

# Sammendrag Informasjonsforval tning IT2802

## Innholdsliste

|  |    |
|--|----|
| Sammendrag.....  | 1  |
| Informasjonsforvaltning.....                                     | 1  |
| IT2802.....  | 1  |
| Informasjonsforvaltning - stikkord/sammendrag.....               | 4  |
| 1.0 XML.....   | 4  |
| 1.1 DTD (Document Type Definition).....                          | 5  |
| 1.2 Well formed (Norsk: "Vel-formulerte").....                   | 6  |
| 2.0 Opphavsrett.....   | 7  |
| 2.1 Samarbeid ved skap av verk.....                              | 7  |
| 2.4 Fotografier.....   | 7  |
| 2.5 Internett.....   | 7  |
| 2.6 Arbeidstakers opphavsrett.....                               | 8  |
| 2.7 Kataloger og databaser.....                                  | 8  |
| 3.0 Preservering - "Langtidslagring av digital informasjon"..... | 9  |
| 3.1 OASIS (Open Archival Information).....                       | 10 |
| 3.2 Identifikatorer.....   | 10 |
| 3.3 Kontrollsiffer.....  | 12 |
| 3.4 Navnerom.....  | 13 |
| 3.5 Administrering.....  | 13 |
| 3.6 Identifikatorsystemer.....                                   | 13 |
| 4.0 Metadata.....  | 14 |
| 4.1 MODS.....  | 14 |
| 4.2 Dublin Core (DC).....  | 15 |
| 4.3 MARC (Machine Readable Cataloguing).....                     | 17 |
| 4.4 Generelt om metadata.....                                    | 21 |

|   |    |
|---|----|
| 4.5 Internasjonale regelverk.....                               | 22 |
| 4.6 ISBD: International Standard Bibliographic Description..... | 22 |
| 4.7 AARC2: Anglo-American Cataloging Rules.....                 | 22 |
| 5.0 Emne og Klassifikasjon.....                                 | 22 |
| 5.1 Dewey Decimal Classification.....                           | 23 |
| 5.2 Universal Decimal Classification (UDK).....                 | 24 |
| 5.3 Thesaurus.....  | 24 |

## Informasjonsforvaltning - stikkord/sammendrag

*Njaal Christoffer Almestad Gjerde*

### 1.0 XML

```
<?xml version="1.0" encoding="UTF-8"?>
<oppskrift navn="Loff" forberedelsestid="5 minutter" totaltid="3 timer">
  <tittel>Vanlig loff</tittel>
  <ingrediens mengde="9" enhet="desiliter">Siktet hvetemel</ingrediens>
  <ingrediens mengde="2,5" enhet="teskje">Gjær</ingrediens>
  <ingrediens mengde="3" enhet="desiliter">Lunket vann</ingrediens>
  <ingrediens mengde="1,5" enhet="teskje">Salt</ingrediens>
  <ingrediens mengde="3" enhet="spiseskje">Vegetabilsk olje</ingrediens>
  <ingrediens mengde="2,5" enhet="spiseskje">Sukker</ingrediens>
  <fremgangsmåte>
    <trinn>Bland alle ingredienser og kna grundig.</trinn>
    <trinn>Dekk over med et klede og sett til heving i et varmt
rom.</trinn>
    <trinn>Kna på nytt, legg i en form og stek ved 200 grader Celsius i 1
time.</trinn>
  </fremgangsmåte>
</oppskrift>
```

Eksempel på xml

### Element

```
<greeting> Hello, world.</greeting>
```

Et element har en start og en -slutttag.

### Atributt

```

```

```
<ingrediens mengde="2,5" enhet="teskje">Gjær</ingrediens>
```

Her er src, alt, mengde og enhet attributter.

### Entitet

```
<!ENTITY NavnPåEntitet "Dette er et eksempel på en intern entitet. All
teksten som står innenfor hermetegn vil brukes når vi benytter
entitetsnavnet NavnPåEntitet">
```

En entitet er som en variabel i et programmeringsspråk. Vi kan sette en entitet for så å bare skrive navnet på entiteten senere når vi vil skrive det innholdet vi har satt entiteten til å være.

