  [Levitt]
                vnd.fdf                                            [Zilles]
                set-payment-initiation                             [Korver]
                set-payment                                        [Korver]
```

```
set-registration-initiation                [Korver]
set-registration                           [Korver]
vnd.seemail                                  [Webb]
vnd.businessobjects                        [Imoucha]
vnd.meridian-slingshot                       [Wedel]
vnd.xara                                 [Matthewman]
sgml-open-catalog                            [Grosso]
vnd.rapid                                  [Szekely]
vnd.enliven                              [Santinelli]
vnd.japannet-registration-wakeup              [Fujii]
vnd.japannet-verification-wakeup              [Fujii]
vnd.japannet-payment-wakeup                   [Fujii]
vnd.japannet-directory-service                [Fujii]
vnd.intertrust.digibox                     [Tomasello]
vnd.intertrust.nncp                        [Tomasello]
prs.alvestrand.titrax-sheet              [Alvestrand]
vnd.noblenet-web                           [Solomon]
vnd.noblenet-sealer                        [Solomon]
vnd.noblenet-directory                     [Solomon]
prs.nprend                                 [Doggett]
vnd.webturbo                                 [Rehem]
hyperstudio                                 [Domino]
vnd.shana.informed.formtemplate            [Selzler]
vnd.shana.informed.formdata                [Selzler]
vnd.shana.informed.package                 [Selzler]
vnd.shana.informed.interchange             [Selzler]
vnd.$commerce_battelle                    [Applebaum]
vnd.osa.netdeploy                             [Klos]
vnd.ibm.MiniPay                            [Herzberg]
vnd.japannet-jpnstore-wakeup               [Yoshitake]
vnd.japannet-setstore-wakeup               [Yoshitake]
vnd.japannet-verification                  [Yoshitake]
vnd.japannet-registration                  [Yoshitake]
vnd.hp-HPGL                                [Pentecost]
vnd.hp-PCL                                 [Pentecost]
vnd.hp-PCLXL                               [Pentecost]
vnd.musician                                  [Adams]
vnd.FloGraphIt                              [Floersch]
vnd.intercon.formnet                          [Gurak]
vemmi                                       [RFC2122]
vnd.ms-asf                                [Fleischman]
vnd.ecdis-update                        [Buettgenbach]
vnd.powerbuilder6                                [Guy]
vnd.powerbuilder6-s                              [Guy]
vnd.lotus-wordpro                       [Wattenberger]
vnd.lotus-approach                      [Wattenberger]
vnd.lotus-1-2-3                         [Wattenberger]
vnd.lotus-organizer                     [Wattenberger]
vnd.lotus-screencam                     [Wattenberger]
vnd.lotus-freelance                     [Wattenberger]
vnd.fujitsu.oasys                             [Togashi]
vnd.fujitsu.oasys2                            [Togashi]
vnd.swiftview-ics                             [Widener]
vnd.dna                                        [Searcy]
prs.cww                                 [Rungchavalnont]
vnd.wt.stf                                    [Wohler]
vnd.dxr                                        [Duffy]
vnd.mitsubishi.misty-guard.trustweb           [Tanaka]
vnd.ibm.modcap                              [Hohensee]
vnd.acucobol                                  [Lubin]
vnd.fujitsu.oasys3                           [Okudaira]
marc                                        [RFC2220]
vnd.fujitsu.oasysprs                          [Ogita]
```

```
                vnd.fujitsu.oasysgp                          [Sugitomo]
                vnd.visio                                      [Sandal]
                vnd.netfpx                                       [Mutz]
                vnd.audiograph                               [Slusanschi]
                vnd.epson.salt                                 [Nagatomo]
                vnd.3M.Post-it-Notes                            [O'Brien]
                vnd.novadigm.EDX                                [Swenson]
                vnd.novadigm.EXT                                [Swenson]
                vnd.novadigm.EDM                                [Swenson]
                vnd.claymore                                    [Simpson]
                vnd.comsocaller                                [Dellutri]
                pkcs7-mime                                      [RFC2311]
                pkcs7-signature                                 [RFC2311]
                pkcs10                                          [RFC2311]
                vnd.yellowriver-custom-menu                      [Yellow]
                vnd.ecowin.chart                                [Olsson]
                vnd.ecowin.series                               [Olsson]
                vnd.ecowin.filerequest                          [Olsson]
                vnd.ecowin.fileupdate                           [Olsson]
                vnd.ecowin.seriesrequest                        [Olsson]
                vnd.ecowin.seriesupdate                         [Olsson]

image           jpeg                                 [RFC1521,Borenstein]
                gif                                  [RFC1521,Borenstein]
                ief             Image Exchange Format            [RFC1314]
                g3fax                                            [RFC1494]
                tiff            Tag Image File Format            [RFC2301]
                cgm             Computer Graphics Metafile       [Francis]
                naplps                                           [Ferber]
                vnd.dwg                                          [Moline]
                vnd.svf                                          [Moline]
                vnd.dxf                                          [Moline]
                png                                    [Randers-Pehrson]
                vnd.fpx                                         [Spencer]
                vnd.net-fpx                                     [Spencer]
                vnd.xiff                                         [Martin]

audio           basic                                [RFC1521,Borenstein]
                32kadpcm                                         [RFC1911]
                vnd.qcelp                                      [Lundblade]

video           mpeg                                 [RFC1521,Borenstein]
                quicktime                                 [Paul Lindner]
                vnd.vivo                                           [Wolfe]
                vnd.motorola.video                              [McGinty]
                vnd.motorola.videop                             [McGinty]

model                                                            [RFC2077]
                iges                                              [Parks]
                vrml                                            [RFC2077]
                mesh                                            [RFC2077]
                vnd.dwf                                           [Pratt]
```

The "media-types" directory contains a subdirectory for each content type and each of those directories contains a file for each content subtype.

