

## Statusrapport

### Prosjekt- og utviklingsmidler for 2016

---

#### Mapping mot norsk WebDewey (Ref #684d0eaa)

Søknadssum: **1 820 000**  
Varighet: **Toårig**  
Periode: **Søknaden gjelder andre prosjektår**  
Saksnummer: **2014/359**

#### 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](mailto: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

## Statusrapport

### Prosjekt- og utviklingsmidler for 2016

---

utenlandske studenter og ansatte. Flerspråklige emneinnganger er dermed svært 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 søkeinnganger 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 søkeinnganger, for eksempel Norsk språkbank eller Språkrådet. Vi ønsker også å dele våre erfaringer i form av konferansebidrag og faglige artikler.

## Statusrapport

### Prosjekt- og utviklingsmidler for 2016

---

#### Start- og sluttdato for prosjektet

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

#### Statusrapport

I oktober 2014 sendte UBO søknad om midler til prosjektet «Mapping mot norsk WebDewey», med formål å mappe vokabularene Humord og Realfagstermer mot Norsk WebDewey. Mappingen skal utføres som en datastøttet intellektuell aktivitet, med bruk av et mappingverktøy (kalt «ccmapper») utviklet i forbindelse med prosjektet. Vi mottok 2 086 000 i støtte til prosjektet i brev fra Nasjonalbiblioteket av 27. januar 2015.

Bakgrunnen for innneværende prosjekt finnes i prosjektet «På vei mot en generell norsk tesaurus», som UBO mottok midler til i 2014. En av hovedkomponentene der var å utvikle en metodikk for mapping av tesaurusen Humord mot WebDewey. Prosjektet sendte rapport til NB fra disse aktivitetene 1. mars 2015, se <http://www.ub.uio.no/for-ansatte/om-ubo/prosjekter/tesaurus-mapping/delte-dokumenter/prosjektrapport-tesaurus-mapping.pdf>

I søknaden til «Mapping mot norsk WebDewey» ble følgende framdriftsplan skissert:

# Utarbeide «mock-ups» som illustrerer hvordan mapping mellom en tesaurus og WebDewey kan forbedre emnesøk i sluttbrukerverktøy (januar- juni 2015)

# Arrangere en workshop i samarbeid med Nasjonalbiblioteket og mappinggruppen i EDUG (European DDC User Group) (mars/april 2015)

# Innarbeide anbefalinger og beslutninger fra EDUG-workshop, samt utredningsarbeid ifm. planlegging av mapping og utvikling av mappingverktøyet (april-juni 2015)

# Utføre mapping (juli 2015-desember 2016). Testmapping vil foregå i forkant av dette

Status er som følger:

Prosjektet avholdt workshop i samarbeid med Nasjonalbiblioteket (ved Elise Conradi) og EDUG (European DDC User Group) i tilknytning til EDUGs årlige møte. Workshopen fant sted 15.-16. april i Napoli. På workshopen deltok et 40-talls personer, vesentlig Dewey-eksperter fra ulike nasjonalbibliotek og andre sentrale Dewey-aktører. Dewey-redaktørene (basert ved OCLC) var også til stede. Siden mapping baserer seg på ISO-25964, var også leder av arbeidsgruppen for standarden, Stella Dextre Clarke, til stede og bidro til diskusjonen. UBO ved Grete Seland presenterte våre erfaringer og teoretiske og praktiske problemstillinger knyttet til mapping av emnesystemer mot Dewey. Are Gulbrandsen presenterte mappingverktøyet ccmapper. Programmet bruker algoritmer, kilder, statistisk mapping og termlikhet for å generere mappingkandidater. Presentasjonene finnes her: <http://edug.pansoft.de/tiki-index.php?page=2015+meeting>.

Dag 2 av workshopen var satt av til gruppearbeid. Basert på dette har det i ettertid blitt utviklet et sett med anbefalinger om mapping av emnevokabular mot Deweys desimalklassifikasjonssystem. Både UBO (ved Grete Seland og Unni Knutsen) og NB (ved Elise Conradi) har sittet i redaksjonskomiteen. Det er første gang det publiseres internasjonale anbefalinger på dette området. Anbefalingene finnes her: <http://edug.pansoft.de/tiki-index.php?page=EDUG+workshops>

## Statusrapport

### Prosjekt- og utviklingsmidler for 2016

---

Det tyske selskapet Pansoft (som blant annet har utviklet oversetterverktøy og dataløsningen for WebDewey) tok i løpet av EDUGs årlige møte kontakt med prosjektet vårt for å undersøke muligheten for å samarbeide om dataløsningen. UBO hadde opprinnelig planlagt – og fått midler til – å utvikle en prototype som kunne brukes i mappingarbeidet ved UBO, et «proof-of-concept». Pansofts forslag gjør det innenfor rekkevidde å få i produksjon et fullverdig mappingverktøy – en web-applikasjon som også vil være et grunnlag for mapping av andre vokabularer, både i Norge og i resten av verden. I etterkant av EDUG-workshopen i april har prosjektet vårt avholdt flere møter og videokonferanser med Pansoft, med gjennomgang av metodikk samt teknologisk tilnærming for et mappingverktøy.

