

Om bruk av semantisk web med mer forvaltningen

Fra papirskjemaer informasjon til elektroniske tjenester og kommunikasjon

Hovedpunkter for forelesningen

- [Gårdagens] problemene og [morgendagens] muligheter
- Semantisk web
 - Hva og hvorfor
 - Eksempel på anvendelser
- Emnekart
 - Kort om ideen og teknologien
 - Noen anvendelser

DRI 2010 H08 0511 Arild Jansen , AFIN

Syntaks, semantikk og grammatikk og ontologi (rep.)

- **Syntaks** er læren om hvordan **ord** settes sammen til større enheter, dvs. setninger og fraser. Ulike språk har *ulike regler* for hvordan dette skjer på både med hensyn til ordstilling og syntaktiske avhengigheter
- **Semantikk** er den grenen av språkvitenskapen som studerer *ordenes betydning*, både enkeltvis og i sammenstilte **strukturer** (setninger).
- **Grammatikk** er studiet av reglene som gjelder et **språk**. Reglene som styrer et spesielt språk kalles språkets grammatikk
- **Ontologi** En spesifisering av hvordan man presenterer objekter, begreper og andre "ting" som eksisterer i et interessefelt og relasjonene mellom disse.
<http://home.hio.no/~toreh/blog/arkiv/docs/Ontologien%20for%20Kulturnett%203.doc>

DRI 2010 H08 0511 Arild Jansen , AFIN

Data, datadefinisjon og metadata (rep.)

Dataelement: Enhet av data som er udelelig, f eks. f. navn, e.navn, p.nr, telefonnr. ...

- **Datadefinisjon:** *Type og formatbeskrivelse av et dataelement*
- **Metadata:** Data om dataelementer, inkl. datadefinisjon, dataeierskap, tilgangsrettigheter,
 - Metadata brukes både i tradisjonelle (relasjons) databaser og andre typer databaser, f eks. XML-baserte databaser.

Metadata omfatter mer enn [rene]datadefinisjoner :

- Bidrar til å opprette logiske sammenhenger, der de ikke finnes fra før
- Bidrar til å gi opplysninger entydige egenskaper
- Bidrar til å knytte informasjon til informasjonens tilhørende sammenheng

DRI 2010 H08 0511 Arild Jansen , AFIN

Problembeskrivelsen

- "Gårdagens" forvaltningen var preget av (blant annet)
 - Papirbaserte skjemaer for søknader for innrapportering saksbehandling hvor skjemaets formål/eier samt forklaringer av felter angir hva slags opplysninger det dreier seg.
 - Skjemaene ble behandlet av mennesker, og lagt inn i databaser.
 - Tabellbaserte databaser, hvor strukturen blir beskrevet i datakatalogene
 - En-veis informasjonsformidling basert på brosjyrer/meldinger.,.
 - Liten grad av tverretattlig og tverrsektoriell samhandlingSe f eks. <http://www.noreg.no/samkart/default.asp?kart=03>

DRI 2010 H08 0511 Arild Jansen , AFIN

Behov for datamaskinlesbare beskrivelser for utveksling av opplysninger (data) mellom systemene

Nettbaserte tjenester

- Papirskjemaer erstattes med skjermbilder som mater datasytemene med data
 - Skjermbilder innebærer at koblingen mellom skjemaets felter og dataene "forsvinner".
 - Tekstlige forklaringer forstås ikke entydig av datamaskinene og må suppleres med *formaliserte, entydige beskrivelser* av dataene gjennom metadata på en standardisert form
- Elektronisk samhandling innebærer
 - Data kan brukes av mange ulike systemer
 - Dette krever at dataene må kunne fortolkes og behandles korrekt på tvers av systemene

DRI 2010 H08 0511 Arild Jansen , AFIN

Litt om organisering av [tekstlig] informasjon (rep)