## Some Web-related RFCs

### These documents are interesting for their historical perspective.

**1689** *A Status Report on Networked Information Retrieval: Tools and Groups*. J. Foster. August 1994. (Also FYI0025, RTR0013) ***(Informational)***

**1935** *What is the Internet, Anyway?.* J. Quarterman & S. Carl-Mitchell. April 1996. *Informational*

### IETF Process:

#### The IETF process itself is documented in RFCs, and goes through a similar consensus and vetting process:

**1396** The Process for Organization of Internet Standards Working Group (POISED). S. Crocker. January 1993. *(Informational)*

**1603** *IETF Working Group Guidelines and Procedures.* E. Huizer & D. Crocker. March 1994. *(Informational)*

**1640** T*he Process for Organization of Internet Standards Working Group (POISED).* S. Crocker. June 1994. *(Informational)*

**1718** *The Tao of IETF - A Guide for New Attendees of the Internet Engineering Task Force.* The IETF Secretariat & G. Malkin. November 1994. (Also FYI0017) *(Informational)*

**1796** *Not All RFCs are Standards*. C. Huitema, J. Postel & S. Crocker. April 1995. *(Informational)*

**1984** *IAB and IESG Statement on Cryptographic Technology and the Internet.* IAB & IESG. August 1996. *(Informational)*

**2026** *The Internet Standards Process -- Revision 3.* S. Bradner. October 1996. (Also BCP0009) *(Best Current Practice)*

**2028** *The Organizations Involved in the IETF Standards Process*. R. Hovey, S. Bradner. October 1996. (Also BCP0011) *(Best Current Practice)*

**2119** *Key words for use in RFCs to Indicate Requirement Levels.* S. Bradner. March 1997. (Also BCP0014) *(Best Current Practice).*

**2282** *IAB and IESG Selection, Confirmation, and Recall Process: Operation of the Nominating and Recall Committees*. J. Galvin. February 1998. (Also BCP0010) *(Best Current Practice)*

### URLs and related RFCs on naming:

**1630** *Universal Resource Identifiers in WWW: A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network as used in the World-Wide Web*. T. Berners-Lee. June 1994. *(Informational)*

**1736** *Functional Recommendations for Internet Resource Locators*. J. Kunze. February 1995. *(Informational)*

Requirements for “what problem are we solving?”

**1737** *Functional Requirements for Uniform Resource Names*. K. Sollins & L. Masinter. December 1994. *(Informational)*

**2276** *Architectural Principles of Uniform Resource Name Resolution*. K. Sollins. January 1998. *(Informational)*

**1738** Uniform Resource Locators (URL). T. Berners-Lee, L. Masinter & M. McCahill. December 1994. *(Proposed Standard)*

**1808** *Relative Uniform Resource Locators*. R. Fielding. June 1995. *(Proposed Standard)*

**2141** *URN Syntax*. R. Moats. May 1997. *(Proposed Standard)*

**2056** *Uniform Resource Locators for Z39.50.* R. Denenberg, J. Kunze, D. Lynch. November 1996. *(Proposed Standard)*

**2288** *Using Existing Bibliographic Identifiers as Uniform Resource Names.* C. Lynch, C. Preston, R. Daniel. February 1998. *(Informational)*

**2168** *Resolution of Uniform Resource Identifiers using the Domain Name System*. R. Danie1, M. Mealling. June 1997. (Status: EXPERIMENTAL)

**2169** *A Trivial Convention for using HTTP in URN Resolution.* R. Danie1. June 1997. (Status: EXPERIMENTAL)

**2240** *A Legal Basis for Domain Name Allocation.* O. Vaughan. November 1997. *(Informational)*

## MIME RFCs on representing and registering media types

**2045** *Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies.* N. Freed & N. Borenstein. November 1996. (Updated by RFC2184, RFC2231) (Status: DRAFT STANDARD)

**2231** *MIME Parameter Value and Encoded Word Extensions: Character Sets, Languages, and Continuations*. N. Freed, K. Moore. November 1997. (Updates RFC2045, RFC2047 RFC2183) *(Proposed Standard)*

**2046** *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*. N. Freed & N. Borenstein. November 1996. (Status: DRAFT STANDARD)

**2048** *Multipurpose Internet Mail Extensions (MIME) Part Four: Registration Procedures*. N. Freed, J. Klensin & J. Postel. November 1996. (Also BCP0013) *(Best Current Practice)*

## HTML:

*Work on HTML moved from IETF to W3C in 1996. These RFCs should be marked as 'HISTORICAL' soon.*

**1866** *Hypertext Markup Language - 2.0.* T. Berners-Lee & D. Connolly. November 1995. *(Proposed Standard)*

**1942** *HTML Tables*. D. Raggett. May 1996. (Status: EXPERIMENTAL)

**1867** *Form-based File Upload in HTML.* E. Nebel & L. Masinter. November 1995. (Status: EXPERIMENTAL)

**1980** *A Proposed Extension to HTML : Client-Side Image Maps.* J. Seidman. August 1996. *(Informational)*

**2070** *Internationalization of the Hypertext Markup Language*. F. Yergeau, G. Nicol, G. Adams, M. Duerst. January 1997. *(Proposed Standard)*

**2318** *The text/css Media Type.* H. Lie, B. Bos, C. Lilley. March 1998. *(Informational)*

## Other web-related content:

**2083** *PNG (Portable Network Graphics) Specification*. T. Boutell. January 1997. *(Informational)*

**1874** *SGML Media Types.* E. Levinson. December 1995. (Status: EXPERIMENTAL)

**1927** *Suggested Additional MIME Types for Associating Documents.* C. Rogers. April 1996. *(Informational)*

**2302** *Tag Image File Format (TIFF) - image/tiff MIME Sub-type Registration.* G. Parsons, J. Rafferty, S. Zilles. March 1998. *(Proposed Standard)*

## Managing web servers using Simple Network Management Protocol:

**2039** *Applicability of Standards Track MIBs to Management of World Wide Web Servers*. C. Kalbfleisch. November 1996. *(Informational)*

## Protocol:

*Work on HTTP is ongoing, with a revision of HTTP/1.1 expected soon.*

**1945** *Hypertext Transfer Protocol -- HTTP/1.0.* T. Berners-Lee, R. Fielding & H. Frystyk. May 1996. *(Informational)*

**2068** *Hypertext Transfer Protocol -- HTTP/1.1.* R. Fielding, J. Gettys, J. Mogul, H. Frystyk, T. Berners-Lee. January 1997. *(Proposed Standard)*

**2069** *An Extension to HTTP : Digest Access Authentication*. J. Franks, P. Hallam-Baker, J. Hostetler, P. Leach, A. Luotonen, E. Sink, L. Stewart. January 1997. *(Proposed Standard)*

**2109** *HTTP State Management Mechanism*. D. Kristol, L. Montulli. February 1997. *(Proposed Standard)*

**2295** *Transparent Content Negotiation in HTTP.* K. Holtman, A. Mutz. March 1998. (Status: EXPERIMENTAL)

**2296** *HTTP Remote Variant Selection Algorithm* -- RVSA/1.0. K. Holtman, A. Mutz. March 1998. (Status: EXPERIMENTAL)

**2145** *Use and Interpretation of HTTP Version Numbers*. J. C. Mogul, R. Fielding, J. Gettys, H. Frystyk. May 1997. *(Informational)*

**2227** *Simple Hit-Metering and Usage-Limiting for HTTP*. J. Mogul, P. Leach. October 1997. *(Proposed Standard)*

**2324** *Hyper Text Coffee Pot Control Protocol (HTCPCP/1.0).* L. Masinter. March 1998. *(Informational)*

## Alternatives can be published:

**2054** *WebNFS Client Specification*. B. Callaghan. October 1996. *(Informational)*

**2055** *WebNFS Server Specification*. B. Callaghan. October 1996. *(Informational)*

## Security:

**2084** *Considerations for Web Transaction Security*. G. Bossert, S. Cooper, W. Drummond. January 1997. *(Informational)*

## Sending HTML in mail:

**2110** *MIME E-mail Encapsulation of Aggregate Documents, such as HTML (MHTML).