

# \*:96 Internet application layer protocols and standards

Compendium with new or modified pages for the fall of 2002

URL for the home page of this course:

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

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

**Last revision:** 10 August, 2002

**Warning:** Versions dated before this date are not valid!

[Schema ht 2002](#)

[Study Plan](#)

**Course presenter:** Professor [Jacob Palme](#)

**phone:** +46-8-16 16 67

**e-mail:** [jpalme@dsv.su.se](mailto:jpalme@dsv.su.se).

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

## Table of contents

- [Segment content](#)
- [Changes since earlier times](#)
- [Prerequisites](#)
- [Mailing list](#)
- [Location](#)
- [Lectures](#)
- [course segment format](#)
- [Requirements for passing this course segment](#)
- [Exams](#)
- [Work task](#)
  - [Submitting the work task](#)
- [Reading](#)
  - [Study plan](#)

## 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 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 Backus-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 (Cascading 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 och 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](#).

## 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.](#)**

## 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](mailto: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](mailto: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](mailto: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](mailto: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 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](mailto: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

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](#)  
[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](#)

**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

**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**

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.

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 (åäöüøé, 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
<a href="#">compendium-1.pdf</a>	Yes	Original standards document for ASN.1, ABNF, DNS, E-mail and LDAP.	Probably all.
<a href="#">compendium-2.pdf</a>	Yes	Original standards on FTP, Cookies, Usenet News, HTTP, URLs and lists of port numbers and media types.	Probably all.
<a href="#">compendium-3.pdf</a>	Yes	HTML 4.01 specification.	Yes.
<a href="#">compendium-4.pdf</a>	No	Descriptions of various standards.	Only page 444-453.
<a href="#">compendium-5.pdf</a>	No	Various descriptions and copies of overheads, coding methods.	All except page 611-854
<a href="#">compendium-6.pdf</a>	No	Introduction, study plan, overview.	All.
<a href="#">compendium-7.pdf</a>	Yes	XML Recommendation.	Yes.
<a href="#">compendium-8.pdf</a>	No	CGI OHs, ASN.1 solution advice, XML OHs.	Yes.
<a href="#">compendium-9.pdf</a>	No	Font size, e-mail basics OHs, messaging overview, etc.	Yes.
<a href="#">p96-diff-020810.pdf</a>	No	Additions to the other compendiums in the fall of 2002.	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/>

## **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>

# **\*:96 Internet application layer protocols and standards**

## **Compendium 5: Not allowed during the exam**

**Last revision: 10 Aug 2002**

Quick HTML Guide .....	592-594
Getting started with HTML .....	595-597
Adding a touch of style.....	598-602
The Bare Bones Guide to HTML.....	603-609
Space: The First Frontier .....	611-612
Top Ten Mistakes in web Design .....	613
PICS Internet Access Controls Without Censorship.....	614-618
A Beginner's Guide to URLs.....	619-620
En stilguide för väven .....	773-779
A Layman's Guide to a Subset of ASN.1, BER and DER.....	821-854
*:96 Overheads Part 1: Basics .....	862-876
*:96 Overheads Part 2a: ABNF, ASN.1, URL, Media Types .....	877-885
*:96 Overheads Part 2b: ASN.1 .....	888-922
*:96 Overheads Part 2bx: ASN.1 solutions to exercises.....	923-932
*:96 Overheads Part 2ca: XML.....	933-944
*:96 Overheads Part 2b: XML solutions to exercises.....	945-946
*:96 Overheads Part 2c: URL, Media types.....	947-950
*:96 Overheads Part 3: E-mail, SMTP, RFC822, MIME .....	955-964
*:96 Overheads Part 4: Message Delivery Protocols (POP and IMAP) .....	965-969
*:96 Overheads Part 5: FTP .....	970-973
*:96 Overheads Part 6: World Wide Web, Hypertext Markup Language (HTML) .....	974-1008
*:96 Overheads Part 7: Hypertext Transfer Protocol (HTTP).....	1009-1017
*:96 Overheads Part 7b: Cookies.....	1018-1019
*:96 Overheads Part 8: Directory systems, PICS.....	1020-1025
Coding Methods .....	1026-1089

*The documents are not ordered in a suitable order for reading them,  
see compendium 6 page 909-911*



### Exercise 4

An identifier in a programming language is allowed to contain between 1 and 6 letters and digits, the first character must be a letter. Only upper case character are used. Write an ABNF specification for the syntax of such an identifier.