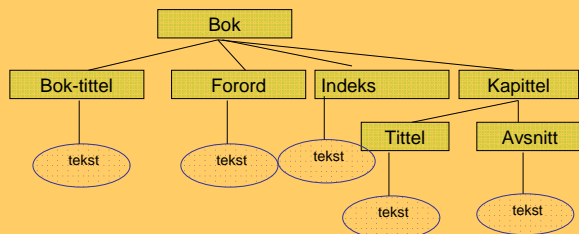
Et tekstlig dokument kan (blant annet) karakteriseres ved:

- *Innhold*: Hva teksten uttrykker/formidler,
 - Eks: Roman, dikt, fagstoff, lovtekst, offentlig rundskriv, brosjyre,
- *Struktur*: Måten innholdet er organisert,
 - Eks. bok, forord, kapitler, avsnitt, nummerering, referanser,...
- *Form/utseende* (Layout, "design")
 - Skrifttyper/størrelser, farger/grafikk, sidestørrelse, spalter, bokser,...
- *Koding og språk* Disse er ikke uavhengige av hverandre

DRI 2010 H08 0511 Arild Jansen , AFIN

XML - Extensible markup language

- XML kan beskrive struktur og innhold
- Eks. en beskrive struktur i en bok



DRI 2010 H08 0511 Arild Jansen , AFIN

Eksempel på XML-kode, inklud HTML-kode

```
<?XML versjon="1.0" Encoding="ISO-8859-1"?>
<bok>
  <description>
    <title> Fra kjernen og ut, fra skallet og inn </title>
    <author>
      <first-name> Gerhard </first name>, <Last-name Skagestein</last-name>
    </author>
  </description>
  <body>
    <Forord > I denne boka vil jeg...</forord>
    <chapter title="Innledning" >
      <p> I dette kapitlet ser vi på .....
    </chapter >
    <Chapter title="systemutviklingsprosessen"> .....</chapter>
  </body>
</book>
(fra Skagestein, kap. 17 forenklet. Fargene er for å synliggjøre teksten)
```

DRI 2010 H08 0511 Arild Jansen , AFIN

Litt mer om XML og semantic web

- Extensible Markup Language (XML) er enkelt språk for å beskrive dataformater (struktur og innhold, og ikke layout-utseende).
- XML kan brukes til å utveksle data mellom systemer
- XML kan brukes til å lagring av semistrukturerte data, f eks. boktekster, web-sider, ...
- XML har en strengere syntaks (grammatikk) enn HTML

Eksempler på bruk av XML ved UiO:

<http://www.uio.no/studier/emner/jus/afin/DRI2010/h08/pensumliste.xml>

Se mer: <http://www.w3.org/XML/>

<http://www.w3.org/2001/sw/>

<http://www.w3.org/2007/03/layerCake.png>

DRI 2010 H08 0511 Arild Jansen , AFIN

Hvorfor er XML viktig i forvaltningen

- Det er ønskelig å effektivisere og forbedre samhandlingen mellom offentlige etater
- Det er et spesielt ønske om å forenkle, samhandling mellom myndigheter og borgere, og med næringslivet
- Vi skal slippe å sende samme data til mange offentlige etater
 - Det er ønskelig å unngå at samme data lagres mange steder
- Målsetningene i St, mld 17 (2006-2007) og spesielt MinSide forutsetter bedre samhandling
- Dette krever standardisering og økt tilgjengelighet av grunddata og andre felles data
- Flere "sentrale" register og tjenester for å oppnå dette er [Brønnøysundregistrene](#) og spesielt [Enhetsregisteret](#) og [Oppgaveregisteret](#).