* J. Palme, A. Hopmann. March 1997. *(Proposed Standard)*

**2111** *Content-ID and Message-ID Uniform Resource Locators.* E. Levinson. February 1997. *(Proposed Standard)*

**2112** *The MIME Multipart/Related Content-type.* E. Levinson. February 1997. *(Proposed Standard)*

## Authoring:

**2291** *Requirements for a Distributed Authoring and Versioning Protocol for the World Wide Web.* J. Slein, F. Vitali, E. Whitehead, D. Durand. February 1998. *(Informational)*

## Character sets:

**2277** *IETF Policy on Character Sets and Languages*. H. Alvestrand. January 1998. (Also BCP0018) *(Best Current Practice)*

**2278** *IANA Charset Registration Procedures.* N. Freed, J. Postel. January 1998. (Also BCP0019) *(Best Current Practice)*

**2279** *UTF-8, a transformation format of ISO 10646.* F. Yergeau. January 1998. *(Proposed Standard)*

# Web-related IETF working groups

## Content Negotiation (conneg)

A number of Internet application protocols need to indicate recipient capabilities, characteristics, and preferences when the resources they handle can vary in form. This working group will finalize registration procedures for distinguishing attributes which cause the media delivered to vary in form. The registration of these "media features" will provide a supplement to the MIME registration of media types and enable the development of a cross-protocol vocabulary for exchanging information on recipient capabilities, characteristics, and preferences. Since these distinguishing attributes commonly occur in related sets, this working group will also describe at least one method for referring to composite media feature sets. Experimental methods for using these features and feature sets within specific protocol contexts may be developed within this group or within the groups standardizing the relevant protocols.

## HyperText Transfer Protocol (http)

Revising HTTP/1.1 to Draft Standard.

## Internet Printing Protocol (ipp)

Using HTTP as a transport layer.

## MIME Encapsulation of Aggregate HTML Documents (mhtml)

World Wide Web documents are most often written using Hyper Text Markup Language (HTML). HTML is notable in that it contains "embedded content"; that is, HTML documents often contain pointers or links to other objects (images, external references) which are to be presented to the recipient. Currently, these compound structured Web documents are transported almost exclusively via the interactive HTTP protocol. The MHTML working group has developed three Proposed Standards (RFCs 2110, 2111 and 2112) which permit the transport of such compound structured Web documents via Internet mail in MIME multipart/related body parts.

## Transaction Internet Protocol (tip)

The task of the TIP working group is to develop an Internet standard two-phase commit protocol specification, to enable heterogeneous Transaction Managers to agree on the outcome of a distributed transaction, based upon the Internet-Draft TIP protocol specification<DRAFT-LYON-ITP-NODES-01.TXT . [Note that since<DRAFT-LYON-ITP-NODES-01.TXT references a modified version of the Session Control Protocol (SCP), the TIP WG will also be responsible for progression of SCP to Proposed Internet Standard.]
In many applications where different nodes cooperate on some work, there is a need to guarantee that the work happens atomically. That is, each node must reach the same conclusion as to whether the work is to be completed (committed or aborted), even in the face of failures. This behaviour is achieved via the use of distributed transactions, employing a two-phase commit protocol (2-pc). The use of distributed transactions greatly simplifies distributed applications programming, since the number of possible outcomes is reduced from many to two, and failure recovery is performed automatically by the transaction service (Transaction Manager).
Key requirements to be met are, 1) the 2-pc protocol be independent of the application-to-application communications protocol, such that it may be used with any application protocol (especially HTTP), and 2) the 2-pc protocol be simple to implement and have a small working footprint (to encourage ubiquitous implementation and offer wide applicability).

## Uniform Resource Locator Registration Procedures (urlreg)

This working group exists for the purpose of creating two documents: The first document, a BCP RFC, will be the process for registering new URL schemes. The second document, an Informational RFC, will be a guideline for the creators of new URL schemes. The purpose of this guideline will be to help ensure that new URL schemes:
- Consistently implement the general syntax of URLs as specified in the URL Generic Syntax and Semantics RFC. - Are compatible with existing URL schemes. - Have clearly specified character encoding rules. - Have a well defined set of operations specified for them. - Properly address security considerations.
The following issues are considered beyond the scope of this working group and shall not be addressed by it:
- Modifications to the URL Generic Syntax and Semantics RFC. - Specific URL schemes, previously proposed or not, except as test cases for the guidelines document. - UR* schemes other than URLs.

## Uniform Resource Names (urn)

The goal of this working group is to define both a Uniform Resource Name (URN) framework and an initial set of components that fit this framework.
URNs are persistent identifiers for information resources. The output of this Working Group will comply with RFC 1737, which defines URNs and gives requirements for them. The framework will define the mechanics for enabling global scope, persistence, and legacy support requirements of URNs; requirements for namespaces to support this structure will also be defined. Although the framework will allow URNs to be defined that vary in terms of degree of scalability and persistance, ensuring "user friendliness" of all resultant identifiers is beyond the scope of this group.

## WWW Distributed Authoring and Versioning (webdav)

This working group will define the HTTP extensions necessary to enable distributed web authoring tools to be broadly interoperable, while supporting user needs.
The HTTP protocol contains functionality which enables the editing of web content at a remote location, without direct access to the storage media via an operating system. This capability is exploited by several existing HTML distributed authoring tools, and by a growing number of mainstream applications (e.g. word processors) which allow users to write (publish) their work to an HTTP server. To date, experience from the HTML authoring tools has shown they are unable to meet their user's needs using the facilities of the HTTP protocol. The consequence of this is either postponed introduction of distributed authoring capability, or the addition of nonstandard extensions to the HTTP protocol. These extensions, developed in isolation, are not interoperable.