Man kan også bruke entiteter som ligger i et eksternt dokument, f.eks:

```
<!ENTITY NavnPåEksternEntitet SYSTEM "http://www.eksempel.no/dokument.txt">
```

### 1.1 DTD (Document Type Definition)

Document Type Definition er en måte å definere hvilke elementer, attributter og i tillegg hvilke attributtverdier som er tillatt for hvert element i et XML dokument.

Eksempel1.dtd

```
<!ELEMENT personliste (person*)>
```

```
<!ELEMENT person (navn, foedselsdato?, kjoenn?, foedselsnummer?)>
```

*På andre linje defineres det at et person element må inneholde navn og i tillegg til dette kan det i nneholde foedselsdato, kjoenn og foedselsnummer, men dette er frivillig da det er brukt spørsmålstejn.*

\* = Kan inneholde uendelig mange av denne typen elementer

? = Trenger ikke, men kan inneholde denne type elementer

Om ingen av disse to er brukt, MÅ elementet inneholde det gitte elementet

Et eksempel på en XML fil som oppfyller DTD-en ovenfor:

Eksempel1.xml

```
<?xml version="1.0" encoding="UTF-8">
  <!DOCTYPE personliste SYSTEM "eksempel.dtd">
  <personliste>
    <person>
      <navn> Ola Nordmann </navn>
    </person>
    <person>
      <navn> Ari Behn </navn>
      <foedselsdato> 30.09.1972 </foedselsdato>
      <kjoenn> Mann </kjoenn>
    </person>
  </personliste>
```

**To betegnelser et XML dokument kan ha:**

## 1.2 Well formed (Norsk: “Vel-formulerte”)

For at et XML-skjema skal være “well formed”, skal det følge **alle** syntaksreglene som er gitt i spesifikasjonen. Dette er en lang liste. Noen eksempler:

- Den inneholder kun ordentlig enkodede lovlige unicodetegn.
- Ingen av de spesielle syntakstegnene som > eller & blir brukt til annet enn de spesielle funksjonene de har i XML.
- Start og slutt tagsene er cas-sensitive, de må matche 100%
- Det er kun et rotelement

## Valid (Norsk: gyldig)

For at et XML-skjema skal være gyldig må det inneholde en referanse til et DTD, og følge reglene som er gitt i DTDen. Eksempel1.xml ovenfor vil da være et valid XML-skjema.

## Generelt om markup languages (ML)

### SGML

- Opprettet i 1986

### HTML

- I likhet med papir er HTML optimalisert for visuell bruk
- Laget for bruk på WWW
- Er et subset av SGML
- Fokuserer på ultimat brukbarhet, men ikke gjennbruk
- Har ingen dokument struktur for å aktivere automatikk

### XML

- Er også et subset av SGML
- Opprettet i 1996
- 

### XHTML

- En kombinasjon av HTML og XML

## 2.0 Opphavsrett

### Opphavsmann

Opphavsmann er alltid en fysisk person, den som har skapt verket.

### 2.1 Samarbeid ved skap av verk

- Flere forfattere skriver hvert sitt kapittel i et læreeverk
  - Opphavsmenn til de delene de selv har skapt
- Et samarbeid der de til slutt ikke klarer skille ut hva den enkelte har bidradd med
  - Sameie i opphavsretten
- Bearbeidelsen: Med utgangspunkt i et åndsverk, tilføyes en ny skapende innsats som selv vil være et åndsverk. (f.eks en bok blir oversatt fra et språk til et annet)

### Internasjonalt perspektiv

Lovgivning er et nasjonalt anliggende og derfor begrenset til enkelt-land, selv om handelen med åndsprodukter er verdensomspennende.

