

Mapping mot Dewey

Kristine Aalrust Kristoffersen
k.a.kristoffersen@ub.uio.no



Mapping mot norsk WebDewey



Agenda

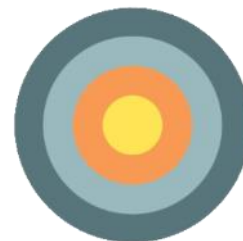
- Litt om prosjektet
- Hva er mapping?
- Relasjonstyper
- Kontekst
- Grunnleggende forskjeller
- Mapping mot Dewey



Hva er mapping?

«Mapping er en aktivitet hvor det etableres relasjoner mellom begreper i to ulike kontrollerte vokabularer. Man mapper fra et **kildevokabular** og finner korresponderende begreper i et **målvokabular**. Mapping brukes også til å benevne resultatet av en mappingaktivitet, dvs. relasjonen mellom et begrep i ett vokabular og en eller flere begreper i et annet vokabular.»

Fra: Styringsdokument for prosjektet «Mapping mot norsk WebDewey» ved Universitetsbiblioteket i Oslo, s.3



Hvorfor mappe?

- Felles verbale innganger
 - Søk på tvers
 - Bedre indekseringskonsistens
- Interoperabilitet
- Internasjonal vinn-vinn



Om prosjektet

- UiO HumSam-biblioteket og Realfagsbiblioteket
- Humord og Realfagstermer mot norsk WebDewey
- Finansiert av Nasjonalbiblioteket



Mål for prosjektet

- Hovedmål: mappe vokabularene mot norsk WebDewey
- Utvikle programvare
 - CCmapper lanseres snart i samarbeid med Pansoft
 - I dag brukes [umapper](#)
- Utvikle metodikk for mapping mot Dewey
 - EDUG recommendations

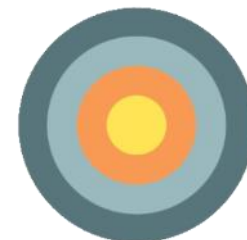
Mapping til Dewey

ISO 25964-2:2013

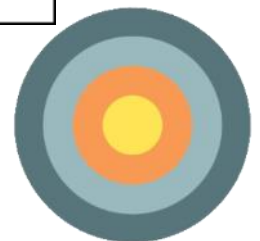
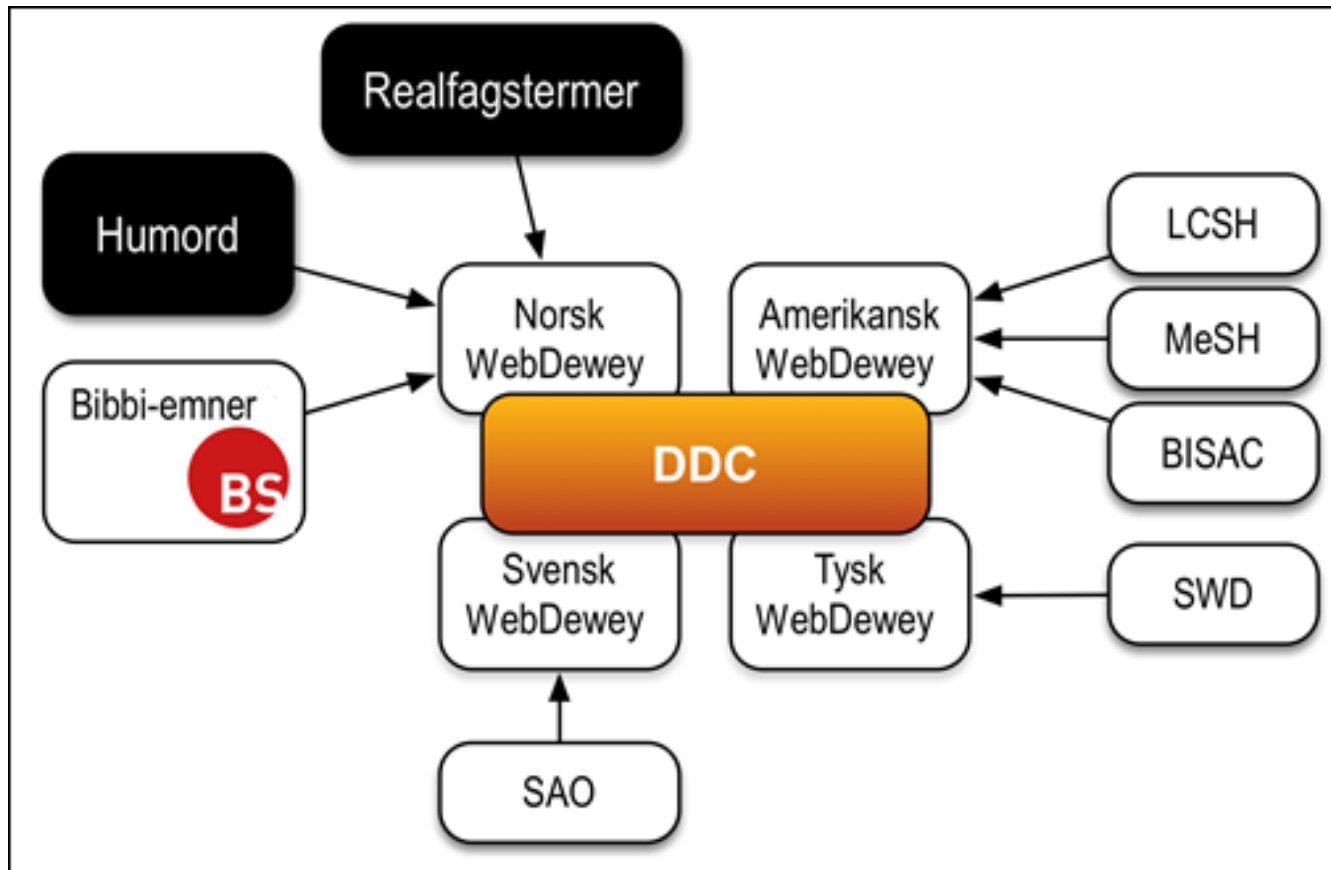
- Interoperabilitet fra tesaurus til andre vokabularer
- Publisert 2013 – mye mapping gjort før den (mer om dette senere)
- Mapping fra Humord og Realfag til norsk WebDewey er blant de første som bruker standarden
- Noe om mapping med prekoordinerte vokabularer, men egentlig ikke noe om klassifikasjonsskjema