## Transport Layer Security (tls)

Several methods of providing a secure and authenticated channel between hosts on the Internet above the transport layer have appeared. The objective of this proposed working group is to write standards track RFC(s) for protocols using the currently available Internet drafts as a basis. The SSL, PCT and SSH protocols are examples of mechanisms of establishing a secure channel for general purpose or special purpose Internet applications running over a reliable transport, usually TCP.
The TLS working group is a focused effort on providing security features at the transport layer, rather than general purpose security and key management mechanisms. The standard track protocol specification will provide methods for implementing privacy, authentication, and integrity above the transport layer.

## Web Transaction Security (wts)

The goal of the Web Transaction Security Working Group is to develop requirements and a specification for the provision of security services to Web transaction, e.g., transactions using HyperText Transport Protocol (HTTP). This work will proceed in parallel to and independently of the development of non-security features in the HTTP Working Group. The working group will prepare two documents for submission as Internet Drafts; an HTTP Security Requirements Specification, and an HTTP Security Protocol Specification. The latter will be submitted as a Standards Track RFC.

Web Security (from http://www.w3.org/Security)

# W3C Security Resources

# Introduction

Web security is a complex topic, encompassing computer system security, network security, authentication services, message validation, personal privacy issues, and cryptography. This page contains links to various aspects of Web and Internet security.

## Overview: The World Wide Web Security FAQ

The World Wide Web Security FAQ (Frequently Asked Questions with answers) provides an overview of Web security issues, security hole alerts, and practical advice for avoiding unpleasant surprises. It is recommended as a starting point for exploration.

# Security Initiatives at the W3C

The W3C is involved in the development of several protocols designed to enhance Web security. Among its activities are the Digital Signature Initiative, the Platform for Internet Content Selection (PICS), Platform for Privacy Preferences (P3P), and the HTTP/1.1 protocol. The W3C also produces software reference implementations that demonstrate the use of security measures.

## Digital Signature Initiative

The Digital Signature Initiative seeks to create a general standard for signing digital documents. Digital signatures are needed anywhere there is an issue of trust or commitment, including, but not limited to applets and other types of active content, political statements, press releases, contracts, and copyrighted material.

## PICS Initiative

The Platform for Internet Content Selection provides an ideology-neutral protocol for describing the contents of Internet documents. Although the first use of PICS has been to provide a way for browsers to filter pages on the basis of content, it is a general facility for labeling documents. It can be used to add many sorts of meta information to documents, including such things as copyright statements and usage information.

## Platform for Privacy Preferences

The Platform for Privacy Preferences (P3P) is a set of standards and protocols that provide end-users with a way of describing their preferences with regards to the use of personal information. Likewise, the P3P gives Web sites a standard way of describing their privacy policies. Software based on P3P will alert users when they encounter sites whose stated privacy policies do not match their personal preferences, and allow them to take action.

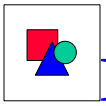
## HTTP/1.1

The HTTP/1.1 protocol includes a much improved scheme for authenticating the identity of users known as Digest Authentication.

## Electronic Commerce Initiatives

The W3C is involved in several initiatives in the realm of electronic commerce and secure payments. More information can be found in the Electronic Commerce Interest Group pages.

## HTML, from http://www.w3.org/Markup/



# HTML 4.0

HTML 4.0 is W3C's recommendation for the latest version of HTML. It includes support for style sheets, internationalization, accessibility to Web pages for people with disabilities, frames, richer tables and forms.

> "Insisting on HTML 4.0 compliance now will preserve your free choice of suppliers of Web software, tools and applications well into the future. With HTML 4.0, any Web application can be vendor independent. There really is no excuse for tying yourselves or your partners to proprietary solutions."
>
> -- Tim Berners-Lee, W3C Director and inventor of the World Wide Web

HTML 4.0 is specified in 3 flavors:

- **HTML 4.0 Strict**Use this when you want really clean markup, free of presentational clutter. Use this together with W3C's Cascading Style Sheet language (CSS) for great looking pages!
- **HTML 4.0 Transitional**: Use this when you need to take advantage of HTML's presentational features because many of your readers don't have the latest browsers that understand CSS.
- **HTML 4.0 Frameset** Use this when you want to use HTML Frames to partition the browser window into two or more frames.

…

# Guidelines for authoring

Don't use HTML purely as a visual formatting language for a single platform and make of browser! Remember that many people will be using different versions or makes of browsers. A large number of people have impairments of one form or another. Some will be browsing using Braille readers, or speech synthesisers. Others will use screen readers.
In particular, remember to include descriptions with each image, and try to avoid server-side image maps. For tables, you should include a summary of the tables structure and remember to associate table data with relevant headers. This will give non-visual browsers a chance to help orient people as they move from one cell to the next. For forms, remember to include labels for form fields. W3C's Web Accessibility Initiative is developing more detailed accessibility guidelines.
Further information will be provided here in the near future on common errors and good practices. We will be including easy to read background material on all aspects of HTML.
/

# XML: from http://www.ucc.ie/xml/

## A.1 What is XML?

XML is the `Extensible Markup Language' (extensible because it is not a fixed format like HTML). It is designed to enable the use of SGML on the World Wide Web.
§ It's actually slightly misnamed: XML itself is not a single markup language: it's a metalanguage to let you design your own markup language. A regular markup language defines a way to describe information in a certain class of documents (*eg* HTML). XML lets you define your own customized markup languages for many classes of document. It can do this because it's written in SGML, the international standard metalanguage for markup languages.

## A.2 What is XML for?

XML is designed `to make it easy and straightforward to use SGML on the Web: easy to define document types, easy to author and manage SGML-defined documents, and easy to transmit and share them across the Web.'
It defines `an extremely simple dialect of SGML which is completely described in the XML Specification. The goal is to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML.'
`For this reason, XML has been designed for ease of implementation, and for interoperability with both SGML and HTML' [quotes from the XML spec].

## A.3 What is SGML?

SGML is the Standard Generalized Markup Language (ISO 8879), the international standard for defining descriptions of the structure and content of different types of electronic document. There is an SGML FAQ at `http://www.infosys.utas.edu.au/info/sgmlfaq.txt` and the SGML Web pages are at `http://www.sil.org/sgml/`.

