

Forslag til metadatamodell for åpne institusjonelle arkiv ved universitetene i Norge.

Arne Jakobsson, Universitetsbiblioteket i Oslo, E-mail: p.a.jakobsson@ub.uio.no

Morten Erlandsen, USIT, E-mail: morten.erlandsen@usit.uio.no

Dato: 2005-04-06.

Oppdatert 2006-05-02

Inledning

Målet med en felles standard for metadata er:

- en felles løsning for import/eksport av metadata mellom Frida (universitetenes system for forskningsdokumentasjon) og åpne institusjonelle arkiv (repositories)
- en felles søketjeneste for åpne institusjonelle arkiv (repositories) i Norge

Metadatamodellen kan med fordel anvendes av andre institusjoner i Norge som bygger opp åpne institusjonelle arkiv.

Overordnet vedtak

Åpne Institusjonelle arkiv skal:

- registrere metadata i følge Dublin Core (DC)
- støtte Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH)

På lokalt nivå vil videreutviklete metadata med all sannsynlighet fortsatt bli laget for sine spesifikke formål og sammenhenger.

Metadatamodellen skal ses som en veiledning for en felles løsning for import/eksport av metadata mellom Frida (universitetenes system for forskningsdokumentasjon) og åpne institusjonelle arkiv (repositories) og en eventuell nasjonal søketjeneste, ikke slik at innholdet nødvendigvis må uttrykkes på samme måte i den lokale tjenesten.

Ettersom en felles/nasjonal søketjeneste ikke skal overta de lokale tjenestenes formål og anvendelse, passer det bra å holde antall felt og søkefunksjonalitet nede. Vi har derfor valgt å gå ut fra Dublin Cores femten felt. OAI går inn for *simple unqualified* Dublin Core, men tillater andre applikasjons- og domenespesifikke format, så lenge de struktureres som XML-data med et tilhørende XML-skjema (OAI går inn for den felles utarbeidede syntaksen i XML-skjemaet http://www.openarchives.org/OAI/2.0/oai_dc.xsd).

Metadatamodellen är baserad på Den norske katalogkomiten dokument DMCI

Metadatatermer http://www.nb.no/katkom/dublincore_metadatatermer.html og har även inspirerats av Metadatamodell för Uppsök– Sverige <http://www.svep-projekt.se/masters-theses/Metadatamodell/>.

Metadatamodell

1. Opphavsmann (forfatter)

Navn	Opphavsmann (forfatter)
Tilsvare i DCMES	Creator
Definisjon	En enhet som er hovedansvarlig for ressursens intellektuelle innhold.
Syntaks	Eksportere flere opphavsmenn i separate tagger. Syntaks for personnavn: "Etternavn, Fornavn" eller "Etternavn, Initial.
Obligatorisk	Nei
Repeterbar	Ja
Best practice for høsting	-
Anvendelse	For søking og presentasjon.
Kommentar	Eksempler på en opphavsmann er en person, en organisasjon eller en tjeneste. Vanligvis angir man her opphavsmannens navn.

2. Bidragsyter

Navn	Bidragsyter
Tilsvare i DCMES	Contributor
Definisjon	En enhet som har bidratt til ressursens innhold.
Syntaks	Eksportere flere bidragsytere i separate tagger. Syntaks for personnavn: "Etternavn, Fornavn" eller "Etternavn, Initial.

Obligatorisk	Nei
Repeterbar	Ja
Best practice for høsting	-
Anvendelse	For søking og presentasjon.
Kommentar	Eksempler på en bidragsyter er en person, en organisasjon eller en tjeneste. Vanligvis angir man her bidragsyters navn.

3. Tittel

Navn	Tittel
Tilsvare i DCMES	Title
Definisjon	Et navn gitt ressursen
Syntaks	-
Obligatorisk	Ja
Repeterbar	Ja
Best practice for høsting	Eksportere ev. alternative titler i separate tagger. Hoved- og undertittel eksporteres i en tagg. Anvend kolonseparator (" : ") mellom hoved- og undertittel.
Anvendelse	For søking og presentasjon. Presenteres som hyperlink til fulltekst i brukergrensesnittet fra den første tittelen om flere titler.
Kommentar	Tittelen er vanligvis navnet som ressursen formelt er kjent under. Både tittel og undertittel skal angis

4. Identifikator

Navn	Identifikator
Tilsvare i DCMES	Identifiser

Definisjon	En entydig henvisning til ressursen i en gitt kontekst.
Syntaks	Identifiserer ressursen ved hjelp av en streng eller et nummer i samsvar med et formelt identifikasjonssystem. Anbefalt praksis: Uniform Resource Identifier (URI).
Obligatorisk	Ja
Repetierbar	Ja
Best practice for høsting	Betjenes av NBs URN-tjeneste
Anvendelse	For presentasjon.
Kommentar	Link til fulltekst, alternativt til det forsatsblad fra hvilket det lenkas til fulltekstfil/-ene

5. Innholdsbeskrivelse (Abstract/sammendrag)

Navn	Innholdsbeskrivelse (Abstract/sammendrag)
Tilsvarende i DCMES	description
Definisjon	En redegjørelse for ressursens innhold.
Syntaks	-
Obligatorisk	Nei
Repetierbar	Ja
Best practice for høsting	Ved både et engelsk og et norsk sammendrag eksporteres disse i separate tagger.
Anvendelse	For presentasjon og søking
Kommentar	En beskrivelse av ressursens innhold kan omfatte: Et sammendrag, en innholdsfortegnelse, en referanse til en grafisk framstilling av innholdet eller en uformell beskrivelse av innholdet, men er ikke avgrenset til dette.