DRI 2010 H08 0511 Arild Jansen , AFIN

Samordningsarbeidet av næringslivets innrapportering til det offentlige

- Opprettelsen av [Enhetsregisteret i 1995](#): De samarbeidende registrene: Arbeidsgiverregisteret, Merverdiavgiftsmanntallet, Statistisk Sentralbyrås bedriftsregister, Foretaksregisteret
- Enhetsregisteret tildeler alle nye bedrifter et ni-sifret [organisasjonsnummer](#), felles for alle offentlige næringslivsregistre.
- [Oppgaveregisteret](#) ble opprettet i 1997. Hovedoppgaven er å holde løpende oversikt over næringslivets oppgaveplikter til det offentlige, og finne muligheter til samordning og forenkling.
 - Oppgaveregisteret sørger for at etatene får kunnskap om hvilke data som allerede finnes hos andre, og opplysningene kan dermed utveksles ved hjelp av organisasjonsnummeret.
- Oppgaveregisteret definerer *meningsinnholdet* i forskjellige spørsmål i statlige spørreskjemaer: nasjonal database for datadefinisjoner.
 - Hvert dataelement blir kategorisert, definert, gitt et navn og tildelt et nummer. Samordning av offentlige registre

Se: <http://www.brreg.no/samordning/>

DRI 2010 H08 0511 Arild Jansen , AFIN

Hvordan brukes dette i praksis Eksemplet ALTINN

- I da skjer en rekke innrapporteringen gjennom den felles tjeneste **ALTINN**, som benytter seg av hos Brønnøysundregistrene
- Se <http://www.altinn.no/>

DRI 2010 H08 0511 Arild Jansen , AFIN

Hva er emnekart ?

- **Emnekart** ([eng. Topic Maps](#)) er en standard for representasjon og utveksling av strukturert og semistrukturert informasjon. Standarden startet som en generalisering av konseptene rundt [stikkordregister](#), [ordbøker](#) og [synonymordlister](#), men har vist seg å ha stor nytteverdi i en rekke andre sammenhenger.
- Emnekart brukes til informasjonsgjenfinning og navigering i informasjon. Standarden er en semantisk teknologi, som kan sammenlignes med [W3C](#)-standardene for [semantisk web](#).
- Emnekart har i tillegg en mekanisme (scope) for å angi kontekst for informasjonen, XMT (XML Topic Maps) gir en syntaks for å beskrive emnekart (ISO/IEC 13250)

Fra <http://no.wikipedia.org/wiki/Emnekart>

DRI 2010 H08 0511 Arild Jansen , AFIN

Kjernen: *emner, forekomster og assosiasjoner* (TAO *Topics, Occurrences og Associations*).

Emnekart representerer en subjekt (tema) - orientert perspektiv

- Et emnekart består av et sett med *emner* av forskjellige typer (innefor et tema)
- *Assosiasjoner* knytter emnene sammen i en grafstruktur
- *Forekomster* er data om emnet
- Eksempel, se ontologi for kulturnett.no:
<http://home.hio.no/~toreh/blog/arkiv/docs/Ontologien%20for%20Kulturnett%203.doc>
- Mer om emnekart, se http://www.topicmaps.com/tmc/presentation.jsp?conf=TM2008&id=Steve_Pepper~TM2007~Everything_is_a_subject

DRI 2010 H08 0511 Arild Jansen , AFIN

Noen eksempler på bruk av emnekart i forvaltningen :

- <http://www.norge.no/>
 - [Forbrukerrådets informasjonsportal](#)
 - [Utdanningsdirektoratet](#)
- Se også
- http://skolenettet.no/moduler/templates/Module_Overviw.aspx?id=22220&epslanguage=NO
- Mer om emnekart på

DRI 2010 H08 0511 Arild Jansen , AFIN

W3C: XML ++ versus Emnekart /Topic Map Noen elementer

- Emnekart (Topic Maps) er en annen Semantic Web-teknologi
- Grunntanken er den samme: Hvordan overføre "betydningen" av data mellom systemer?
- Forskjeller:
 - Forskjellige eiere (W3C vs ISO)
 - Forskjellig utgangspunkt (logikk vs informasjonsgjenfinning)
 - Forskjellig formål (generelt dataformat vs mer spesifikt "stikkordsregister på steroider")

DRI 2010 H08 0511 Arild Jansen , AFIN