### Bern konvensjonen

- Sikrer opphavsrettslig vern i alle land tilknyttet, uansett opphavsmannens nasjonalitet.
- Aka, her kan man være med for å sikre opphavsrett på tvers av landegrenser
- Norge er med

### Eiendomsrett og opphavsrett

### 2.4 Fotografier

Fotografiske verk

-Følger åndsverksloven

Fotografiske bilder

- Følger egne vilkår, mindre vernetid osv enn vanlige åndsverk

### 2.5 Internett

“pekeransvaret”

- Ingen verk taper opphavsrettslig vern selv om de gjøres tilgjengelig på web

Fotografier beholder sin rett som fotografi, og musikk, tekst og tilsvarende opprettholder tilsvarende vern i henhold til sin type

- De ulike komponenter i en webside (bilder, tekst, lyd) har selvstendig vern
- Mulig med vern av siden som en sammensetning (Må tilfredsstillende alminnelige krav til vern)

## **2.6 Arbeidstakers opphavsrett**

- Opphavsretten til en ansatts arbeid påfaller ikke nødvendigvis automatisk til arbeidsgiver med mindre annet er avtalt i ansettelsesavtalen UNNTAK: datamaskinprogrammer
- Journalister og kringkastingsmedarbeidere avtaler som oftest dette i sin ansettelsesavtale

## **Datamaskinprogrammer**

- Metoder, formler og lignende er ikke opphavsrettlig vernet.
- Et programmeringsspråk er (normalt) ikke vernet, men en kompilator vil være vernet.
- Datamaskinprogram kan være vernet som åndsverk
- Det er måten programmet er skrevet på som er vernet, ikke ideer som programmet bygger på

## **2.7 Kataloger og databaser**

- Åndsverksloven §43 verner “[f]ormular, katalog, tabell og lignende arbeid som sammenstiller et større antall opplysninger”
- Opprettet i 1961
- Var da ment for å verne prislister, radioprogrammer, rutetabeller, telefonkataloger osv
- Arbeider som ikke var resultat av noen skapende innsats, og derfor ikke åndsverk
- Krevde større investeringer, derfor krav på en viss beskyttelse
- Vernet varer i 10 år fra utgivelsen og verner “ettergjøring” aka kopiering av hele greia eller kopiering av større deler av den
- Da databaser kom fant man ut at denne loven også dekker databaser (I Norge)



- Bare de nordiske land har katalogvern i dag

### **Eget database direktiv, skiller databaser i to kategorier:**

- Databaser som på grunn av sitt utvalg eller ordning av innhold representerer opphavsmannens egen intellektuelle produksjon, skal gis opphavsrettslig vern som åndsverk, med en beskyttelsestid på 70 år etter opphavsmannens død.
- Databaser som ikke oppfyller kriteriene for opphavsrettslig vern skal gis et særlig vern med en vernetid på 15 år fra tidspunktet for den første utgivelse av databasens innhold.

### **Personvern**

Personregisterloven krever konsesjon(tillatelse/innrømmelse) for å etablere visse personregistre.

“personregister” – personopplysninger organisert slik at de kan gjenfinnes

I datamaskinbaserte samlinger av personopplysninger antas gjenfinningskriteriet å være oppfylt

## **3.0 Preservering - “Langtidslagring av digital informasjon”**

### **Problem:**

Teknologien går så fort. Medietypene og formatene endrer seg. Selv om vi kan tyde runer fra flere tusen år siden kan vi har problemer med å lese en fil skrevet for ti år siden. Hvor lett er det for eksempel å hente ut innholdet ifra en 8 tomers diskett i dag? Dagens datamaskiner har til og med begynt å komme uten cdrom. I tillegg endrer filformatene seg hyppig.

### **Løsning: Preservering**

#### **Preserveringsstrategier:**

- **Refreshing** – Fornyng av lagringsmedium, f.eks overføring ifra en cdplate til en ny
- **Migration** – Overføring ifra et gammelt system til et nytt system. F.eks ifra en filtype til en annen (.txt til .doc). Men kan også være ifra et medium til et annet, f.eks ifra CDROM til ekstern harddisk.
- **Replication** – Reprodusering av filer til et eller flere steder. Utrykt å bare ha en kopi av en fil!