### Solution

```
ALPHA = "A" / "B" / "C" / "D" / "E" / "F" / "G" /  
"H" / "I" / "J" / "K" / "L" / "M" / "N" / "O" / "P"  
/ "Q" / "R" / "S" / "T" / "U" / "V" / "X" / "Y" /  
"Z"  
DIGIT = "0" / "1" / "2" / "3" / "4" / "5" / "6" /  
"7" / "8" / "9"  
Identifier = ALPHA *5( ALPHA / DIGIT )
```

# **\*:96 Internet application layer protocols and standards**

## **Compendium 6: *Not allowed during the exam***

**Last revision: 8 Aug 2002**

Introduction and basic concepts.....	876-908
Study plan for this course.....	909-911
An overview of common Internet protocols and services.....	912

# Internet Application Protocols and Standards

## Recommended Study Plan

File name:

<http://dsv.su.se/jpalme/internet-course/p96-study-plan.html>

Latest change: 2001-12-21

Home page for this course:

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

## Course Description, Work Task Description

4	448-453	Course Segment Description
6	912	An overview of common Internet protocols and services

## Basic Concepts, DNS

5	862-876	Overhead pictures for the introductory lectures
6	884-908	Introduction and basic concepts
5	1026-1089	Coding methods
4	465	Cache Consistency Mechanisms
4	472-494	The Domain Name System
1	19-46	RFC 1034: Domain Names
4	463-464	Distributed File Systems
5	619-620	A Beginner's Guide to URLs
2	425-437	Uniform Resource Locators (URL)
5	947-950	Overheads on URL and media types
2	445-460	IANA Registry of Port Numbers
2	461-468	IANA Registry of Media Types

## ABNF

5	1035-1041	Coding methods
5	877-887	*:96 Overheads Part 2: ABNF
1	12-18	RFC 2234: Augmented BNF for Syntax Specifications: ABNF

**ASN.1**

5	1041-1062	Coding methods
5	821-854	A Layman's Guide to a Subset of ASN.1, BER and DER
1	2-11	ASN.1 syntax
4	459-462	Övningsuppgifter på ASN.1 och BER

**XML**

5	1066-1079	XML Introduction
5	933-946	XML OHs
7	913-929	XML Recommendation (=Standard)

**E-mail**

5	955-969	*:96 Overheads Part 3: E-mail
4	495-508	Applications: Electronic Mail
4	701-705	*:96 Overheads Part 4: Message Delivery Protocols
4	738-743	Post Office Protocol
4	744-745	Interactive Mail Access Protocol
1	47-305	RFC 2821, 2822, 2197, 2045, 2047, 2048, 2049, 1891, 1892, 1894, 1725, 2060

**NNTP, Usenet News Transfer Protocol**

4	749-750	News and Usenet
4	509-519	NNTP, Network News Transfer Protocol
2	299-308	RFC 1036: Standard for Interchange of Usenet Messages
2	411-424	RFC 977: NNTP

**X.500, LDAP = Light-Weight Directory Access Protocol**

4	558-564	The Organisation of the Directory
1	307-341	RFC 2251-2252 LDAP

**FTP = File Transfer Protocol**



# An overview of common Internet protocols and services

It is a pre-requisite for [this course](#) that you have some experience with using Internet services as an ordinary user. The course is **not** on how to use the services, it is about the protocols used to provide them.

Last revision: 10 August, 2002 by [Jacob Palme](#).

