

Prosjekt- og utviklingsmidler for 2015

Mapping mot norsk webDewey (Ref #684d0eaa)

Søknadssum: **2 086 000**

Varighet: **Toårig**

Kategori: **Innsatsområder**

- Nasjonalbibliotekets digitale tjenester som grunnlag for nye tilbud

Opplysninger om søker

Organisasjonsnavn / nr

Universitetsbiblioteket i Oslo / 971035854

Postboks 1085 Blindern

0317 Oslo

<http://www.ub.uio.no/>

Institusjonens leder

Bente Andreassen

Ansvarlig kontaktperson (søker)

Unni Knutsen

22855673/95777978 / Unni.Knutsen@ub.uio.no

Mål for prosjektet

Gjennom prosjektet vil Humord og Realfagstermer bli mappet til norsk webDewey. Erfaringene fra mappingen i det foreslåtte prosjektet vil ha høy overføringsverdi for andre bibliotek som planlegger å mappe sine vokabular mot webDewey.

Resultatet av mapping av Humord/Realfagstermer mot Dewey-systemet vil bli tilgjengeliggjort som åpne, lenkede data i et standardisert format (SKOS/RDF) og dermed som rådata for andre utviklere.

Mappingen av Humord/Realfagstermer til norsk webDewey vil gjøre tesaurusen mer tilgjengelig og synlig, også for folkebibliotekene.

Mappingen av termer fra humanistiske fag, samfunnsvitenskapene, real- og teknologifag vil gi norsk webDewey økt terminologitilfang. Antall bygde numre i webDewey med relevans for norske forhold vil øke. Økt tilgang på termer og bygde numre vil gi flere emnemessige innganger til Dewey og lette klassifikasjons- og emneordsarbeidet i bibliotekene.

En norsk webDewey tilkoblet emneordssystem som Biblioteksentralens emneord, MeSH og Humord/Realfagstermer vil bidra til at norsk webDewey får økt funksjonalitet og relevans.

Mapping mellom Humord/Realfagstermer og Dewey-systemet vil gi forbedret emnesøk i sluttbrukerverktøy.

De ulike oversettelsene av Deweys desimalklassifikasjonssystem samt tilmappede vokabular danner grunnlag for å utvikle flerspråklige emnesøk i sluttbrukersystemene. Universitets- og høgskolebibliotekene har et økende antall utenlandske studenter og ansatte. Flerspråklige emneinnganger er dermed svært

Prosjekt- og utviklingsmidler for 2015

viktige i denne sektoren. Inneværende prosjekt vil legge et grunnlag for økt verbal navigasjon via Dewey-systemet, og samtidig gi et mer fullstendig resultat ved navigering i tesaurusstrukturen.

Bibliotekenes metadata utveksles mer over landegrensene enn tidligere og det er derfor viktig å kunne mappe eksempelvis engelskspråklige og norske termer. Store deler av de metadataene som mottas i e-bok-pakker er utstyrt med Dewey-nummer og emneord på engelsk. En mapping via Dewey vil gi norskspråklige søkeinn ganger til disse dokumentene. Dette vil radikalt forbedre gjenfinnbarheten til dokumentene.

Prosjektet vil bidra til økt samarbeid rundt arbeidet med emneord og klassifikasjon.

Det er et mål at det skjer en overføring av kompetanse mellom data-/utviklerpersonale og bibliotekarene som deltar i prosjektet. Vi vil også ha nytte av kunnskaps- /erfaringsutveksling med andre miljøer som jobber med terminologi og søkeinn ganger, for eksempel Norsk språkbank eller Språkrådet. Vi ønsker også å dele våre erfaringer i form av konferansebidrag og faglige artikler.

Prosjektbeskrivelse

Universitetsbiblioteket i Oslo (UBO) fikk i 2014 prosjektmidler fra Nasjonalbiblioteket (NB) for å delta med NB i et forprosjekt som skal gi beslutningsgrunnlag for hvorvidt det skal etableres en generell, norsk tesaurus med utgangspunkt i Humord. Dette arbeidet skal resultere i en rapport som de to institusjonene vil behandle i 2015.

Deltakere i indekseringssamarbeidet rundt Humord per i dag er universitetsbibliotekene i Oslo, Bergen og Tromsø samt biblioteket ved Senter for studier av Holocaust og livssynsminoriteter. Også Ibsenbibliografien bruker Humord som grunnlag for sin emnebeskrivelse. HumSam-biblioteket ved Universitetsbiblioteket i Oslo koordinerer samarbeidet. Dersom NB og UBO ønsker å samarbeide om å utvikle tesaurusen videre, er det naturlig at NB tar rollen som koordinator.

UBO mottok også midler for å utrede metodikk for mapping av Humord til norsk webDewey. Dette arbeidet er i full gang og bygger videre på de metoder som er utviklet i tidligere prosjekter i Realfagsbiblioteket.