- **Emulation** - Emulering av et gammelt system for å kunne kjøre ting som er avhengig av den typen system
- **Metadata attachment** - Data på en separat fil som inneholder info om: opprettelse, tilgangsrettigheter, restriksjoner, preserveringshistorie, rettighetshåndtering.
- **Trustworthy digital objects**

### **3.1 OASIS (Open Archival Information)**

Arkiv som består av en organisasjon av mennesker og systemer som har godkjent ansvaret for å preservere informasjon og gjøre den tilgjengelig for et designert samfunn.

## **Eksempler på preservering:**

### **Google Books**

Legger inn bøker i fulltekst i sin egen database og tilbyr søking i fulltekst i disse. Drømmen er at man skal kunne søke i bøker også via googlesøk.

### **WaybackMachine (archive.org)**

Lagrer sider fra nettet slik at man kan se på gamle nettsider.

### **3.2 Identifikatorer**

- Navngivning som gir mening
- Navn skal være beskrivende
- F.eks variabelnavn i programmeringsspråk
- f.eks navn på printer

Men, det er forskjell på navn og identifikator:

- **Navn**

- • www.idi.ntnu.no
- • \\printers.idi.ntnu.no\itv254farge
- • takhirov (@idi.ntnu.no)

- • **Identifikator**

- • Ikke nødvendigvis "lesbar" for mennesker; kan like gjerne være
- basert på nummer, el kombinasjon av nummer og tall
- • ISSN - 1500-4538

- • 00000236678976543785
- • 129.241.56.200

Så, et navn er en identifikator. Men en identifikator er ikke nødvendigvis et navn.

## Identifikatorer og navn må kunne

- Genereres manuelt og automatisk
- Være lesbare/prosesserbare for maskiner og/eller mennesker
- Oversettes (f.eks ifra id til lokaliseringsinformasjon)

## Adresse / lokatorer

- Nummer/navn på register
- Plass i minnet
- IP-adresse
- Navn på harddisk ( [C:\](#) )

## Eksempler på identifikatorer som er i bruk:

- ISBN, ISSN
- URL
- Identifikasjon basert på attributter (metadata)
- Bruk av nummer som er unike

## Aspekter ved identifikator systemer

- Syntaks - Hvordan den skal utformes og hvilke egenskaper den skal ha
- Administrering - Hvem og hvordan
- Bruksområde - Hva kan identifiseres/navngis vha identifikatoren
- Systemer og tjenester - Nødvendig for administrering og bruk av identifikatoren

## Syntaks

- Regler for hvordan en identifikator skal kodes og tolkes
- Hvilke komponenter en identifikator kan/skal bestå av
- Hvilke tegn som er tillat og lengde på strengen

## Unike identifikatorer

- Unik betyr at det bare kan finnes en
- Unik, ikke bare i dag, men også i fremtiden = utvidbar, skalerbar

- Navnerom(hierarkiske) kan forenkle generering av globalt unike identifikatorer

## Lesbarhet

- Kan designes for å fremme lesbarhet (for mennesker)
- Bruk av forkortelser og ord i dns
- f.eks idi.ntnu.no
- Lange sekvenser av tall og tegn kan designes for å være lesbare ved hjelp av skilletegn, eks: ISBN, ISSN og SICI

## Smarte / Dumme

### Smarte

- Inneholder info om objektet
- Kan avlede info fra selve identifikatoren

### Dumme

- Meningsløse strenger
- Gir oss ingen info

## 3.3 Kontrollsiffer

Eksempel: ISSN bruker 8 siffret kontrollsiffer

Journal of Digital Information

ISSN: 1368-750?

$$1*8 + 3*7 + 6*6 + 8*5 + 7*4 + 5*3 + 0*2 = 148$$

$$148 / 11 = 13 \text{ (rest 5)}$$

$$11 - 5 = 6$$

ISSN: 1368-750**6**

### 3.4 Navnerom

Navnerom utgjøres av det sett av unike navn vi kan ha i en gitt kontekst

#### Eksempel:

Hierarkisk oppbygd (trestruktur)