Protocol name	Main usage	Clients	Servers
DNS	<ol style="list-style-type: none"> <li>1. Translating domain names to numerical host addresses.</li> <li>2. Finding the e-mail server handling mail to e-mail addresses with certain domain names.</li> </ol>	All kinds of clients and name servers.	(DNS) name servers.
HTTP (and HTML, GIF, JPEG as formats)	Downloading web pages in the WWW. Can also be used to send in filled in forms and to send in files. Also used for many specialized protocols based on HTTP.	Web browsers.	HTTP servers.
SMTP (and RFC822 and MIME as formats)	Sending and forwarding of e-mail to and between MTAs (Message Transfer Agents).	Mail clients and SMTP servers.	SMTP servers.
(SMIME and PGP)	Adding security features to e-mail.	Mail clients.	SMTP servers, certificate servers.
POP and IMAP	Downloading of incoming e-mail to the mail clients for their recipients.	Mail clients.	POP or IMAP servers.
NNTP	Submission, downloading and forwarding of Usenet News articles.	News clients and news servers.	News servers.
FTP	Anonymous downloading of files, non-anonymous transfer of files between directories protected by log in.	FTP clients, Web browsers.	FTP servers.

Gopher	An old, nowadays not much used protocols, which can be seen as a limited subset of HTTP.	Web browsers, Gopher clients.	Gopher servers.
PICS	Protection of children from material on the net regarded as unsuitable for them, and other kinds of controlled censorship.	All kinds of clients.	PICS servers, PICS labels on, for example, HTML pages.
LDAP	Searching in directories.	LDAP clients, often built into e-mail clients.	LDAP servers.
WebDav	Uploading and downloading of web pages edited by a group of people.	WebDav clients.	WebDav-compliant HTTP servers.

# Bildformat i datorer

När bilder skall hanteras i en dator, måste bilderna, liksom all annan information en dator, omvandlas till en serie ettor och nollor. En sådan etta eller nolla kallas för en bit (se vidare ordet *bit* i NE). Det finns ett flertal olika metoder att lagra bilder i datorer.

Två huvudmetoder är *linjebilder* och s.k. *bitmappade* bilder. Linjebilder beskriver en bild logiskt genom att lagra kommandon av typ ”dra en linje som är 0,1 mm tjock från punkt 12,45 till punkt 26,72”. Linjebilder använder utöver raka linjer också cirklar/ellipser eller delar av cirklar/ellipser (båglinjer). *Bezierkurvor* är ett avancerat format att beskriva en komplex linje som en serie av delsegment av båglinjer och raka linjer. Linjegrafik kombinerast ofta med fyllning av hela ytan inom en sluten linje med en viss färg.

Bitmappade bilder delar upp bilden i ett mönster av bildpunkter, och lagrar färg och ljushet för varje sådan bildpunkt. En sådan bildpunkt kallas för en pixel. För bitmappade bilder gäller att tätare pixlar ger en skarpere bild. Dataskärmar har ofta en pixeltäthet av 72-96 pixels per tum, medan färgbilder i tidningar ofta har en pixeltäthet av 150-300 pixels per tum och tryckt text i tidningar ofta produceras med en pixeltäthet av 1200 pixels/tum eller mer.

Om man bara vill hantera helt svarta och helt vita bildpunkter, räcker det med en bit per pixel. Med 8 bitar per pixel får man 256 olika färgnyanser, och med 24 eller 32 bitar får man miljontals olika färgnyanser.

Bara 8 bitar per pixel kan göra det svårt att bra återge bilder med många färgnyanser. Det finns två varianter av hur man kan återge en bild med bara 8 bitar per pixel. Med den ena metoden använder man en standarduppsättning av höst 256 färger, samma för alla bilder. Den metoden, som används på en del mycket enkla och billiga datorer, ger ibland en kraftig förvrängning. Med den andra metoden väljer man de 256 nyanser som bäst passar för en viss bild. Den metoden ger mycket bättre bild, ofta kan det mänskliga ögat inte skilja en sådan bild från en bild med mycket fler bitar/pixel. Men om man konverterar

en bild med fler bitar/pixel och fler än 256 färgnyanser till en bild med bara 8 bitar/pixel, så förstör man information, man kan inte få tillbaka exakt den ursprungliga bilden igen när man öppnar den lagrade bilden.

När man omvandlar en bild med mer än 256 färgnyanser till 256 färger, använder man sig ofta av s.k. *gittring* (eng. *Dithering*), där ett mönster av bildpunkter, sedd på långt avstånd, kan se ut som en mellanliggande blandfärg.

När man ombildar en linjebild till en bitmappad bild, kan man använda sig av en metod som kallas för *kantutjämning* (eng. *antialiasing*) och som gör att sneda linjer inte blir så taggiga.

