*:96 Internet application layer protocols and standards

Changed pages in the compendiums September to December 2002

*:96 Internet application protocols and standards

Contents and references for a technical university course on Internet application layer protocols and standards.

Last revision: 3 December, 2002
Warning: Versions dated before this date are not valid!

Schema ht 2002

Study Plan

Which overheads are shown in which lecture?

Course presenter: Professor Jacob Palme
phone:+46-8-16 16 67
e-mail: jpalme@dsv.su.se.

course segment codes: SU: *.96, KTH: 2I1263.

Table of contents

Goal

To describe the most common application layer protocols on the Internet, and to describe the general principles and methods used in these protocols.

This course segment is given twice a year, once in the spring and once in the fall. KTH students might only be allowed to participate at one of these two times.

Swedish, less complete info (Svensk information, kortare och ej helt

Compendium 4 page 443

aktuell).

Segment contents

- Methods and solutions in design of distributed application layer protocols: Architecture, ports, layering, protocols, authentication, connection-less and connection-oriented protocols, chaining, referral, multicasting, replication and caching.
- Extensibility in distributed applications.
- How software based on old and new versions of a protocol can interact.
- DNS (Domain Naming System) - Globally unique names, specifying and searching.
- ABNF (Augmented Bachus-Naur Format) - A common way of specifying the syntax in many Internet application layer protocols.
- An alternative encoding method: ASN.1 (Abstract Syntax Notation One).
- A third encoding method: XML: Extended Markup Language.
- The Basic Encoding Rules (BER) for ASN.1 for example Kerberos, LDAP, SNMP and S/MIME.
- URL, URI, URN, URC (Uniform Resource Locator, Identifier, Name, Characteristics).
- Media types.
- SMTP (Simple Mail Transport Protocol) - A protocol for transport of e-mail.
- RFC822 - A protocol for the format of e-mail.
- MIME (Multipurpose Internet Mail Extensions) - A protocol for content and media types, used by both e-mail and the WWW.
- POP (Post Office Protocol) and IMAP - E-mail delivery protocols.
- NNTP (Network News Transfer Protocol) - A protocol for the transmission of Network News.
- FTP (File Transfer Protocol) - A protocol for file transmission.
- HTML (Hypertext Markup Language) - A formatting language for WWW documents. How to produce your own web pages. Tools for web page generation.
- CSS (Cascacing Style Sheets) a language for controlling the layout of HTML and XML documents.
- HTTP (Hypertext Transfer Protocol) - The access protocol of the World Wide Web.
- Directory systems: X.500, Whois and Whois++.
- CGI (Common Gateway Interface) - Standard for interfacing applications to WWW servers.
- PICS (Platform Independent Content Selection) - a protocol for voluntary access control.
- Adaptation of Internet to other languages than English.
- The process of developing standards in IETF (Internet Engineering Task Force), ITU (International Telecommunications Union) and ISO (International Standards Organisation).
- The golden rule: "Be liberal in what you accept, be conservative in what you produce.

Course segment format

The course segment format is lectures, reading materials, a small work task, and an exam.

Compendium 4 page 444

Changes since previous instances of this course segment

Changes in the fall of 2000: There will be more about e-mail standards and technology, and a new compendium 9 with information about e-mail has been added to the course segment documentation.

Changes in the spring of 2000: Much more complete coverage of XML.The work task is changed and is based on XML instead of on HTML. New additional compendiums with XML information.

Changes in the spring of 1999.

Changes in the year 2000.

Changes in the spring of 2001.

Changes in September 2002.

Changes in January 2003.

To get a complete view of the changes, just take any two of the files below and compute a diff between them, for example using the **Tools -> Track changes -> Compare documents** command in MS Word.

2002-08-29
2002-03-05
2002-02-21
2001-12-21
2001-08-28
2001-03-11
2000-11-05
2000-10-22
2000-10-11
2000-02-18

Prerequisites

In order to understand this segment, you have to have some experience in using the Internet, especially the World Wide Web and e-mail. If you do not have such experience, you will have problems following the segment. This segment is not primarily a course segment about usage of the Internet, it is a segment about how Internet works, for students who already know the basics of how to use the Internet. If you do not know the basics of how Internet works, read some book about it and try it out yourself before the start of the course segment.

You should also have some knowledge about how computers work, such as a programming course segment or some other technically oriented course segment about computers.