Basert på de erfaringer som er gjort i tidligere prosjekter, søkes det om midler til å planlegge og gjennomføre mapping av Humord og Realfagstermer mot norsk webDewey. Oversettelsen av norsk webDewey vil ferdigstilles i 2015. Humord, som opprinnelig var en tesaurus for humaniora, har siden 2011 blitt utvidet slik at også samfunnsvitenskap med tilgrensende fagområder er blitt inkorporert. Humord består i dag av om lag 26 500 termer (ca 18 000 hovedtermer, 8500 se-henvisninger). Innenfor rammen av dette prosjektet skal Realfagstermer inkludert fellesvokabularet fra prosjektet Realfagstermer og TEKORD (2012) også mappes til norsk webDewey. (Dette utgjør ca. 15 000 hovedtermer, ca 2000 se-henvisninger).

En mapping mellom Humord, Realfagstermer og norsk webDewey vil medføre en kombinasjon av maskinelt og intellektuelt arbeid. I prosjektet Felles terminologi for klassifikasjon med Dewey (2014) ble det utviklet et mappingverktøy. Inneværende prosjekt (På vei mot en generell, norsk tesaurus) bygger videre på dette arbeidet, men vil innlemme flere kilder samt basere valg på termlikhet, statistisk mapping og potensielt også regelstyring som gjør de intellektuelle beslutningene lettere.

Prosjekt- og utviklingsmidler for 2015

Verktøyet vil utvikles i en SKOS/RDF-plattform.

Siden mapping skal gjøres mot den kommende norske webDewey, er det nødvendig og også sterkt ønskelig å jobbe tett med Nasjonalbiblioteket i hele prosjektperioden.

Mappingen vil utføres i tråd med ISO 25964-2 Information and documentation : thesauri and interoperability with other vocabularies : part 2 : interoperability with other vocabularies. Siden ISO-standarden beskriver mappingprosesser generelt, er det behov for tolkninger i et praktisk arbeid. I den europeiske Dewey-brukergruppen EDUG finnes en interessegruppe for mapping. Vi ønsker å bidra til og arrangere en workshop i tilknytning til EDUGs årlige møte hvor man diskuterer og tar stilling til konkrete problemstillinger i praktisk mappingarbeid mellom en tesaurus og Deweys desimalklassifiseringssystem. Vi vil, som et forarbeid til dette møtet, utarbeide ulike «mock-ups» (illustrasjoner) for å synliggjøre hvordan valg ved mapping kan påvirke funksjonalitet i sluttbrukerverktøy. Det er sluttbrukerperspektivet som må være styrende for selve utførelsen av mappingarbeidet.

Dersom Nasjonalbiblioteket og Universitetsbiblioteket i Oslo går videre i arbeidet med å utvikle Humord-tesaurusen til å bli en generell tesaurus, vil denne måtte utvides. Mappingen er en selvstendig aktivitet som kan skje uavhengig av dette. Dersom mappingen skjer parallelt med videreutviklingen av hierarkier i Humord, vil arbeidet være til gjensidig berikelse.

Både kildevokabular (Humord/Realfagstermer) og målvokabular (Deweys desimalklassifiseringssystem) vil imidlertid over tid være gjenstand for endringer. Mapping kan ikke bare utføres én gang, men er en vedvarende prosess. UBO vil vedlikeholde mappingen etter at prosjektperioden er over.

Det vil i dette prosjektet arbeides med å øke IKT-kompetansen blant bibliotekarene som blir involvert. En tett integrering mellom utviklere/datapersonale og emneordseksperter er derfor nødvendig gjennom hele perioden.

Prosjekt- og utviklingsmidler for 2015

Start- og sluttdato for prosjektet

Prosjektvarighet	Fra dato	Til dato
Toårig	01-01-2015	31-12-2016

Aktivitetsplan for prosjektet

Beskrivelse	Fra dato	Til dato
Utarbeide mock-ups av emnesøk i sluttbrukerverktøy	01-01-2015	30-06-2015
Arrangere workshop i samarbeid med NB og EDUG	01-03-2015	30-04-2015
Innarbeide anbefalinger og beslutninger fra workshop	01-04-2015	30-06-2015
Utføre mapping. Testmapping vil foregå i forkant av dette	01-07-2015	31-12-2016

Andre opplysninger

Samarbeidspartnere

Det er ikke registrert noen samarbeidspartnere.

Prosjekt- og utviklingsmidler for 2015

Budsjett**Søknadsbeløp fra Nasjonalbiblioteket 2 086 000****Utgifter**

Beskrivelse	Denne periode	Totalt for prosjektet
Lønn (inkl. overhead)	3 309 000	6 434 000
Deltakelse på seminarer/møter	110 000	220 000
Totale utgifter	3 419 000	6 654 000

Inntekter

Beskrivelse	Denne periode	Totalt for prosjektet
Tilskudd fra Nasjonalbiblioteket	2 086 000	4 036 000
Egenfinansiering	1 333 000	2 618 000
Totale inntekter	3 419 000	6 654 000

Vedleggsoversikt**Beskrivelse**