Om man vill använda en standarduppsättning av färger som fungerar på nästan alla datorer med färgskärm, använder man den s.k. Netscape eller webpaletten, som består av 216 färger.

Om man förstörar en bitmappad bild, blir pixeltätheten lägre och bilden blir suddigare eller hackigare. En linjebild däremot blir lika skarp även om man förstörar den.

En linjebild måste alltid förr eller senare omvandlas till en bitmappad bild när den skall visas på en skärm eller tryckas på papper. Processen att omvandla en linjebild till en bitmappad bild kallas för *rippning*, från engelska RIP = Raster Image Processor (se ordet *RIP* i NE). Ett närliggande ord är *rendering* (se ordet *datorgrafik* i NE), som handlar om att omvandla en förenklad logisk beskrivning av en ofta tredimensionell bild med belysning och skuggor till en färdig bitmappad tvådimensionell bild.

Bitmappade bilder kräver ofta mycket minne i datorn. En hel A4-sida med 300 pixels per tum och med 24 bitar/pixel kräver 26 megabyte och tar 14 minuter att ladda ner via modem med 32 kbit/sekund. För att spara lagringsutrymme och nedladdningstid använder man sig därför ofta av komprimering. Två vanliga komprimeringsmetoder är varianter av LPV (Lempel-Ziv-Welch) och JPEG. LPV är en icke-förstörande komprimeringsmetod – exakt samma bild kan återskapas när komprimeringen upphävs. LPV ger en kraftig komprimering för

bilder som innehåller återkommande mönster, vilket är vanligt när ursprungligen linjebilder omvandlas till bitmappade bilder.

JPEG bygger på kunskap om det mänskliga ögat. T.ex. sitter de punkter på näthinnan, som kan skilja på olika färger, mycket glesare än de som kan skilja på olika ljushetsgrad. Därför kan man ta bort mycket av färginformationen ur en bild utan att det mänskliga ögat märker något. Bl.a. upptäcker JPEG gradvisa övergångar mellan två färger och ersätter dem med en mera reguljär gradvis övergång, vilket kan lagras med färre bitar. JPEG ger en kraftig komprimering vid bilder som innehåller mjuka övergångar mellan olika färger, vilket är vanligt för t.ex. fotografier. JPEG är inte bra för vissa bilder med stora jämnt färgade ytor. För sådana bilder kan det uppstå s.k. artefakter, alltså onaturliga förvrängningar, när man sparar dem i JPEG-format.

Eftersom JPEG tar bort information ur en bild, får man inte tillbaka samma bild igen när komprimeringen upphävs. Varje gång man öppnar, ändrar i och lagrar en JPEG-bild, blir bilden suddigare och mer förvrängd. Man bör därför vara försiktig med att använda JPEG som mellanformat för en bild som man kan tänkas ändra på i framtiden.

MPEG-formatet för komprimering av video liknar JPEG, men tar också hänsyn till att det ofta är små skillnader mellan två på varandra följande bildrutor i en film.

Några av de vanligaste förekommande bildformaten är:

GIF passar för bilder med mönster och jämnt färgade ytor. Maximalt 256 färgnyanser kan användas, men man kan välja de 256 nyanser som bäst passar för en viss bild. GIF tillåter också delvis genomskinliga bilder, och bilder med *animering*, alltså rörliga bilder bestående av en serie av bildrutor. GIF kan också använda s.k. *interlacing*. Därmed menas att bilden sänds till en dator så att man först kan visa en suddigare bild som sedan blir allt skarpare när mer information har överförts. GIF använder LPV-metoden för komprimering.

PNG liknar GIF, men tillåter mer än 256

nyanser. Även PNG använder en LPV-liknande metod för komprimering, men inte just den ursprungliga LPV-metoden, bl.a. därför att denna metod är patentskyddad till år 2003.

TIFF är en metod som används mycket för att sända bitmappade bilder till tryckerier och för att lagra bilder vid produktion av tryckta skrifter. TIFF kan använda LPV-metoden för komprimering. TIFF är en icke-förstörande lagringsmetod.

BMP är ett på Windowsdatorer vanligt Microsoft-format med egenskaper delvis liknande GIF och PNG.