Special for KTH students

Information about registration, etc. is provided by Pia Haapala, phone +46-

8-16 49 43, e-mail <piah@dsv.su.se>.

All KTH students must register their participation in this course segment to the study counseling ("studievägledning") at KTH.

Students in KTH courses of type self-contained and further education ("fristående", "fortbildning" and "vidareutbildning") must use application forms provided by the education and research administrative unit, Valhallavägen 79, KTH. The forms can be ordered from utbildningsexpeditionen, phone +46-8-790 70 07 or +46-8-790 70 08, fax +46-8-790 68 87.

Mailing List, First Class Conference

Important information about this segment will be distributed through an e-mail mailing list, a First Class conference and a KOM 2000 forum. The same information will be sent to all three media, so students can choose to subscribe to either.

All course segment participants should either register for this mailing list or regularly access the First Class conference or regularly access the KOM 2000 forum.

Registering for the mailing list

To register for the mailing list, send a message to: listserv@su.se which contains the text:

```
SUB INTPROT Nils Nilsson
```

where `"Nils Nilsson"` is replaced by your name (not your e-mail address).

Important: This text should be in the text of the message, *not* in the Subject!

You should get a confirmation message from LISTSERV saying that you have been added to the list.

After doing this, you can send questions about the course segment by e-mail to intprot@su.se. The question and my answer will be published both through the mailing list and in First Class, even if you send the question only to the mailing list.

At the end of the course segment, you can unsubscribe from the mailing list, by sending a message to listserv@su.se which contains the text:

```
UNS INTPROT
```

Using First Class

If you prefer to get the information through First Class, get an account at

Compendium 4 page 446

the DSV First Class server. If you are registered in this course segment, the INTPRO First Class conference should be on your desktop. If you prefer to use First Class, move all submeetings (INTPRO Dokumentation, INTPRO Handledning and INTPRO Fritt forum), to your First Class desktop. However, if you prefer to use First Class, you should still send your questions to the teacher by e-mail to intprot@su.se and not write them in the First Class conference. Your questions, and the teachers's answers, will still appear in the First Class conference.

Get a working e-mail address

In order to take part in this course segment, you have to have a working e-mail address, where you read your mail regularly. Note that DSV can give you an e-mail address on our computers, if you do not have any. You can also get free e-mail accounts with a number of Internet providers.

If you have an e-mail account outside from DSV, but also have a Unix account at DSV (such an account automatically includes an e-mail account), and if you read your e-mail more regularly on the account outside of DSV, then you should arrange for forwarding of e-mail from the DSV account to the external mail account. There is a simple description on how to do this.

Location

The location for the course segment is the department of computer and systems sciences, Stockholm University and KTH, in Electrum, Kista, north of Stockholm. Some exams may be in other locations.

Lectures

Lectures

The lectures are not mandatory, but there may be questions in the exam which ask about things mentioned in the overheads but more fully described in the lectures. Thus, if you do not go to the lectures, read the overheads carefully and ensure that you understand what is behind each overhead, not only what is on the overhead.

Exercise times

You can go to the DSV terminal rooms at any time, but if you go at times when the terminal rooms are booked for this course segment, you have priority of using the computers. At other times, other course segments may have priority in using the computers.

Requirements for passing this course segment

Compendium 4 page 447

For passing this course segment you must pass the exam (3.5 points), and you must submit the XML work task as described below (0.5 points).

Exams

Times for the next exams.

For those exams, which are in the Electrum building, room will be announced on a notice board on lift A floor 6 (usually outside room 601) immediately before the exam and on a telephone answerer on number 674 70 04 the day before the exam.

Do not forget that you must register in advance, in order to participate in the exam.

Note that many of the documents for this course segment can be brought and used during the exam. For more info on which documents can be brought to the exam, see the list of reading material.

A few copies of the reading material, which are allowed to use during the exam, will be available for loan during the exam, for those who have not bought their own copies of these documents.

Most of the exam questions have been collected in a web site where they are ordered in logical order and with references to the replies, in a format suitable for self-studies.

Note: The contents of the course segment has shifted over the years, so old exams may not give a correct representation of the current course segment content. For example, XML was added in the year 2000, exams before 2000 do not ask such detailed questions on XML.

