Metatilnærming til metadata

Tore Hoel
eStandard-prosjektet
16. aug 2004

Referanser og mer informasjon

- Aalberg,T (2003): Supporting Relationships in Digital Libraries
- Thomas Baker, Christophe Blanchi, et al. Principles of Metadata Registries, DELOS whitepaper, 2003, http://delos-noe.iei.pi.cnr.it/activities/standardizationforum/Registries.pdf
- Understanding metadfdata, NISO, 2004, ISBN 1-880124-62-9 http://www.niso.org
- JISC IE Metadata Schema Registry Project - http://www.ukoln.ac.uk/projects/iemsr/

Veien videre...

- Større vekt på arkitektur
- Er ISO/IEC 11179 løsningen?
  - Løpende diskusjon om hva som er den beste teknologi for å uttrykke rike semantiske sammenhenger, jf. semantisk web/RDF/OWL... og TopicMaps
- Registre er løsningen - som **organisatorisk** virksomhet
  - Desentral, distribuert virksomhet forutsetter sentrale tiltak (registre)
  - Konsekvenser for organisering av programmet
  - Konsekvenser for nasjonal organisering av arbeidet med semantisk interoperabilitet
  - Ikke bare et spørsmål om forretningsprosesser - kultur- og kunnskapssektorene må sterkere med!
  - Endringer av bilbiotekskulturen???

ISO/IEC 11179-1:1999: **Metadata Registries standard**

- Gjennomgang av 11179, basert på sides fra WG2-convenor i SC32
- Se også UKOLN-presentasjon av ISMSR

MOF = Meta-Object Facility Specification fra Object Management Group

While ISO 11179 is emphasized because of considerable legacy work and MOF is considered more useful currently, the Semantic Web technologies of RDF and OWL have much more mature and capable data models for semantic integration and interoperability and provides a convergence of the four data communities (document, Web, database, and programming).

(*Suggestions provided to the AIC Governance and Components Subcommittees, August 25, 2003, (US Government organisasjon)*

Disclaimer

AP1 - revurdering av mandatet?

- Metadataformater - enes om et nasjonalt utvekslingformat
- Anbefale framtidig nasjonalt utvekslingsformat
- Anbefale format(er) til bruk for innholdsprodusenter

STATISK VERSUS DYNAMISK TILNÆRMING TIL METADATA

Mål: Interoperabilitet

Meta til hva - avgrensningsproblemet: *Hva skal være på hyllene i det digitale bibliotek?*

Står vi alene med problemstillingene?
- OASIS - business-verdenen
- Brønnøysundregisterets EU-seminar i juni om Semantisk interoperabilitet
- Aalberg om utfordringene for digitalt bibliotek (doktorgradsarbeidet fra 2003)

Må se vårt arbeid i en større nasjonal sammenheng!

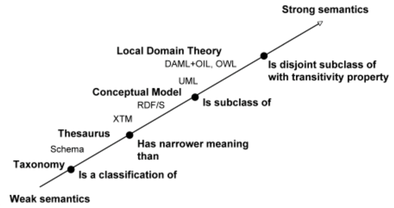
Hvor går vi?

Mot standardisering på spesifikke **metadata skjema**?

I retning av den **semantiske web**, der målet er å øke interoperabiliteten av metadata?

Skifte i fokus på grunn av at Internett er blitt allemannseie

Vil **Arkitektur for elektronisk samhandling i offentlig sektor** ha noen betydning for NDB-arbeidet?



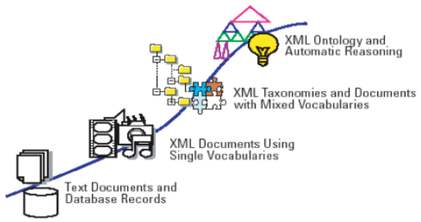
Precisely because metadata language communities now find themselves on the common ground of a global network – an Internet Commons – attention is shifting from *refining the metadata languages of particular communities* towards *defining design principles, data structures, and algorithms that will facilitate interoperability among them*. (DEMOS-hvitboka)

Sentralisert til desentralisert dataforvaltning: Vi må ta innover oss hva Internett betyr!

Mer vekt på **data-semanitikk**

Data Semantics is the discipline that facilitates the delivery of the right information based on the requestors requirements. Semantic agreement within Community's of Practice is essential to facilitating meaningful and effective data exchange.

The Smart Data Continuum

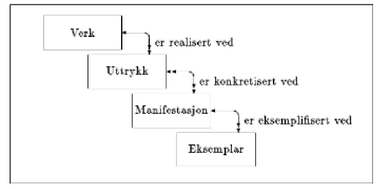


The trend is to put the "smarts" in the data, not in the applications.

Metadata er...