6. Utgiver

Navn	Utgiver
Tilsvarende i DCMES	publisher
Definisjon	En enhet som er ansvarlig for å ha gjort ressursen tilgjengelig.
Syntaks	-
Obligatorisk	Ja
Repetierbar	Ja
Best practice for høsting	
Anvendelse	For søking og presentasjon. I den grad innholdet i feltet presenteres på en ensartet måte, kan informasjonen også benyttes til browsing mellom utgivere.
Kommentar	Eksempler på en utgiver er en person, en organisasjon eller en tjeneste. Vanligvis angir man her utgiverens navn. Normalisering av utgiver ligger i set- specification.

7. Stikkord og emne

Navn	Stikkord og emne
Tilsvarende i DCMES	Subject
Definisjon	Et uttrykk for ressursens emnemessige innhold.
Syntaks	-
Obligatorisk	Nei
Repetierbar	Ja
Best practice for høsting	Eksportere stikkord/emne i separate tagger.
Anvendelse	For presentasjon og søking
Kommentar	Fritt valgte stikkord/emne som beskriver arbeidets faginnhold.

8. Ressurstype

Navn	Ressurstype
-------------	--------------------

Tilsvarende i DCMES	Type
Definisjon	En egenskap ved ressursen eller en sjangerbetegnelse som beskriver ressursens innhold..
Syntaks	
Obligatorisk	Ja
Repetierbar	Ja
Best practice for høsting	
Anvendelse	For presentasjon og browsing
Kommentar	Verdi velges fra DCMI Type Vocabulary [DCMITYPE] . http://dublincore.org/documents/dcmi-type-vocabulary/ Norsk oversettelse: http://www.nb.no/katkom/dublincore_metadatatermer.html#DCMI_typer

9. Dato

Navn	Dato
Tilsvarende i DCMES	Date
Definisjon	En dato knyttet til en hendelse i levetiden til en ressurs.
Syntaks	ISO8601
Obligatorisk	Ja
Repetierbar	Nei
Best practice for høsting	Uttrykkes numerisk ÅÅÅÅ.
Anvendelse	For søking og presentasjon og browsing på år.
Kommentar	Dato knyttes vanligvis til en ressurs' tilblivelse (se tittelbladet). Dato oppgis slik det er angitt i ISO 8601 [W3CDTF] i formen ÅÅÅÅ-MM-DD.

10. Språk

Navn	Språk
Tilsvarende i DCMES	Language
Definisjon	Språket som ressursens intellektuelle innhold hovedsakelig er skrevet på.
Syntaks	Uttrykkes i overensstemmelse med ISO639-2 - Tags for the identification of language, 3 bokstav-versjonen (for eksempel nor og eng).
Obligatorisk	Ja
Repeterbart	Ja
Best practice for høsting	-
Anvendelse	Anvendes til søking
Kommentar	-

11. Dekning

Navn	Dekning
Tilsvarende i DCMES	Coverage
Definisjon	Utstrekningen til eller omfanget av ressursens innhold.
Syntaks	ISO 3166
Obligatorisk	Nei
Repeterbart	Ja
Best practice for høsting	-
Anvendelse	Anvendes til søking
Kommentar	Dekning inkluderer vanligvis uttrykk for rom (et geografisk navn eller geografiske koordinater), uttrykk for tid (navn på en tidsperiode, dato, eller tidsintervall) eller jurisdiksjon (som for eksempel en navngitt administrativ enhet). Anbefalt praksis: Velg en verdi fra en kontrollert liste. Når det er mulig, velges navn på steder eller tidsperioder framfor numeriske identifikatorer som koordinater eller tidsintervall.

12. FELLES EMNEHIERARKI

Som supplement til de frie stikkordene foreslås at det anvendes et felles emnehierarki basert på f.eks. Norsk inndeling av vitenskapsdisipliner.

Et felles emnehierarki implementeres som OAI Sets.

Et felles emnehierarki tjener to hensikter

- Å muliggjøre selektiv høsting (jfr: <http://www.openarchives.org/OAI/openarchivesprotocol.html#ListRecords>).
- Å berike metadatamodellen med et element som skal gjøre det mulig for brukere å browse/navigere med hensyn til emne kategorisering.

Flere SetSpecs leveres i separate tagger.

13 Metadata i tillegg til DC

Faste vokabularer legges i set-specifications

Vedtakk:

- 1) Opplysninger om universitet legges i set-specification
- 2) Opplysninger om dokumenttype legges i set-specification.