Exam 960920 with correct answers.
Exam 961026 with correct answers.
Exam 970111 with correct answers. (in Swedish)
Exam 971031 with correct answers

Note that since the exams above were given, this course segment has been extended from 3 to 4 weeks (points), by adding part from a previous course segment *:73. Below are some examples of exams from *:73:

Exam 960429 in Swedish and English
Exam 960615 only in Swedish

Below are exams given after the change from 3 to 4 points

Exam 970919 with correct answers
Exam 980508 with correct answers
Exam 980604 with correct answers
Exam 980926 with correct answers
Exam 981121 with correct answers
Exam 990601 with correct answers
Exam 990823 with correct answers
Exam 990918 with correct answers
Exam 991109 with correct answers

Compendium 4 page 448

Exam 000419 with correct answers
Exam 000522 with correct answers
Exam 001114 with correct answers
Exam 010116 with correct answers
Exam 010221 with correct answers
Exam 010831 with correct answers
Exam 011020 with correct answers
Exam 011208 with correct answers
Exam 020306 with correct answers
Exam 020522 with correct answers
Exam 021019 with correct answers

Important note: Since XML was not included in this course segment before the year 2000, exams in the year 2000 or later may contain questions about XML, even though there are no such questions in the old exams.

Observe especially the following exam questions:
001114-4, 980508-1, 990601-4, In these questions, you are not asked to specify something which already exists, but to specify extensions to standard. Such extensions should be written in the format used in related standard, and it should contain both a syntax and a semantics section, unless the question text says that only syntax is required.

Work task

About cheating: It is not allowed to download an existing DTD or XML from the Internet, or from a task done by someone else, and supply it, either as is or modified. It is not allowed to falsely report the validation result. Cheating may result in exclusion from future studies at the university for a period of several weeks or months.

Old work task: The work task is entirely different from the spring 2000 onwards, compared to 1999 and earlier.

Scheduled times: The times scheduled for the work tasks ("laborationer") are only scheduled so that you have priority to the workstations at those times. There will be no tutor present, you can do the work tasks with other workstations or home computers if you so prefer, there is no scheduled plan for the content of these times.

Submission time: Work tasks should be submitted on or before the Monday after the first exam. I will collect late arriving tasks and process them every second month.

1. Write an XML DTD specification, and an XML text following this specification. A hint: Take any document you have, such as a web page, a ticket, a pamphlet, a time-table, and try to encode the information on it using XML.

2.

   Your dtd and xml documents must make use of the following facilities

Compendium 4 page 449

1. **Multiple nested XML elements.**
2. **Different types of attributes on elements.**
3. **You must include a report from a program which checks the correctness of the DTD and XML used, such as the XML validator at http://www.stg.brown.edu/service/xmlvalid/**
   How to use the validator:
   a. Create a subdirectory, to the WWW directory on your unix account at DSV or your personal web site. The subdirectory can, for example, have the name xml.
   b. Put both the DTD and the XML files on this subdirectory.
   c. The XML file should contain the file name of the DTD file as a relative URL in the <!DOCTYPE> tag.
   d. Access the XML validator, and put the URL of the XML file into the URI field in the validator window.
   e. If your account at DSV is "my-name", and the XML file has the path "WWW/xml/xml-file.xml" then the URL to this file will be "http://dsv.su.se/%7Emy-name/xml/xml-file.xml"
      More info on how to use the validator.
4. **Include the report from the validator when you submit the work task.**
5. **Minimum length 20 lines for both the DTD and the XML file.**

3. Use of CSS is permitted, but not required. This means that your XML need not be very readable when printed with a web browser.
4. There will not be any supervision during the allocated lab times. If you have problems, you are welcome to contact me by phone or e-mail or after the lectures, or come to my office room (but I am seldom there).
5. Deliver the completed task and put it into the box for "inlämningsuppgifter" outside studentexpeditionen at DSV in Electrum, or send it by postal mail (**not e-mail**) to Jacob Palme not later than the first Monday after the first exam after the start of the course segment.

   Your submission of the work task **must** contain the following items:

   a. the DSV preface for work tasks (available in HTML or RTF-format).
   b. Your social security number ("personnummer").
   c. A print-out on paper of the DTD TEXT.
   d. A print-out of the XML text using this DTD.
   e. A print-out from a service which validates the correctness of the DTD and XML.

Submit the work task in the box for work task submissions outside Studentexpeditionen, DSV, or send it by postal mail to Jacob Palme, DSV, Forum 100, 164 40 Kista.

If you need extra fast handling of the work task, send it by postal mail to Jacob Palme, Skeppargatan 73, 11530 Stockholm, and mark the work task with the words "URGENT" or "BRÅTTOM" in large visible letters on the first page of the work task submission.