EDUG recommendations

- European DDC Users Group
- Anbefalinger for mapping med DDC
- Anbefaler å følge ISO-standard
- Anbefaler nav-modellen



Mot Dewey: Nav-modellen



Enkelt eksempel

Semantisk web

Semantisk web i WebDewey

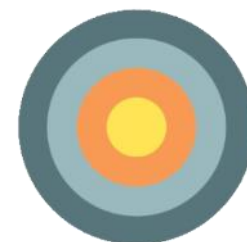


Relasjoner



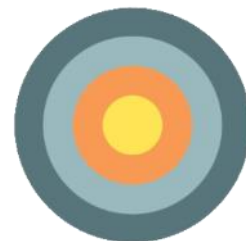
Relasjoner: ISO-standard

=EQ	(EQ = Equivalence). Likhetstegnet angir at mappingen er eksakt
~EQ	Tilden angir at mappingen ikke er eksakt. Det innebærer at begrepene kan være like i noen sammenhenger, men ikke i alle eller at begrepene kan være delvis overlappende eller avvike noe i betydningsinnhold
BM	Broader mapping. Termen i målvokabularet har en videre betydning enn termen i kildevokabularet
NM	Narrower mapping. Termen i målvokabularer har en mer spesifikk betydning enn termen i kildevokabularet
RM	Related mapping. Termen i målvokabularet assosieres med termen i kildevokabularet, men er ikke et synonym, et kvasisynonym eller en bredere eller smalere term



Eksempel på bruk av relasjoner

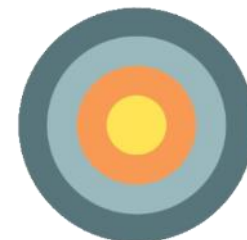
- Eksempel: Internett



Flere typer relasjoner

- CrissCross (Tyskland)
 - Mappet SWD til Dewey
 - Degrees of determinacy
 - D4 til D1
 - D4: Identisk
 - D3: Nesten identisk
 - D2: Likt en del av begrepet i målvokabularet
 - D1: Likhet med en liten del av målbegrepet

Gödert, Hubrich & Nagelschmidt, 2014, s. 250



Flere typer relasjoner

- Chaplan (1995): 19 kategorier av relasjoner
- FinnOnto: fra tesaurus til ontologi
 - mer semantikk i relasjonene



Hva betyr en relasjon?

Kometer

Halley's comet

Verdensrommet

Kometer

- ISO-relasjonene tilpasset tesaurus
- Gjennomgående i mappingprosjekter:
vanskelig og krevende å velge relasjon



Mening



Hva betyr begrepet?

- Emneord i kontekst
 - Se- og se også-henvisninger
 - Litteraturbelegg
 - Plassering
 - Individuell vurdering (?)