### 3.5 Administrering

Identifikatorsystemer er forpliktende og stiller krav til administrering

**Navneautoritet** = Enhet som er autorisert til å opprette og tilordne en identifikator

#### Oppgaver med identifikatorer

- Oppretting / navngiving
- Registrering av info
- Langtidslagring av informasjon
- Delegering av rettigheter

### 3.6 Identifikatorsystemer

- ISO-standarder: ISBN, ISSN, ISMN, ISRN, ISRC, ISWC, ISAN
  - **ISBN** - International Standard Book Number
    - Bøker (og andre monografier)
  - **ISSN** - International Standard Serial Number
    - Serier: tidsskrifter, aviser, årbøker...
  - **ISMN** - International Standard Music Number
    - Trykte musikk publikasjoner (noter ...)
  - **ISRN** - International Standard Technical Report Number
    - Tekniske rapporter (forskningsrapporter, undersøkelser ...)
  - **ISRC** - International Standard Recording Code
    - Audio og musikk video opptak
  - **ISAN** - International Standard Audiovisual Number
    - Audio-visuelle verk (uavhengig av fysisk medium)
  - **ISWC** - International Standard Musical Work Code
    - Musikkverk uavhengig av fysisk manifestasjon
- Forlagsverdenen: SICI, BICI

- Identifikatorsystemer for Internett: URI, URL, URN, The Handle System, DOI

## 4.0 Metadata

Informasjon om et objekt

I utsagnet «prisen er kr. 5,-» er «kr. 5,-» data, mens «prisen» er metadata, ettersom «prisen» beskriver hva slags data «kr. 5,-» er.

### Katalogisering

#### Katalog

- Fortegnelse over en samling (bøker, gjenstander osv)
- Funksjoner som gir brukeren mulighet til å finne ressurser i en samling
- I gamledager brukte man kortkatalogen:
  - Et kort = en innførsel
  - Lett å sette inn nye innførsler
  - Organisert etter tittel, forfatter, emne osv

#### 4.1 MODS

Metadata Object Description Schema. Inneholder de feltene som allerede ligger i MARC, men har muligheten til å legge til flere. Mods har flere og rikere elementer enn Dublin Core. Lettere å bruke enn MARC. Bruker vanlig språk istedenfor tall og dollartegn som variabler for de forskjellige feltene.

## 4.2 Dublin Core (DC)

- Et metadata format
- Er en ISO standard
- 15 elementer (ingen obligatoriske felt, gjenbrukbare, rekkefølge er ikke viktig)

**Metadata skjema. Hadde først 13 felt, ble økt til 15 felt. Disse feltene er gitt, de kan ikke endres, men man kan selv velge hvilke man vil bruke og hvor mange ganger man vil bruke de.**

| Innhold     | Opphavsrett | Manifestasjon |
|-------------|-------------|---------------|
| Title       | Creator     | Date          |
| Subject     | Publisher   | Type          |
| Description | Contributer | Format        |
| Source      | Rights      | Identifiser   |
| Language    |             |               |
| Relation    |             |               |
| Coverage    |             |               |

### Ulemper med DC

- Ingen konsistent måte å definere data på
  - f.eks, language=no, eller language=bokmål
  - f.eks 29.09.88 eller 09.29.88 ?
- Dato elementet, kan være brukt til flere ting, hvordan skal vi vite hva som er blitt ment med det?
  - F.eks kan det være dato på når boken ble skrevet, eller når den ble gitt ut osv.....

## Qualifiers (norsk: kvalifikatorer). En måte å være mer spesifikk på i Dublin Core

- Kan definere feltene mer spesifikk
- f.eks Creator.autor
- f.eks Date.created eller Date.published
- Vi kan også spesifisere hvordan innholdet skal tolkes ved å angi hvilken standard vi bruker
  - `<meta name="DATE.published" scheme="W3CDTF" content="2004-11-7" />`

## Dumb-Down-principle (DDP) (VIKTIG!!)

- Man skal kunne forstå innholde selv om man fjerner alle kvalifikatorene
- Man vil kanskje ha et tap av presisjon, men datene vil forstatt kunne brukes i søk osv...