## A.4 What is HTML?

HTML is the HyperText Markup Language (RFC 1866), a specific application of SGML used in the World Wide Web.

## A.5 Aren't XML, SGML, and HTML all the same thing?

Not quite. SGML is the `mother tongue', used for describing thousands of different document types in many fields of human activity, from transcriptions of ancient Sumerian scrolls to the technical documentation for stealth bombers, and from patients' clinical records to musical notation.
HTML is just one of these document types, the one most frequently used in the Web. It defines a single, fixed type of document with markup that lets you describe a common class of simple office-style report, with headings, paragraphs, lists, illustrations, *etc*, and some provision for hypertext and multimedia.
XML is an abbreviated version of SGML, to make it easier for you to define your own document types, and to make it easier for programmers to write programs to handle them. It omits the more complex and less-used parts of SGML in return for the benefits of being easier to write applications, easier to understand, and more suited to delivery and interoperability over the Web. But it is still SGML, and XML files may still be parsed and validated the same as any other SGML file (see the question on XML software).
Programmers may find it useful to think of XML as being SGML-- rather than HTML++.

## A.6 Who is responsible for XML?

XML is a project of the World Wide Web Consortium (W3C), and the development of the specification is being supervised by their XML Working Group. A Special Interest Group of co-opted contributors and experts from various fields contributes comments and reviews by email.
XML is a public format: it is not a proprietary development of any company.

## A.7 Why is XML such an important development?

It removes two constraints which are holding back Web developments:

1. dependence on a single, inflexible document type (HTML);
2. the complexity of full SGML, whose syntax allows many powerful but hard-to-program options.

XML simplifies the levels of optionality in SGML, and allows the development of user-defined document types on the Web.

## A.8 How does XML make SGML simpler and still let you define your own document types?

To make SGML simpler, XML redefines some of SGML's internal values and parameters, and removes a large number of the more complex and sometimes less-used features which made it harder to write processing programs (see Appendix A of the XML specification).
But it retains all of SGML's structural abilities which let you define your own document type. It also introduces a new class of document which does not require you to use a predefined document type. See the questions about `valid' and `well-formed' documents, and how to define your own document types in the Developers' Section.

## A.9 Why not just carry on extending HTML?

HTML is already overburdened with dozens of interesting but often incompatible inventions from different manufacturers, because it provides only one way of describing your information.
XML will allow groups of people or organizations to create their own customized markup languages for exchanging information in their domain (music, chemistry, electronics, hill-walking, finance, surfing, linguistics, mathematics, knitting, history, engineering, rabbit-keeping *etc*).
HTML is at the limit of its usefulness as a way of describing information, and while it will continue to play an important role for the content it currently represents, many new applications require a more robust and flexible infrastructure.

## A.10 Why do we need all this SGML stuff? Why not just use *Word* or *Notes*?

Information on a network which connects many different types of computer has to be usable on all of them. Public information cannot afford to be restricted to one make or model or manufacturer, or to cede control of its data format to private hands. It is also helpful for such information to be in a form that can be reused in many different ways, as this can minimize wasted time and effort.
SGML is the international standard which is used for defining this kind of application, but those who need an alternative based on different software are entirely free to implement similar services using such a system, especially if they are for private use.

## About PNG (from http://www.cdrom.com/pub/png/pngintro.html)

The Portable Network Graphics (PNG) format was designed to replace the older and simpler GIF format and, to some extent, the much more complex TIFF format. (See the main page or news/history page for background information.) Here we'll concentrate on two major uses: the World Wide Web (WWW) and image-editing.
For the Web, PNG really has three main advantages over GIF: alpha channels (variable transparency), gamma correction (cross-platform control of image brightness), and two-dimensional interlacing (a method of progressive display). PNG also compresses better than GIF in almost every case, but the difference is generally only around 10% to 30%, not a large enough factor to encourage folks to switch on that basis alone. One GIF feature that PNG does *not* try to reproduce is multiple-image support, especially animations; PNG was and is intended to be a single-image format only. (A very PNG-like extension format called MNG is currently being designed, but MNGs and PNGs will have different file extensions and different purposes.)
For image editing, either professional or otherwise, PNG provides a useful format for the storage of intermediate stages of editing. Since PNG's compression is fully lossless--and since it supports up to 48-bit truecolor or 16-bit grayscale--saving, restoring and re-saving an image will not degrade its quality, unlike standard JPEG (even at its highest quality settings). And unlike TIFF, the PNG specification leaves no room for implementors to pick and choose what features they'll support; the result is that a PNG image saved in one application is readable in any other PNG-supporting app. (Note that for transmission of finished truecolor images--especially photographic ones--JPEG is almost always a better choice. Although JPEG's lossy compression can introduce visible artifacts, these can be minimized, and the savings in file size even at high quality levels is much better than is generally possible with a lossless format like PNG.)

## JPEG and GIF (from http://www.cis.ohio-state.edu/hypertext/faq/usenet/jpeg-faq/top.html)

### What is JPEG?

JPEG (pronounced “jay-peg”) is a standardized image compression mechanism. JPEG stands for Joint Photographic Experts Group, the original name of the committee that wrote the standard.
JPEG is designed for compressing either full-color or gray-scale images of natural, real-world scenes. It works well on photographs, naturalistic artwork, and similar material; not so well on lettering, simple cartoons, or line drawings. JPEG handles only still images, but there is a related standard called MPEG for motion pictures.
JPEG is “lossy,” meaning that the decompressed image isn’t quite the same as the one you started with. (There are lossless image compression algorithms, but JPEG achieves much greater compression than is possible with lossless methods.) JPEG is designed to exploit known limitations of the human eye, notably the fact that small color changes are perceived less accurately than small changes in brightness. Thus, JPEG is intended for compressing images that will be looked at by humans. If you plan to machine-analyze your images, the small errors introduced by JPEG may be a problem for you, even if they are invisible to the eye.

### When should I use JPEG, and when should I stick with GIF?

JPEG is *not* going to displace GIF entirely; for some types of images, GIF is superior in image quality, file size, or both. One of the first things to learn about JPEG is which kinds of images to apply it to.
Generally speaking, JPEG is superior to GIF for storing full-color or gray-scale images of “realistic” scenes; that means scanned photographs, continuous-tone artwork, and similar material. Any smooth variation in color, such as occurs in highlighted or shaded areas, will be represented more faithfully and in less space by JPEG than by GIF.