Students enrolled in the course segment in 1999 or earlier, can instead perform the old work task which was used in the course segment during those years.

Some hints on common problems with the XML task

Problems with National characters (àäöüøé¿, etc.) in XML: You must choose to encode the National characters using either UTF-8 or ISO 8859-1. If you use ISO 8859-1, then the processing instruction

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
```

must be placed at the top of the XML file. If you use national characters in the DTD file, you may have to put a similar instruction in that file.

The validator cannot find my files: Check with an ordinary web browser that you can access the DTD and XML files using the same URL as sent to the validator.

How to get neat layout of the XML when shown with a web browser: This is *not* required for the work task. But if you want to do it, you can use CSS, see the example in compendium 8, section 1.35.

Combining DTD and XML in one file: It is possible to combine DTD and XML in a single file. This means that you do not have to upload the DTD file to a server before testing your XML.

Reading

The reading material consists of nine compendiums.

They are sold by Studerandeexpeditionen, DSV.

In the autumn of 2002, we will send the compendiums for the spring of 2002 plus the additions-compendium with new information added in the autumn of 2002.

Table of contents for the reading materials sold by Studentexpeditionen: http://dsv.su.se/jpalme/internet-course/p96-compendiums.pdf

Incomplete texts for the compendiums are available online in PDF format as follows:

Compendium No. and link	Allowed during the exam?	Content	Included and not included
compendium-1.pdf	Yes	Origiginal standards document for ASN.1, ABNF, DNS, E-mail and LDAP.	Probably all.
compendium-2.pdf	Yes	Original standards on FTP, Cookies, Usenet News, HTTP, URLs and lists of port numbers and media types.	Probably all.
compendium-3.pdf	Yes	HTML 4.01 specification.	Yes.
compendium-4.pdf	No	Descriptions of various standards.	Only page 444-453.

Compendium 4 page 451

compendium-5.pdf	No	Various descriptions and copies of overheads, coding methods.	All except page 611-854
compendium-6.pdf	No	Introduction, study plan, overview.	All.
compendium-7.pdf	Yes	XML Recommendation.	Yes.
compendium-8.pdf	No	CGI OHs, ASN.1 solution advice, XML OHs.	Yes.
compendium-9.pdf	No	Font size, e-mail basics OHs, messaging overview, etc.	Yes.

List of additional (partly not mandatory, and partly not up-to-date) reading materials for this course segment:
http://dsv.su.se/jpalme/internet-course/P96-litt.html

Study Plan:
http://dsv.su.se/jpalme/internet-course/p96-study-plan.html

Additional overheads for the lectures might be added at URL:
http://dsv.su.se/jpalme/internet-course/extra-overheads.html

Introductory chapters with basic concepts for this course segment:
http://dsv.su.se/jpalme/abook/basics-chapter.pdf
http://dsv.su.se/jpalme/abook/coding.pdf

An explanation of how to think when solving ASN.1 question:
http://dsv.su.se/jpalme/internet-course/solving-asn-1-exercise.html

A comparison of ASN.1, ABNF, RFC822 and XML

Web font comparisons:
http://dsv.su.se/jpalme/internet-course/font-comparisons.html

Font size comparisons:
http://dsv.su.se/jpalme/internet-course/font-size-comparisons.html

Why bitmapped screen dumps sometimes get ugly when shown on the screen
http://dsv.su.se/jpalme/internet-course/why-bitmapped-ohs-are-ugly/

A web-based ASN.1 series of exercises with automatic checking that the answer you write is correct can be found at
http://asn1.elibel.tm.fr/tutorial.

Additional Reading (If you want to learn more)

"ASN.1 - Communication between heterogeneous systems" by Olivier Dubuisson, available on the net and from bookstores. A very comprehensive and detailed book about ASN.1.

Compendium 4 page 452

Other course segments which may be of interest

URL: http://dsv.su.se/jpalme/courses.html

URL for this web page

http://dsv.su.se/jpalme/internet-course/Int-app-prot-kurs.html

Compendium 4 page 453

P A R T 1

```
--1430317162
Content-Type: TEXT/PLAIN; CHARSET=ISO 8859-1
Content-Transfer-Encoding: QUOTED-PRINTABLE

Test message containing 8-bit characters.
AE=3D=80, OE=3D=85.