## Eksempel på DC i HTML

```
<html>
  <head>
    <meta name = "DC.Creator" content = "Engels, F.">
    <meta name = "DC.Title" content = "Kapitalen">
  </head>
  <body>
  </body>
</html>
```

## Eksempel på DC i XML (kan lage egne elementer)

```
<DC>
  <Title>Kapitalen</Title>
  <Creator>Engels, F</Creator>
</DC>
```



### 4.3 MARC (*Machine Readable Cataloguing*)

- Det finnes flere «dialekter»
- Katalogisering
- 
- Machine Readable Catalogging record. En datamaskin skal kunne lese de tog det skal inneholde de samme feltene som et katalog kort ville inneholdt. Bruker tall, dollartegn og bokstav som variabler for feltene. Eks: 210 \$a = tittel. Har utrolig mange forskjellige felt som kan brukes.

#### Felt (MARC)

|     |                                   |
|-----|-----------------------------------|
| 0xx | Koder, numre, klassifikasjon m.m. |
| 1xx | Hovedordningsord                  |
| 2xx | Tittel m.m.                       |
| 3xx | Fysisk beskrivelse                |
| 4xx | Serie                             |
| 5xx | Noter                             |
| 6xx | Emneord                           |
| 7xx | Biinførsler                       |
| 8xx | Serie i annen form                |
| 9xx | Se-henvisninger                   |