EPS och Postscript (två varianter av samma format) är det vanligaste formatet för linjebilder, men även bitmappade bilder och text (med angivande av typsnitt och storlek) kan ingå som del i en EPS-bild. Därför är EPS ett vanligt format vid produktion av tryckta skrifter. EPS används för enstaka bilder, Postscript för hela färdiga sidor som skickas till en skrivare. Postscriptfiler kan bli mycket stora jämfört med andra format.



PDF (Portable Document Format) är ett format som komprimerar EPS och Postscript. Det användes ursprungligen av programmet Adobe Acrobat, men stöds numera av flera andra program från olika tillverkare. PDF innehåller också en funktion att typsnitt som inte finns på den mottagande datorn kan ersättas av andra, liknande typsnitt, som även kan ändras för att ännu mer likna det ursprungliga typsnittet. PDF innehåller också funktioner för bl.a. hyperlänkar (klickbara länkar som i World Wide Web) och formulär med ifyllbara fält. PDF används ofta som alternativ till HTML för webbsidor, därför att PDF kan ge en mera exakt återgivning av skaparens intentioner än HTML. En nackdel med PDF, jämfört med HTML, är att sidor inte enkelt kan omvandlas genom att ändra antal tecken/rad för att passa visning i små fönster, t.ex. på mobila datorer och läsare för elektroniska böcker. PDF kan ställas in på mer komprimering (passar för webbdokument) eller mindre komprimering (passar för trycksaksproduktionen). PDF används mycket inom trycksaksproduktion.






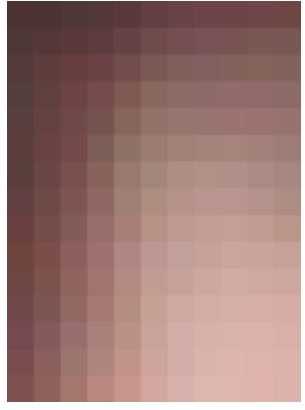
## Kort sammanställning av vanliga format:


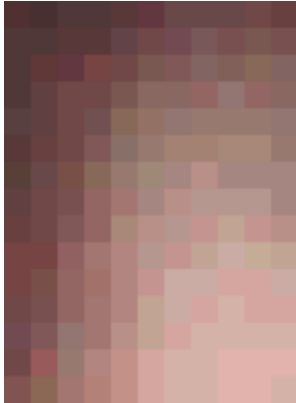

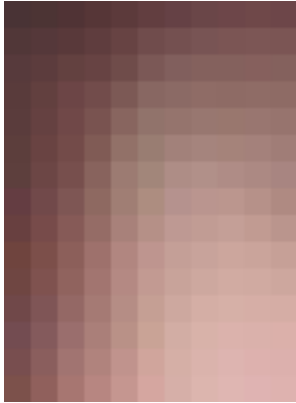

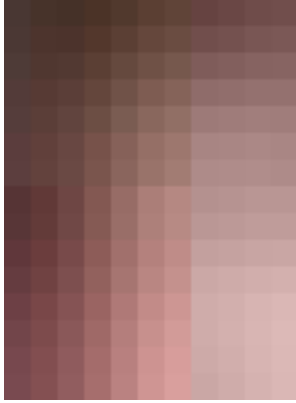

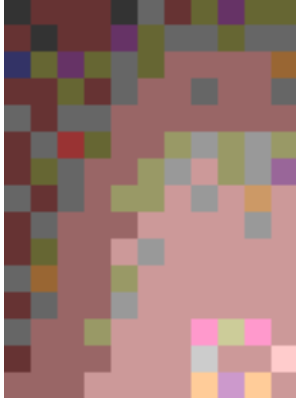
Namn	Vanligt slut på filnamnet	Egenskaper	Användning
GIF	.gif	Max 256 färger, LPV-komprimering, animering och delvis genomskinliga bilder. Om den ursprungliga bilden innehåller mer än 256 färgnyanser förstörs information vid omvandling till GIF-format.	Webbilder, speciellt om de innehåller jämnt färgade ytor.
PNG	.png	Liknar GIF men kan hantera mer än 256 färger och har andra fördelar.	Webbilder, stöds dock inte av äldre webbläsare.
JPEG	.jpeg .jpg	Förstörande komprimering som tar hänsyn till det mänskliga ögats begränsningar. Kan ge mycket kraftig komprimering utan att det mänskliga ögat märker att bilden förstörts.	Fotografier, webbilder.
MPEG	.mpeg .mpg	Format för video, liknar JPEG men utnyttjar även det förhållandet att två på varandra följande bildrutor ofta liknar varandra för att uppnå ännu bättre komprimering.	Video.
EPS	.eps	Linjegravik, men kan även innehålla text och bitmappad grafik.	Trycksaksproduktion.
Postscript	.ps	Variant av EPS för att återge hela sidor.	Trycksaksproduktion, kommunikation mellan dator och skrivare.
PDF	.pdf	Komprimerad version av Postscript med tillägsfunktioner.	Trycksaksproduktion, trycksaksliknande webb-dokument.