Test message containing 8-bit characters.
AE=3D=80, OE=3D=85.
```
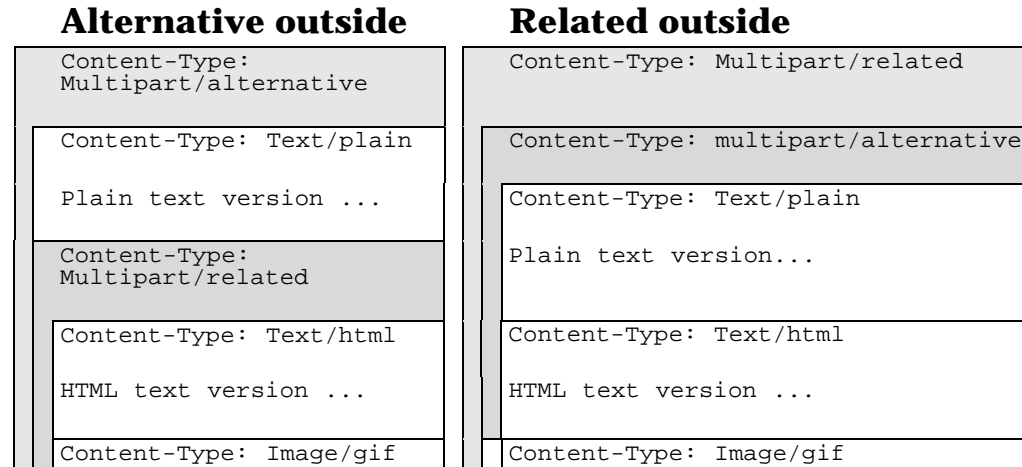
P A R T 2

```
--1430317162
Content-Type: TEXT/PLAIN; CHARSET=ISO 8859-1
Content-Transfer-Encoding: BASE64

VGVzdCBtZXNzYWdlIGNvbnRhaW5pbmcgOC1iaXQgY2hhcmFjdGVy
gCwgT0U9hS4K
--1430317162--
```

Multipart/related
(HTML with inline objects)

RFC 2110-2112

Method 1

```
Content-Type: Multipart/related
```

```
Content-Type: Text/html

Here is a picture
<img src="cid:12345@jp@host.nu">
```

```
Content-Type: Image/gif
Content-ID: :12345@jp@host.nu

VGVzdCBtZXNzYWdlIGNvbnRhaW5pbm
```

Method 2

```
Content-Type: Multipart/related
```

```
Content-Type: Text/html

Here is a picture
<img src="my-picture.gif">
```

```
Content-Type: Image/gif
Content-location: my-picture.gif

VGVzdCBtZXNzYWdlIGNvbnRhaW5pbm
```

Combining multipart/alternative with multipart/related

Alternative outside

```
Content-Type:
Multipart/alternative
```
```
Content-Type: Text/plain

Plain text version ...
```
```
Content-Type:
Multipart/related
```
```
Content-Type: Text/html

HTML text version ...
```
```
Content-Type: Image/gif
```

Related outside

```
Content-Type: Multipart/related
```
```
Content-Type: multipart/alternative
```
```
Content-Type: Text/plain