**2XX - Tittel**

240 Standardtittel

241 Originaltittel

245 Tittel

246 Parallelltittel

250 Utgave

254 Musikktrykkets fysiske presentasjon

255 Kartografisk materiale

256 Filkarakteristika

260 Utgivelse, distribusjon

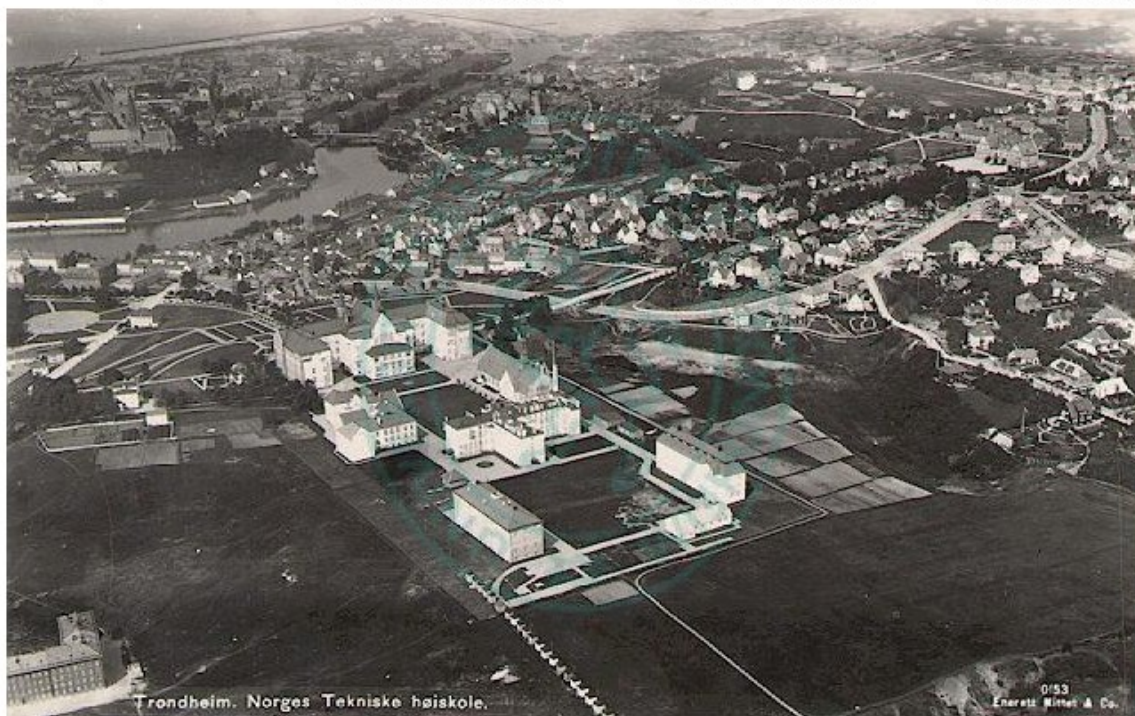
**Delfelter for 245 tittel**

|     |                                |
|-----|--------------------------------|
| \$a | Hovedtittel                    |
| \$b | Annen tittelinformasjon        |
| \$c | Ansvarsangivelse               |
| \$h | Generell materialbetegnelse    |
| \$n | Nummer for del av verk         |
| \$p | Tittel for del av verk         |
| \$w | Sorteringsfelt for delfelt \$a |

**Eksempel på MARC:**

# Flyfoto over Trondheim, Norges tekniske høyskole i forgrunnen.

|                  |  |
|------------------|--|
| Tittel           | Trondheim. Norges Tekniske høyskole.                                   |
| Beskrivelse      | NTH er fotografert fra syd. Store deler av Singsaker er med på bildet. |
| Generell merknad | Prospektkort   |
| Topografisk      | Uhjr. Uhjb   |
| Sted             | Trondheim. Gløshaugen.<br>Trondheim.<br>Norges tekniske høyskole.      |
| Emneord          | flyfoto.   |
| Fotograf         | Mittet & Co.   |
| Format           | H 9.00 x B 14.00   |
| Negativnummer    | F II 4469  |
| Eksemplar        | UBiT, topografisk  |
|                  | VII-Uhj r-004376   |
|                  | Prospektkort s/h   |



|         |  |
|---------|--|
| ID      | UBT-TO-004376  |
| 012 \$e | 2004-08-23   |
| 012 \$k | IJGR   |
| 088 \$a | Uhjr   |
| 088 \$b | Uhjb   |
| 096 \$a | UBIT   |
| 096 \$b | topografisk  |
| 096 \$c | VII-Uhj r-004376   |
| 096 \$f | Prospektkort s/h   |
| 100 \$a | Mittet & Co.   |
| 100 \$c | forlegger  |
| 245 \$a | Flyfoto over Trondheim, Norges tekniske høgskole i forgrunnen.         |
| 300 \$c | H 9.00 x B 14.00   |
| 500 \$a | Prospektkort   |
| 520 \$a | NTH er fotografert fra syd. Store deler av Singsaker er med på bildet. |
| 591 \$a | F II 4469  |
| 651 \$c | Trondheim  |
| 651 \$h | Gløshaugen   |
| 651 \$a | Trondheim  |
| 651 \$a | Norges tekniske høgskole   |
| 687 \$a | flyfoto  |
| 740 \$a | Trondheim. Norges Tekniske høiskole                                    |

## 4.4 Generelt om metadata

### Andre metadata formater

Dublin Core og MARC er bare to eksempler på metadata formater. Andre eksempler er:

- CIMI
- EAD
- GILS
- MODS
- FGDC
- LOM,
- MPEG-7
- MPEG21
- RFC1807

### Metadataformat «kategorier»

| Type 1   | Type 2  |
|--|---|
| <b>Lokalisering</b><br>Enkle formater/Ustrukturerte Robot-generert Proprietært Fulltekst-indeksering | <b>Utvelgelse</b><br>Strukturerte formater Robot pluss manuell input Standarder (under utvikling) Felt struktur |
| Søkemotorer<br>-Google<br>-Yahoo   | Dublin Core<br>RFC1807  |

### Metadata sine funksjoner:

- Søking
- Evaluering
- Identifikasjon og lokasjon
- Utvelgelse
- Basis for informasjon organisering
- Preservering og behandling

## 4.5 Internasjonale regelverk

### 4.6 ISBD: International Standard Bibliographic Description

ISBD er et sett med regler som beskriver et et bredt område av bibliotekmaterialer, innen konteksten til en katalog.