Core meaning og conceptual meaning

Core meaning

- Mening uavhengig av litteratur tilknyttet begrepet
- Core meaning → Focused mapping (én mapping pr. begrep)

Conceptual meaning

- Core meaning + man tar hensyn til litteraturbelegget
- Conceptual meaning → comprehensive mapping

Eksempel: Ljåer

Gödert, Hubrich & Nagelschmidt, s. 108-110



Mening i plassering

- Dewey-hierarki

629.134351 Brensel	
600	Teknologi
620	Teknikk (ingeniørfag)
629	Andre områder innen teknikk
629.1	Luft- og romfartsteknikk
629.13	Aeronautikk
629.134	Luffartøykomponenter og generelle teknikker
629.1343	Deler
629.13435	Motorer og brensel
629.134351	Brensel

- Spesielt viktig v/homonymi, nærsynonymi, etc.

- Humord-hierarki

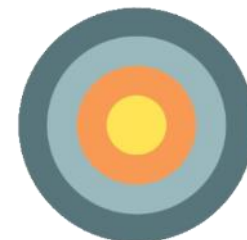
HUME: «Rokker»

Broader

04440 Redskaper
 04429 Utstyr
 04305 Ingeniørfag
 04660 Artefakter
 04657 Etnologi
 04654 Tradisjonsforskning
 04452 Kulturkunnskap

Other relationships

↗ rejected to WDNO: 641.692 «Fisk»
 ↗ rejected from REAL: «Rokker»
 + New mapping

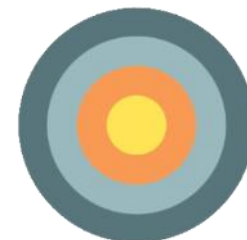


Grunnleggende prinsipper



Forskjell i formål

- Kan man egentlig si at et emneord og et klassenummer er «det samme»?
- Et emneordssystem er samlingsbasert
- Norsk WebDewey oversatt fra amerikansk WebDewey
- Altså: emneordssystem oppstår, utvikles og vedlikeholdes i takt med samlingen, og med de som jobber for å vedlikeholde dem
- WebDewey er et generelt system med plass til «alt»
 - (men er også preget av universet det eksisterer i)
- BS emneord
 - Også tatt inn i WebDewey
 - Preget av norske folkebiblioteks «felles samling»?



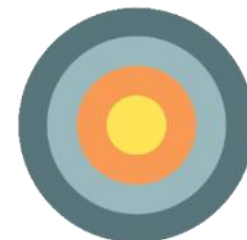
Sortering på fag og emne

- Tradisjonelt: klassifikasjonsskjema sorteres etter fagområder, tesaurus sorteres etter emner
 - Ett begrep mappes flere steder
- Eksempel: ananas



Sammensatte begreper

- Dewey åpner for nummerbygging
- Prekoordinerte emneord, eksempel fra BS emneord:
Landssvik - Norge - Kunstnere - 1940-1945 : historie
- Igjen: hva betyr relasjonene? Hva står det «mellom linjene?»
- Mapping av strenger
- Postkoordinerte vokabularer kan også ha sammensatte begreper
- Tverrspråklig mapping – forskjell i regler for orddeling



Ett begrep, mange mappings

- Eksempel: Landbruk



Menneskelig faktor

- F.eks.: Humord har topptermer etter fag
- Intern praksis
 - F.eks. termer som bare skal brukes i kombinasjon med andre termer
- «Brukt for» brukes for synonymer og underordnede begreper



Øvelse: Fellesmapping

- Diamanter



Takk for meg!

- All info om prosjektet og daglig statistikk på prosjektsidene:
<http://www.uio.no/for-ansatte/enhetssider/ub/prosjekter/mapping-mot-webdewey/>
- Alle mapper åpent tilgjengelig:
<https://lambda.biblionaut.net/relationships>
- Ta kontakt!

Litteratur

Chaplan, M. A. (1995). Mapping Laborline thesaurus terms to Library of Congress subject headings: implications for vocabulary switching. *Library Quarterly* (65). doi: 10.1086/602752

Gödert, W., Hubrich, J, & Nagelschmidt, M. (2014). *Semantic knowledge representation for information retrieval*. Berlin: Walter De Gruyter

Hyvönen, E., Viljanen, K., Tuominen, J. Kauppinen, T., Ruotsalo, T., Valkeapää, O., ...Kurki, J. (2007). *Elements of a National Semantic Web Infrastructure: Case Study Finland on the Semantic Web*. Helsinki: Helsinki University of Technology. Hentet fra

<http://www.seco.tkk.fi/publications/2007/hyvonen-et-al-elements-2007.pdf>

Knutsen, U., & Gulbrandsen, A. (2015). På randen av mapping. I G. H. Aalstad (Red.), *Kunnskapsorganisering* (s. 36–46). Oslo: Nasjonalbiblioteket.

Styringsdokument for prosjektet «Mapping mot norsk WebDewey» ved Universitetsbiblioteket i Oslo (UBO) (2015) Hentet fra http://www.uio.no/for-ansatte/enhetssider/ub/prosjekter/mapping-mot-webdewey/delte-dokumenter/styringsdokument_mappingmotwebdewey.pdf