Plain text version...
```
```
Content-Type: Text/html

HTML text version ...
```
```
Content-Type: Image/gif
```

See also compendium 9 page 57

Privacy Enhanced Mail (RFC 1421-1424)

- Control who may read messages;

- Ensure that the message content has not been modified or corrupted;

- Ensure that a message comes from the stated sender; and

- Eliminate the possibility that the sender can repudiate sending a message.

Security services: confidentiality, integrity, sender authentication, recipient authentication and sender's nonrepudiation. It may be implemented using either secret key or public-key cryptography.
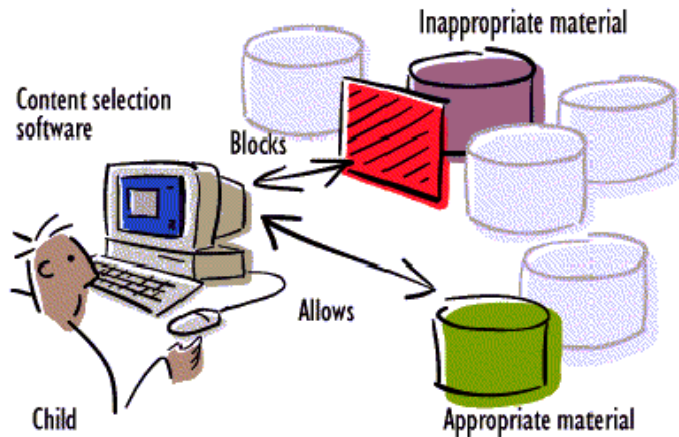
PEM in MIME: MOSS (RFC 1847, RFC 1848)

Alternative: PGP (Pretty Good Privacy)
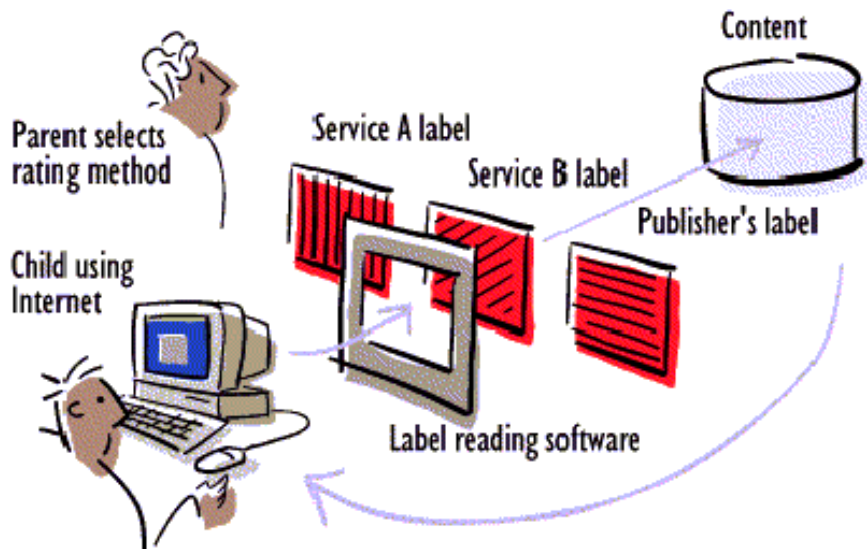
PICS 1 – Platform for Internet Content Selection

From a presentation by Jim Miller and Paul Resnic, found at URL:

`http://www.chg.ru/w3/PICS/951030/JM/talk.htm` **and**
`http://www.bilkent.edu.tr/pub/WWW/PICS/iacwc.htm`

PICS 3

PICS Technical Mission, excerpt

The technical working group will devise a values-neutral infrastructure for Internet content labeling. The three primary goals are to:

1. enable content providers to voluntarily label the content they create and distribute.
2. enable third-party labeling services to associate additional labels with content created and distributed by others. Services may devise their own labeling systems, and the same content may receive conflicting labels from different services.
3. enable parents and teachers to use the labels to control the information that children under their supervision receive.

See also compendium 9 page 57

PICS 2

PICS 4

A rating *service* is an individual, group, organization or company that produces labels for information. A *rating system* is a way of rating information, consisting of one or more *categories* and a *scale* for each category.

The **Motion Picture Association of America** (MPAA) is a rating service, which uses a well-known (in the United States) rating system for rating movies. Other organizations also provide rating systems or services, such as **SafeSurf** and **SurfWatch** (both PICS founding members) and the **Recreational Software Advisory Council** (a PICS supporting member).
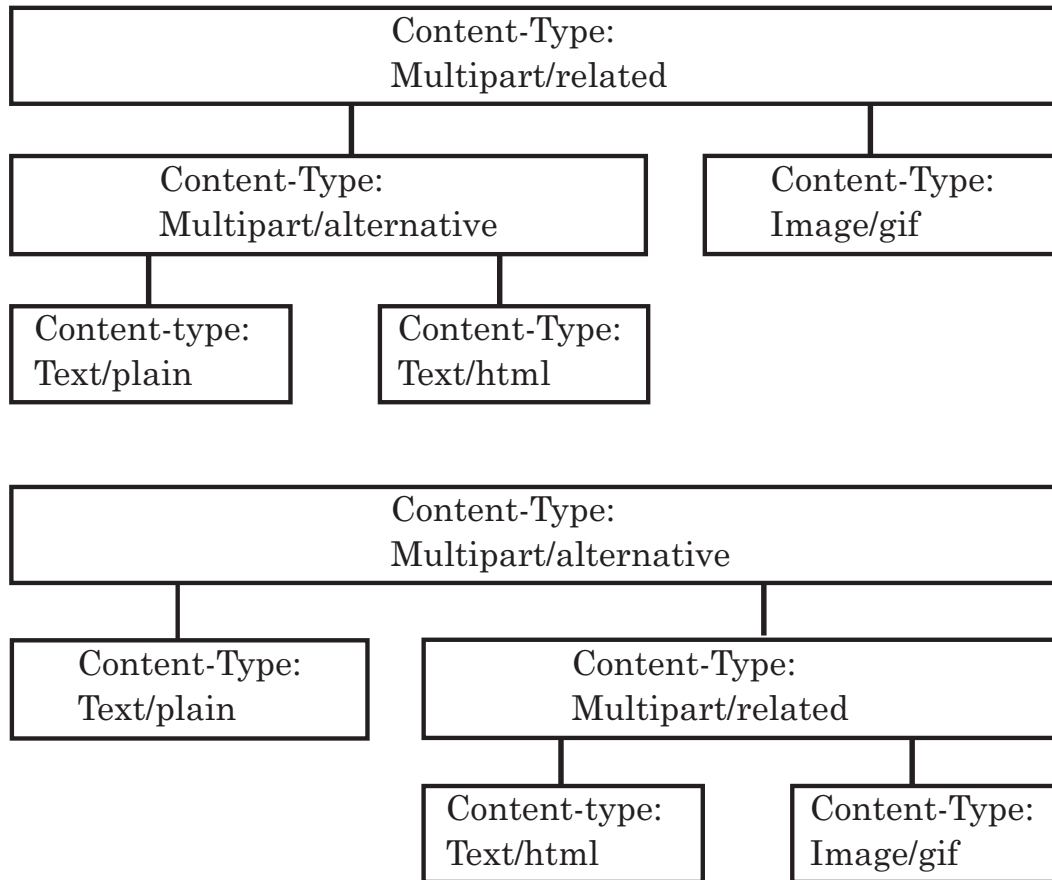
A rating system provides a number of *categories* (or *dimensions*) along which information can be rated.

The MPAA rating system has only one category, the overall rating of the movie. The RSAC rating system has three categories: *nudity/sex*, *violence*, and *language*.

Combining multipart/related and multipart/alternative