GIF does significantly better on images with only a few distinct colors, such as line drawings and simple cartoons. Not only is GIF lossless for such images, but it often compresses them more than JPEG can. For example, large areas of pixels that are all *exactly* the same color are compressed very efficiently indeed by GIF. JPEG can’t squeeze such data as much as GIF does without introducing visible defects. (One implication of this is that large single-color borders are quite cheap in GIF files, while they are best avoided in JPEG files.)
Computer-drawn images, such as ray-traced scenes, usually fall between photographs and cartoons in terms of complexity. The more complex and subtly rendered the image, the more likely that JPEG will do well on it. The same goes for semi-realistic artwork (fantasy drawings and such).
But icons that use only a few colors are handled better by GIF.

# Style Sheets: http://www.w3.org/Style

## What are style sheets?

Style sheets describe how documents are presented on screens, in print, or perhaps how they are pronounced. Style sheets are soon coming to a browser near you, and this page and its links will tell you all there is to know about style sheets.
By attaching style sheets to structured documents on the Web (e.g. HTML), authors and readers can influence the presentation of documents without sacrificing device-independence or adding new HTML tags. Style sheets have been an W3C activity since the consortium was founded and has resulted in the development of CSS. Recently, a Working Group on XSL was launched.
The easiest way to start experimenting with style sheets is to find a browser that support CSS. Discussions about style sheets are carried out on the www-style@w3.org mailing list and on comp.infosystems.www.authoring.stylesheets.

- …

## CSS

Cascading Style Sheets (CSS) is a style sheet mechanism that has been specifically developed to meet the needs of Web designers and users.

- How can I learn more about CSS?
- What browsers support CSS? And what features do they support?
- What authoring tools support CSS?
- Where are the CSS specifications?

## DSSSL

Where CSS1 is simple, DSSSL is advanced. DSSSL, an ISO standard, is a document tree transformation and style language with many adherents in the SGML community.

> *It's a mistake to put DSSSL into the same bag as scripting languages. Yes, DSSSL is turing-complete; yes, it's a programming language. But a script language (at least the way I use the term) is procedural; DSSSL very definitely is not. DSSSL is entirely functional and entirely side-effect-free. Nothing ever happens in a DSSSL stylesheet. The stylesheet is one giant function whose value is an abstract, device-independent, nonprocedural description of the formatted document that gets fed as a specification of display areas to downstream rendering processes.*
> *-- Jon Bosak*

DSSSL resources on the Web:

- The final Document Style Semantics and Specification Language (DSSSL) draft is available from the Novell Publications Server.
- DSSSList is the DSSSL Users' Mailing List. The archives of the now dysfunctional dsssl-lite mailing list are also available.
- Jade is James Clark's DSSSL engine.

## XSL

W3C has recently launched a Working Group to develop the eXtensible Style Language (XSL). XSL builds on DSSSL and CSS and is primarily targeted for highly structured XML data which e.g. needs element reordering before presentation. One feature of XSL is that it can be user-extended through the ECMAScript language. For more information on XSL see the W3C XSL resource page.

## Dynamic HTML

Dynamic HTML is a term used to describe HTML pages with dynamic content. CSS is one of three components in dynamic HTML; the other two are HTML itself and JavaScript (which is being standardized under the name EcmaScript). The three components are glued together with DOM, the component object model. Dynamic HTML is still in its infancy and current implementations are experimental.

## Synchronized Multimedia (from http://www.w3.org/AudioVideo/Activity.html)

Web technology is limited today when it comes to creating continuous multimedia presentations. For these applications, content authors need to express things like "five minutes into the presentation, show image X and keep it on the screen for ten seconds". More generally speaking, there must be a way to describe the synchronization between the different media (audio, video, text, images) that make up a continuous multimedia presentation.

There is an imminent danger that a plethora of non-interoperable solutions for integrating real-time multimedia content into the Web architecture will emerge. These different solutions will most likely not result from a healthy competition advancing technological progress. In contrast, they will result from a simple lack of communication between the three very different communities involved, namely the Web community, the CD-ROM community and the community working on Internet-based audio/video-on-demand.

# Dublin Core, from http://purl.org/metadata/dublin_core

## Metadata for Electronic Resources

The Dublin Core is a 15-element metadata element set intended to facilitate discovery of electronic resources. Originally conceived for author-generated description of Web resources, it has also attracted the attention of formal resource description communities such as museums and libraries.
The Dublin Core Workshop Series has gathered experts from the library world, the networking and digital library research communities, and a variety of content specialties in a series of focussed, invitational workshops. The building of an interdisciplinary, international consensus around a core element set is the central feature of the three-year evolution of the Dublin Core. The progress represents the emergent wisdom and collective experience of many stakeholders in the resource description arena. An open mailing list supports ongoing work.
The characteristics of the Dublin Core that distinguish it as a prominent candidate for description of electronic resources fall into several categories.

**1. Title Label: "Title"**

The name given to the resource, usually by the Creator or Publisher.

**2. Author or Creator Label: "Creator"**

The person or organization primarily responsible for creating the intellectual content of the resource. For example, authors in the case of written documents, artists, photographers, or illustrators in the case of visual resources.

**3. Subject and Keywords Label: "Subject"**

The topic of the resource. Typically, subject will be expressed as keywords or phrases that describe the subject or content of the resource. The use of controlled vocabularies and formal classification schemes is encouraged.

**4. Description Label: "Description"**

A textual description of the content of the resource, including abstracts in the case of document-like objects or content descriptions in the case of visual resources.

**5. Publisher Label: "Publisher"**

The entity responsible for making the resource available in its present form, such as a publishing house, a university department, or a corporate entity.

**6. Other Contributor Label: "Contributor"**

A person or organization not specified in a Creator element who has made significant intellectual contributions to the resource but whose contribution is secondary to any person or organization specified in a Creator element (for example, editor, transcriber, and illustrator).

**7. Date Label: "Date"**

A date associated with the creation or availability of the resource. Such a date is not to be confused with one belonging in the Coverage element, which would be associated with the resource only insofar as the intellectual content is somehow about that date. Recommended best practice is defined in a profile of ISO 8601 [Date and Time Formats (based on ISO8601), W3C Technical Note, http://www.w3.org/TR/NOTE-datetime] that includes (among others) dates of the forms YYYY and YYYY-MM-DD. In this scheme, for example, the date 1994-11-05 corresponds to November 5, 1994.

**8. Resource Type Label: "Type"**

The category of the resource, such as home page, novel, poem, working paper, technical report, essay, dictionary. For the sake of interoperability, Type should be selected from an enumerated list that is currently under development in the workshop series.