### 4.7 AACR2: Anglo-American Cataloging Rules

- Bygger på ISBD, går videre i detaljering

#### Eksempler:

De gales hus : roman / Karin Fossum. - [Oslo] : Cappelen , 2000 - 371 s. - (Cappelens storbøker). - (839.823[S]). - 1. utg. 1999. - Storbok-utg., 2. oppl. 2002 ISBN 82-02-20107-1 (ib.) : Nkr 149.00

De unges forbund [lydopptak] / Henrik Ibsen ; [regi: Gerda Ring]. - [Melhus] : NRK lydbokforl. ; [Oslo] : Radioteatret , 1999 - 2 plater (CD) (2 t, 23 min) : digital ; 12 cm, i eske. - (Fra Radioteatrets arkiv). - (839.822[S]). - I rollene: Tore Foss et al. ; musikk: Richard Rodgers: Carousel fantasy. - "Produsert og redigert i NRK, Oslo 1952" - Esken ISBN 82-421-0870-6

#### MPEG

- ISO standard for kodet representasjon av lyd og bilde
- MPEG-1 - til MPEG-2
- Jobber med ny standard MPEG-21

## 5.0 Emne og Klassifikasjon

- Klassifikasjon er å dele ett eller annet inn i klasser
- En klasse etableres fordi elementene i klassen deler ett eller flere kjennetegn

Kan skille mellom

- **Klassifikasjonssystemer:** Bruker notasjon
  - Basert på tall, bokstaver etc

- Thesaurus, **emneord-systemer**
  - en (kontrollert) liste av ord som kan benyttes

### **Samme formål, men forskjellig form**

Klassifikasjonssystemer kan være spesifikke for en type informasjon eller et fag, eller annet formål.

- Geografisk informasjon
- Medisin
- Zoologi og botanikk

Eller de kan være generelle og dekke alle former for kunnskap, f.eks Dewey:

### **5.1 Dewey Decimal Classification**

Delt opp i 10 hovedklasser eller kategorier, hver av dem er igjen delt opp i 10 sekundærklasser, som igjen har subdivisjoner.

Dewey er det mest brukte klassifiseringssystemet i verden

- 000 – Computer science, information, and general works
- 100 – Philosophy and psychology
- 200 – Religion
- 300 – Social sciences
- 400 – Languages
- 500 – Science and Mathematics
- 600 – Technology and applied science
- 700 – Arts and recreation
- 800 – Literature
- 900 – History and geography and biography

Den siste klassen er enda ikke definert...

---

Dewey undernivå:

**500** Science

**500** general works on the sciences

**510** mathematics

**520** astronomy

**530** physics

---

**530** Science, physics

530 general works in physics

531 classical mechanics

532 fluid mechanics

533 gas mechanics

---

Det er også mulig å ha et uendelig antall undernivå. Eksempel:

600 Technology (Applied sciences)

630 Agriculture and related technologies

636 Animal husbandry

636.7 Dogs

636.8 Cats

## 5.2 Universal Decimal Classification (UDK)

- Indekserings og gjennfinningspråk
- Nesten likt som Dewey
- Bruker også spesialtegn og kombinasjoner av klassifikasjonskoder for å uttrykke sammensatte emner

## 5.3 Thesaurus

### Problem:

Søkeren har ofte et annet vokabular enn dokumentene han prøver å finne

- Mener det samme men bruker forskjellige ord
  - Da finner vi ikke relevante dokumenter
  - =Problem med recall
- Mener ikke det samme men bruker like ord
- Dokumentene som er funnet er ikke relevante
  - =Problem med presisjon

Tesaurus skal hjelpe deg til å velge riktig emneord

Samme konsept skal sorteres under samme emneord