## Bildexempel

Här visas några exempel på bilder som illustrerar effekten av olika bildformat. Alla bilderna är i 72-96 pixels/tum, alltså det normala formatet för bilder i webbsidor (medan tryckta skrifter oftast använder mellan 160 och 1200 pixels/tum). "Förstorad 10 gånger" visar bilden i förstoring, så att man tydligare skall kunna se hur den är uppbyggd. Normalt ser inte människor bilder i sådan förstoring, och bilder som normalt ser bra ut kan se underliga ut vid en så stor förstoring.

Beskrivning	Storlek, bytes (oktetter)	Utseende som man normalt ser bilden	Del av bilden förstörd 10 ggr för att visa hur bilden är uppbyggd.
Linjeritad bild (text i oval), t.ex. eps-format, konverterat till bitmappad bild. Observera taggigheten hos sneda linjer. Anmärkning: Georgia är ett typsnitt gjort för att bli snyggt i låg upplösning utan kantutjämning.	PDF-format: 2 329 GIF-format: 938 TIFF-format: 7 554		

Beskrivning	Storlek, bytes (oktetter)	Utseende som man normalt ser bilden	Del av bilden förstörd 10 ggr för att visa hur bilden är uppbyggd.
<p>Samma bild men kantutjämnad vid konvertering till bitmappad bild. Observera att text blir snyggare i den vänstra bilden, men fulare vid förstoringen till höger.</p> <p>Observera att taggigheten motverkas genom att sätta in punkter med mellanliggande färger i hacken i den sneda linjen.</p> <p>Denna bild blir likadan i GIF-format med en anpassad palett, eftersom antalet nyanser är färre än 256.</p>	<p>Tiff-format: 9 518</p> <p>GIF-format: 2 854</p>		
<p>Samma bild med GIF-format och begränsat till Netscape-paletts 216 färger.</p> <p>Observera att den blå ytan, som hade en av standardfärgerna, inte kräver gittring, men att triangeln, som hade en färg mitt emellan två standardfärger, orsakar gittring för att färgpunkterna, sedda på långt avstånd, skall smälta ihop till rätt nyans.</p>	2 852		
<p>Fotografi med en teknik som tillåter miljontals färgnyanser (24 eller 32 bitar per pixel).</p>	<p>i TIFF-format 43 094</p> <p>Helt omprimerat 397 473</p> <p>i Photoshop-format 53 932</p> <p>i PICT-format utan komprimering 51 310</p>		

Beskrivning	Storlek, bytes (oktetter)	Utseende som man normalt ser bilden	Del av bilden förstörd 10 ggr för att visa hur bilden är uppbyggd.
Samman bild i GIF-format och med en specialanpassad palett av 256 nyanser som passar för just denna bild. Viss gittring finns, se t.ex. de vita punkterna på näsans högersida.	16 750		
Samma bild i JPEG-format med JPEG inställd på låg komprimering, vilket ger bättre bildkvalitet än högre komprimering.	7 663		
Samma bild i JPEG-format med med JPEG inställd på mycket kraftig komprimering. Här ser man komprimeringen främst i suddighet i de fina linjerna, se t.ex. håret till höger om hakan. Vid förstoring 10 ggr ser man även ett större rutnmönster.	1 371		
Samma bild i GIF-format och begränsat till Netscape-paletten, som inte passar särskilt bra för nyanserna i denna bild. Observera att ett gitter av färgpunkter används för att smälta samman till rätt nyans när man ser det på längre avstånd.	8 687		

# 1 Other Application Layer Standards

## 1.1 Introduction

This chapter gives an overview of miscellaneous standards, which are not covered in separate chapters in this book.