**9. Format Label: "Format"**

The data format of the resource, used to identify the software and possibly hardware that might be needed to display or operate the resource. For the sake of interoperability, Format should be selected from an enumerated list that is currently under development in the workshop series.

**10. Resource Identifier Label: "Identifier"**

A string or number used to uniquely identify the resource. Examples for networked resources include URLs and URNs (when implemented). Other globally-unique identifiers, such as International Standard Book Numbers (ISBN) or other formal names are also candidates for this element.

**11. Source Label: "Source"**

Information about a second resource from which the present resource is derived. While it is generally recommended that elements contain information about the present resource only, this element may contain a date, creator, format, identifier, or other metadata for the second resource when it is considered important for discovery of the present resource; recommended best practice is to use the Relation element instead. For example, it is possible to use a Source date of 1603 in a description of a 1996 film adaptation of a Shakespearean play, but it is preferred instead to use Relation "IsBasedOn" with a reference to a separate resource whose description contains a Date of 1603. Source is not applicable if the present resource is in its original form.

**12. Language Label: "Language"**

The language of the intellectual content of the resource. Where practical, the content of this field should coincide with RFC 1766 [Tags for the Identification of Languages, http://ds.internic.net/rfc/rfc1766.txt ]; examples include en, de, es, fi, fr, ja, th, and zh.

**13. Relation Label: "Relation"**

An identifier of a second resource and its relationship to the present resource. This element permits links between related resources and resource descriptions to be indicated. Examples include an edition of a work (IsVersionOf), a translation of a work (IsBasedOn), a chapter of a book (IsPartOf), and a mechanical transformation of a dataset into an image (IsFormatOf). For the sake of interoperability, relationships should be selected from an enumerated list that is currently under development in the workshop series.

**14. Coverage Label: "Coverage"**

The spatial or temporal characteristics of the intellectual content of the resource. Spatial coverage refers to a physical region (e.g., celestial sector); use coordinates (e.g., longitude and latitude) or place names that are from a controlled list or are fully spelled out. Temporal coverage refers to what the resource is about rather than when it was created or made available (the latter belonging in the Date element); use the same date/time format (often a range) [Date and Time Formats (based on ISO8601), W3C Technical Note, http://www.w3.org/TR/NOTE-datetime] as recommended for the Date element or time periods that are from a controlled list or are fully spelled out.

**15. Rights Management Label: "Rights"**

A rights management statement, an identifier that links to a rights management statement, or an identifier that links to a service providing information about rights management for the resource.

Platform for Internet Content Selection: http://www.w3.org/PICS


# Platform for Internet Content Selection

**PICS** is an infrastructure for associating labels (metadata) with Internet content. It was originally designed to help parents and teachers control what children access on the Internet, but it also facilitates other uses for labels, including code signing and privacy. PICS is a platform on which other rating services and filtering software have been built.

## Flexible Blocking (from http://www.w3.org/PICS/iacwcv2.htm)

Not everyone needs to block reception of the same materials. Parents may not wish to expose their children to sexual or violent images. Businesses may want to prevent their employees from visiting recreational sites during hours of peak network usage. Governments may want to restrict reception of materials that are legal in other countries but not in their own. The "off" button (or disconnecting from the entire Net) is too crude: there should be some way to block only the inappropriate material. Appropriateness, however, is neither an objective nor a universal measure. It depends on at least three factors:

3. The supervisor: parenting styles differ, as do philosophies of management and government.
4. The recipient: what's appropriate for one fifteen year old may not be for an eight-year-old, or even all fifteen-year-olds.
5. The context: a game or chat room that is appropriate to access at home may be inappropriate at work or school.

Computer software can implement access controls that take into account all these factors. The basic idea, illustrated in Figure 1, is to interpose selection software between the recipient and the on-line documents. The software checks labels to determine whether to permit access to particular materials. It may permit access for some users but not others, or at some times but not others.

Prior to PICS there was no standard format for labels, so companies that wished to provide access control had to both develop the software and provide the labels. PICS provides a common format for labels, so that any PICS-compliant selection software can process any PICS-compliant label. A single site or document may have many labels, provided by different organizations. Consumers choose their selection software and their label sources (called *rating services*) independently, as illustrated in Figure 2. This separation allows both markets to flourish: companies that prefer to remain value-neutral can offer selection software without providing any labels; values-oriented organizations, without writing software, can create rating services that provide labels.

## What are URNs? (from http://www.acl.lanl.gov/URN/)

Uniform Resource Names are being developed as an adjunct to Uniform Resource Locators (URLs). URLs, the addresses used in the World Wide Web, typically specify a particular file on a particular machine. This makes it difficult to reorganize the files on your web server, or to replicate the resource on multiple machines in order to distribute the load or to provide fault-tolerance. URNs are intended to overcome these problems. The basic idea is that URNs will be assigned to specify the identity of a resource, rather than its location. ISBNs are an example of what is meant by specifying identity. Given a URN, a client, such as a browser, contacts a resolver to discover locations of the resource. The resolver is a level of indirection that allows us to change the location of a resource, or add multiple locations, without changing the identifier. The resolver may also be able to provide additional information on the resource, such as its price, bibliographic information, etc.

### From http://www.acl.lanl.gov/URN/urn-overview.html

When was the last time you clicked on a link and got a message that "this resource has moved, plesae make a note of its new location"? When was the last time that you accessed a popular site and found yourself fetching the resource at a snail-like pace? When was the last time you clicked a link and the server on the other end was not responding?

These problems, and a few others, arise because the current mechanism for addressing resources on the web (URLs) specify where to go to get what you want, rather than just specifying what is desired. Because they hard-code a machine and path, it is difficult to replicate the resource onto multiple machines to avoid long network delays, overloaded servers, etc. URLs merge the notions of identification and location. If we were able to specify resources by their identity, and had a way of mapping that identify to locations, then we could do several useful things. We could move resources around the network without having to update all the pages in the world that pointed to the resource. We could replicate the resource to multiple locations, in different parts of the world, so that traffic on the backbones is reduced. This also allows us to improve the fault-tolerance of web access. It would also allow very popular resources to be placed onto many different servers to cope with the expected load.

Uniform Resource Names are just such a method for specifying resources by their identity rather than their location. The current system, URLs, is akin to citing a book by its location on the shelves of a particular library. URNs would be akin to citing the work by its ISBN number. With that number, your local library or bookstore can get a copy of the book for you.