Metadata er data om data
- Ikke noe spesielt - bare vanlige, hverdagslige data
- Finnes i dokumenter, meldinger, bilder, lydstrømmer, videoer, datasett, nettsteder osv.
- Gir informasjon om dataene eller ressursen....

Hva er det vi beskriver? **FRBR-modellen**



Hva skal vi bruke metadataene til?
- Fremmer klarere forståelse
- Hjelper oss å finne informasjon
- Fremmer mer effektiv bruk og gjenbruk av informasjon

Begreper vi bruker (fra DELOS whitepaper, 2003)

Meningsbærende enhet som er sammenligbar og mulig å mappe med elementer i andre standarder

Metadata-elementer
- **Attributt-rom** (LOM, DC osv.)
- **Verdi-rom** (f.eks. ontologier, classifikasjoner, kontrollerte vokabularer, taksonomier)

Metadata-instans

Interoperabilitet i flere lag
- Nivå 3: metadata-elementer
  - Attributt-rom
  - Verdi-rom
- Nivå 2: Representasjon
  - XML, RDF, DAML-OIL
- Nivå 1: Transport og utveksling
  - http GET, AOI PMH

Interoperabilitet gjennom standardisering av...
- Grammatikken for metadata-språket
  - Datamodeller
  - DC vs. IEEE LOM (flat vs. hierarkisk modell)
- Måten å Identifisere elementer
- Skjema (på nivå 2)
- Applikasjonsprofiler
- Semantisk web
- Metadata Crosswalk
  - En crosswalk er en mapping av elementene, semantikk og syntaks fra ett metadataopplegg til et annet
- Modell for register-tjenester
  - Metadata-elementene i kontekst: ISO/IEC 11179 Metadata Registry
  - Oppslagsverk om metadata
  - In terms of this model, metadata registries are applications that use metadata languages (Layer 3) in a form processable by machines (Layer 2) in order to make those languages available for use by both humans and machines.
  - Ulike foki

En "**organisert**" tilnærming til metadata

Hvis vi organiserer metadataene - hvordan vil dette fremme **bedre dataforvaltning**?
- Gir en konsistent modell for å "fange" data om data
- Registrert ett sted
- Gir detaljer om hvordan dataene er ordnet
- Lenker direkte til online data
- Reduserer risikoen for dublisering av datasamlinger

Hvilken rolle kan et **dataregister** spille?
- Et sted for å finne nødvendige fakta for å beskrive, få oversikt over, analysere og klassifisere data
- Støtter utviklingen og utberedelse av data standarder
- Er en kilde for velformede dataelementer og verdirom
- Fremmer gjenbruk av data i forvaltningssystemer og i andre sammenhenger
- Gjør det mulig å sammenligne, dele og integrere data
- Tjener som en sentral kilde for informasjon om metadata som brukes av ulike fagsystemer

Metadataregistre

Oppslagsverk over metadataspråk

Mål med metadataregistre (JISC IE Metadata Schema Registry)
- Fremme eksisterende metadataskjema-løsninger
- Økt interoperabilitet mellom skjema som et resultat av gjenbruk på tvers av applikasjoner
- Mindre dobbeltarbeid for utviklere
- Administrert utvikling av skjema
- Oppmuntre til harmonisering mellom "konkurrerende" standarder
- Mer spesifikt...
  - Vise metadata vokabularer og termer
  - Vise/velge (og gjenbruke) vokabularer, termer
  - Verifisere opphav/status til vokabularer og termer
  - Navigere i relasjonene mellom termer
  - Mapping, avledninger (inferencing)
  - Peke til relaterte ressurser
    - retninglinjer, veiledninger
    - bindings
    - transformasjoner

Bruker**grupper**
- Informasjonssøkere
- Utviklere av metadataskjema
- Utviklere av andre registre som ønsker å forsikre seg om at deres registre inneholder autoritative skjema
- Vedlikeholdere av metadatastandarder
- Utviklere av programvare som kan bruke registrerte skjema for å konfigurere metadataverktoy, flette metadata fra ulike kilder, eller konvertere data fra ett format til et annet

Bruk**stilfeller**
- En som skal klassifisere ressurser ønsker å vite hva som er god praksis på et bestemt område
- En gruppe informasjonstilbydere ønsker å harmonisere metadata seg imellom
- Informasjonstilbudere trenger å oversitte deres metadata til et felles format
- En utvikler ønsker å lage et skjema som gjenbruke elemener i størst mulig grad
- En programvareutvikler ønsker metadataverktøy som oppdaterer deres systemer automamtisk
- Ti år fra nå ønsker et arkiv å tolke og konvertere metadataposter fra 2004
- Kinesere ønsker å lese og bruke metadata utviklet i Norge
- Flere brukstilfeller - se Usage Scenarios for the JISC IE Metadata Schema Registry, http://www.ukoln.ac.uk/projects/iemsr/