## 1.2 Scalable Vector Graphics (SVG)

According to the original World Wide Web standards, documents were sent in HTML format with embedded pictures in either the JPEG or the GIF format. Both these formats are bitmapped formats, they describe a picture by a twodimensional matrix of pixels, with the color given for each pixel.

Bitmapped formats are good for photographs and paintings. But they are not good for pictures whose content can be logically described. For such pictures, they have several disadvantages:

1. Pictures get less and less sharp when you enlarge them. And pictures which look all right on the screen (typical resolution 72-96 pixels/inch) will look blurry when printed on paper (typical resolution 160-2400 pixels/inch).
2. File sizes are often unnecessarily large, which also gives longer download time.

Because of this, the World Wide Web Consortium has developed a standard named Scalable Vector Graphics (SVG). This standard is based on a logical description of the content of a picture, i.e. describing it as consisting of straight and curved lines and polygons, and areas bordered by such lines. Plug ins are available for displaying SVG in web browsers, and future web browsers may have SVG built in.

SVG uses XML to encode the logical description of a picture. For example, the following code defines an ellipse:

```
<ellipse cx="200" cy="300" rx="60" ry="30"
style="fill:none;stroke:#000099;
stroke-width:6;opacity:none" />
```

The full description of SVG can be found in [W3C 2001A] and a tutorial can be found at [Adobe 2002].

## 1.3 XML Schema

An important property of XML is that you can use XML without any syntax specification. The DTD language for syntax specification of XML is not mandatory. This also means that instead of DTD, another syntax specification language can be used. One other such syntax specification language, developed by the World Wide Web consortium, is XML Schema [W3C 2001B].

XML Schema has the following advantages compared to DTD:

1. An XML Schema is itself written in XML.
2. The format is more user friendly, avoids some cryptic encodings in DTD.
3. XML Schema gives more detailed control of the data. It is, for example, possible to define that a value should be an integer, or an integer between 1 and 100.

XML Schema has many similarities to ASN.1.

Here is an example of an XML Schema:

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:simpleType name="positiveInteger">
    <xsd:Restriction base="xsd:integer">
      xsd:minInclusive value="1" /
    </xsd:Restriction>
  </xsd:simpleType>
  <xsd:complexType name = "Person" >
    <xsd:sequence>
      <xsd:element name="name" type="xsd:string" />
      <xsd:element name="age" type="positiveInteger" />
    </xsd:sequence>
  </xsd:complexType>
</xsd:schema>
```

This schema defines a data type "Person" which consists of two elements, a name as a string, and an age as a positive integer.

XML Schema defines also some non-XML formatting. For example, the following XML Schema:

```
<xsd:simpleType name="SimpleListOfIntegers">
<xsd:list itemType="myInteger" />
</xsd:simpleType>
```

specifies data which is sent like this in XML:

```
<SimpleListOfIntegers>1 2 3 5 8 13</listOfIntegers>
```

Note that the syntax separating the integers is spaces, not XML encoding. So XML Schema is in fact a standard specifying a somewhat richer standard than XML alone.

A full XML encoding of a list of integers with all syntax separators in XML format, might be:

```
<complexTypeOfIntegers>
  <integer>1</integer>
  <integer>2</integer>
  <integer>3</integer>
  <integer>5</integer>
  <integer>8</integer>
  <integer>13</integer>
</complexTypeOfIntegers>
```

## 1.4 References

- [W3C 2001A] Scalable Vector Graphics (SVG) 1.0 Specification, World Wide Web Consortium, September 2001. <http://www.w3.org/SVG/>.
- [Adobe 2002] SVG Zone, Adobe 2001, <http://www.adobe.com/svg/>.
- [W3C 2001B] XML Schema. Part 0 Primer (<http://www.w3.org/TR/xmlschema-0/>), Part 1 Structures (<http://www.w3.org/TR/xmlschema-1/>), Part 2 Datatypes (<http://www.w3.org/TR/xmlschema-2/>), World Wide Web Consortium, September 2001,