Under årets WLIC (den årlige IFLA-konferansen) møttes UBO ved Unni Knutsen og NB ved Jonny Edvardsen, Svein Arne Tinnesand og Elise Conradi for å diskutere mulighetene for at UBO/NB og Pansoft samarbeider om utviklingen og lager et fullt fungerende mappingverktøy som gjenbraker UBO-utviklet metodikk, interaksjonsdesign og arkitektur. På møtet var det enighet om at det er fornuftig å utvikle en web-applikasjon i full produksjon framfor en prototype. Det ble besluttet at UBO kontakter Pansoft med våre kravspesifikasjoner, og ber dem lage et reelt kostnadsoverslag. Når forhandlinger om pris er ferdige sett fra UBOS side, involveres NB. Før en eventuell inngåelse av kontrakt, må det avklares om det er UBO eller NB som er kontraktspartnerne. NBs videre involvering vil avhenge av om de ser på utviklingen av verktøyet som et nasjonalt anliggende. UBO arbeider i skrivende stund aktivt med å utvikle skisser til dataløsningen som ivaretar våre behov for arbeidsflyt og tilgang til aktuelle kilder. Basert på disse og videre samtaler med Pansoft, vil det kunne utarbeides et kontraktsforslag som kan presenteres for NB.

I påvente av at mappinganbefalingene skulle bli ferdigstilte – og fordi nye problemstillinger har fortsatt å dukke opp – har vi fram til 1. oktober ansett mappingaktiviteten vår som testmapping. Totalt har vi foretatt testmapping av ca. 340 begreper. Hvert begrep får i gjennomsnitt to mappinger mot WebDewey, siden ett emne kan ha flere faglige plasseringer i WebDewey. Disse blir nå gjennomgått i lys av EDUGS publiserte mappinganbefalinger, og får status som ferdigmappede. I påvente av et ferdigutviklet mappingverktøy, har vi gjort forbedringer i et program (kalt «µmapper») som er utviklet av Realfagsbiblioteket. Dette har vært brukt i tidligere mappingprosjekter. Verktøyet µmapper er basert på termsammenfall, og vi bruker dette til å mappe overlappsvokabularene mellom våre kildevokabularer og Norsk WebDewey. Dette utgjør ca. 30 % av hvert vokabular. Foreløpig er ca. 200 begreper mappet, hvert med drøyt to mappinger mot WebDewey.

I tillegg til presentasjonene i forbindelse med workshopen i Napoli, gjorde Elise Conradi og Unni Knutsen en fellespresentasjon under årets kunnskapsorganisasjonsdager: «Dewey fra menneske til maskin» hvor også mappingprosjektet var i fokus.

Foreløpig består gruppen som gjør mapping av UBOS egne prosjektmedarbeidere. Øvrige tilsetninger i prosjektet vil skje når det nye mappingverktøyet er ferdig/nær ferdigstillelse

## Statusrapport

### Prosjekt- og utviklingsmidler for 2016

---

#### Budsjett

Søknadsbeløp fra Nasjonalbiblioteket 1 820 000

#### Utgifter

Beskrivelse	1. prosjektår	2. prosjektår	3. prosjektår	Totalt
Lønn (inkl. overhead)	3 309 000	<b>1 400 000</b>	0	4 709 000
Deltakelse på seminarer/møter	110 000	<b>100 000</b>	0	210 000
Utvikle ccmapper som verktøy i WebDewey. Kontraktingåelse med Pansoft, jfr medfølgende rapport.	0	<b>1 000 000</b>	0	1 000 000
<b>Totale utgifter</b>	<b>3 419 000</b>	<b>2 500 000</b>	<b>0</b>	<b>5 919 000</b>

#### Inntekter

Beskrivelse	1. prosjektår	2. prosjektår	3. prosjektår	Totalt
Tilskudd fra Nasjonalbiblioteket	2 086 000	<b>1 820 000</b>	0	3 906 000
Egenfinansiering ( <b>bekreftet tilskudd</b> )	1 333 000	<b>680 000</b>	0	2 013 000
<b>Totale inntekter</b>	<b>3 419 000</b>	<b>2 500 000</b>	<b>0</b>	<b>5 919 000</b>

#### Vedleggsoversikt

Beskrivelse

---