```
┌─────────────────────────────────────────────────────────┐
│                    Content-Type:                          │
│                    Multipart/related                      │
└─────────────────────────────────────────────────────────┘
         │                                    │
┌──────────────────────────────┐    ┌──────────────────────┐
│        Content-Type:          │    │    Content-Type:     │
│        Multipart/alternative  │    │    Image/gif         │
└──────────────────────────────┘    └──────────────────────┘
       │                  │
┌──────────────┐   ┌──────────────┐
│ Content-type:│   │ Content-Type:│
│ Text/plain   │   │ Text/html    │
└──────────────┘   └──────────────┘
```

```
┌─────────────────────────────────────────────────────────┐
│                    Content-Type:                          │
│                    Multipart/alternative                  │
└─────────────────────────────────────────────────────────┘
         │                                    │
┌──────────────────────┐    ┌──────────────────────────────┐
│    Content-Type:      │    │        Content-Type:          │
│    Text/plain         │    │        Multipart/related      │
└──────────────────────┘    └──────────────────────────────┘
                                    │                  │
                            ┌──────────────┐   ┌──────────────┐
                            │ Content-type:│   │ Content-Type:│
                            │ Text/html    │   │ Image/gif    │
                            └──────────────┘   └──────────────┘
```

PICS Terminology

VAC	Voluntary Access Control
PICS	Platform for Internet Content Selection
WWWC, W^3C	World Wide Web Consortium
Rating service	Service provider providing rating services
Rating system	Schema for a rating service
Content label	The rating of a particular object (document, site, domain)

Where are the Overheads?

This document will tell you where in the compendiums for the course segment *:96 2I1263, 2I4063 you can find the overhead pictures used in the lectures.

Compendium	Page	Lecture
5	862	Basics
5	877	Coding - Basics
5	881	Coding - ABNF
5	889	Coding - ASN.1
5	974	HTML
5	947	URI, media types
8	115	XML
5	1018	Cookies
5	1009	HTTP
9	3	E-mail - basics
5	955	E-mail - SMTP
5	959	E-mail - DSN
5	960	E-mail - Message format, MIME
9	57	E-mail - Message format, MIME
5	1020	X.500
5	1024	PICS
9	57	PICS
8	105	CGI