Just as your bookstore uses something like "Books in Print" to discover how to order a book given an ISBN, the URN system needs a "resolver" to take a URN and return the URLs where the resource can be found. The hardest part of the URN system is the mechanism for locating the resolver, since no information on its location can be embedded in the URN. The NAPTR proposal is one way to find a resolver from a URN. It offers the considerable advantage of allowing the resolver, as well as the resource, to be replicated. This is an important point, since the failure of a resolver could cause a large number of resources to become impossible to locate.

Another one of the benefits of the NAPTR system is that it allows us to use the wide variety of existing identification schemes that are already in use identifying things like books, movies, magazine articles, and music. The ISBN scheme is most familar to people, but there are identification schemes for all the other types as well.

## What are URCs? (http://www.acl.lanl.gov/URC/ and ad_charter.txt)

Uniform Resource Characteristics (URCs) are descriptions, such asbibliographic or configuration control records, of Internet-accessibleresources. Particular communities expend considerable effort todevelop descriptions that are appropriate for their resources. Many ofthese descriptions are notable more for their commonalties and theirgratuitous differences than for their necessary, community-specific,distinctions.

## HTTP (from RFC 2068)

The Hypertext Transfer Protocol (HTTP) is an application-level protocol for distributed, collaborative, hypermedia information systems. It is a generic, stateless, protocol which can be used for many tasks, such as name servers and distributed object management systems, through extension of its request methods. A feature of HTTP is the typing and negotiation of data representation, allowing systems to be built independently of the data being transferred.

# 1992 definition of HTTP (http://www.w3.org/Protocols/HTTP/HTTP2.html)

*The protocol is basically stateless, a transaction consisting of*

*Connection*

> *The establishment of a connection by the client to the server - when using TCP/IP port 80 is the well-known port, but other non-reserverd ports may be specified in the URL;*

*Request*

> *The sending, by the client, of a request message to the server;*

*Response*

> *The sending, by the server, of a response to the client;*

*Close*

> *The closing of the connection by either both parties.*

*The format of the request and response parts is defined in this specification. Whilst header information defined in this specification is sent in ISO Latin-1 character set in CRLF terminated lines, object transmission in binary is possible.*

## HTTP/1.1

...1 Status of this Memo...1 Abstract...1 Table of Contents...3 1 Introduction ...8 1.1 Purpose ...8 1.2 Requirements ...8 1.3 Terminology ...9 1.4 Overall Operation ...12 2 Notational Conventions and Generic Grammar ...13 2.1 Augmented BNF ...13 2.2 Basic Rules ...15 3 Protocol Parameters ...16 3.1 HTTP Version ...16 3.2 Uniform Resource Identifiers ...17 3.2.1 General Syntax ...17 3.2.2 http URL ...18 3.2.3 URI Comparison ...18 3.3 Date/Time Formats ...18 3.3.1 Full Date ...18 3.3.2 Delta Seconds ...19 3.4 Character Sets ...20 3.4.1 Missing Charset ...20 3.5 Content Codings ...21 3.6 Transfer Codings ...22 3.6.1 Chunked Transfer Coding ...22 3.6.2 Identity Transfer Coding ...23 3.7 Media Types ...23 3.7.1 Canonicalization and Text Defaults ...24 3.7.2 Multipart Types ...25 3.8 Product Tokens ...25 3.9 Quality Values ...26 3.10 Language Tags ...26 3.11 Entity Tags ...26 3.12 Range Units ...27 4 HTTP Message ...27 4.1 Message Types ...27 4.2 Message Headers ...28 4.3 Message Body ...28 4.4 Message Length ...29 4.5 General Header Fields ...30 5 Request ...31 5.1 Request-Line ...31 5.1.1 Method ...31 5.1.2 Request-URI ...32 5.2 The Resource Identified by a Request ...33 Fielding, et al [Page 3]_ INTERNET-DRAFT HTTP/1.1 Friday, March 13, 1998 5.3 Request Header Fields ...33 6 Response ...34 6.1 Status-Line ...34 6.1.1 Status Code and Reason Phrase ...34 6.2 Response Header Fields ...36 7 Entity ...36 7.1 Entity Header Fields ...36 7.2 Entity Body ...37 7.2.1 Type ...37 7.2.2 Entity Length ...37 8 Connections ...38 8.1 Persistent Connections ...38 8.1.1 Purpose ...38 8.1.2 Overall Operation ...38 8.1.3 Proxy Servers ...39 8.1.4 Practical Considerations ...40 8.2 Message Transmission Requirements ...41 8.2.1 Persistent Connections and Flow Control ...41 8.2.2 Monitoring Connections for Error Status Messages ...41 8.2.3 Automatic Retrying of Requests ...41 8.2.4 Use of the 100 (Continue) Status ...41 8.2.5 Client Behavior if Server Prematurely Closes Connection ...43 9 Method Definitions ...44 9.1 Safe and Idempotent Methods ...44 9.1.1 Safe Methods ...44 9.1.2 Idempotent Methods ...44 9.2 OPTIONS ...45 9.3 GET ...46 9.4 HEAD ...46 9.5 POST ...47 9.6 PUT ...47 9.7 DELETE ...48 9.8 TRACE ...49 9.9 CONNECT ...49 10 Status Code Definitions ...49 10.1 Informational 1xx ...49 10.1.1 100 Continue ...50 10.1.2 101 Switching Protocols ...50 10.2 Successful 2xx ...50 10.2.1 200 OK ...50 10.2.2 201 Created ...51 10.2.3 202 Accepted ...51 10.2.4 203 Non-Authoritative Information ...51 10.2.5 204 No Content ...51 10.2.6 205 Reset Content ...52 10.2.7 206 Partial Content ...52 10.3 Redirection 3xx ...53 10.3.1 300 Multiple Choices ...53

## Protocols based on HTTP

### IPP

Internet Printing Protocol: Use HTTP for firewall support

### WebDAV

Support authoring on same servers for same resources.

### HTCPCP (RFC 2324)

a protocol for controlling, monitoring, and diagnosing coffee pots….
HTTP 1.1 ([RFC2068]) permits the transfer of web objects from origin servers to clients. The web is world-wide. HTCPCP is based on HTTP. This is because HTTP is everywhere. It could not be so pervasive without being good. Therefore, HTTP is good. If you want good coffee, HTCPCP needs to be good. To make HTCPCP good, it is good to base HTCPCP on HTTP.

## W3C Copyright Notice: http://www.w3.org/Consortium/Legal/copyright-documents.html