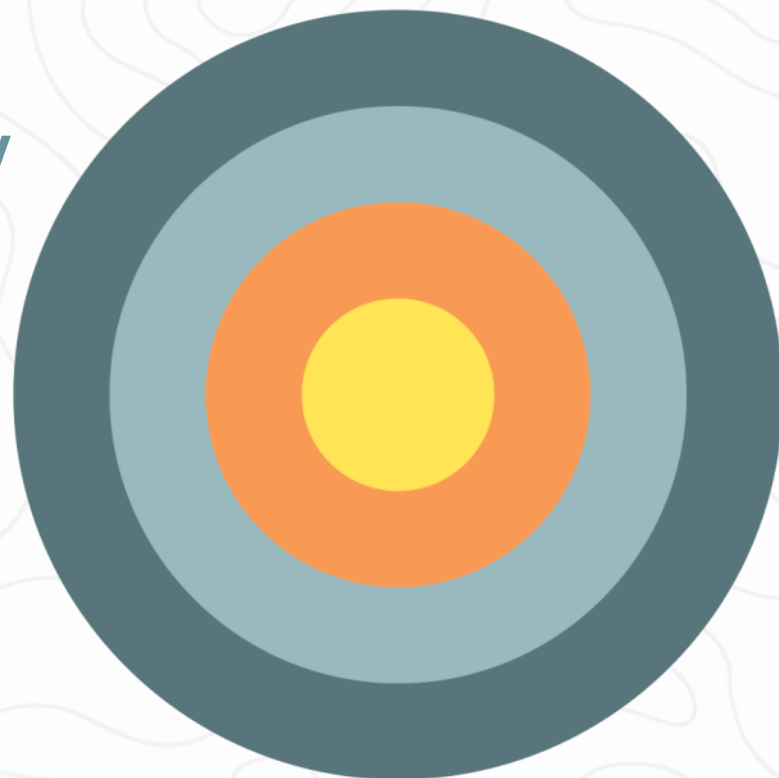


Mapping mot norsk WebDewey

Viola Kuldvere, Heidi Konestabo,
Dan Michael O. Heggø,
Kristine Aalrust Kristoffersen

Mapping mot norsk WebDewey



Hva – Hvorfor – Hvordan

- Hva er mapping?
- Hvorfor mapper vi?
- Hvordan mapper vi?
- Prosjektet er finansiert av Nasjonalbiblioteket
- Les mer i brosjyren, styringsdokumentet og tidligere presentasjoner på prosjektsiden:
ub.uio.no/om/prosjekter



Hva – Hvorfor – Hvordan

Hva er mapping?

Mapping vil si å etablere relasjoner mellom begreper i to ulike vokabularer

Realfags-
termer
og
Humord



Norsk
WebDewey

Mål for prosjektet er å mappe Realfagstermer og Humord mot norsk WebDewey

Relasjonstyper

Relasjonstyper

Vi bruker fire ulike relasjonstyper for å angi graden av samsvar mellom begrepene som mappes

Eksakt ekvivalens

Begrepet tilsvarer hele deweyklassen

Kildevokabularer
Humord og Realfagstermer

Målvokabular
Norsk WebDewey

Transcendentalfilosofi

141.3 Transcendentalisme

Tilnærmet ekvivalens

Begrepet har stort overlapp med deweyklassen

Symmetrigrupper

512.2 Grupper og gruppeteori

Overordnet mapping

Begrepet er snevrere enn deweyklassen

Leksikalsk syntaks

415 Standardspråkets grammatikk

Relatert mapping

Begrepet er assosiert med deweyklassen

Boformer

643.1 Boliger

Et emne kan mappes til flere deweyklasser

Relasjonstyper

Vi bruker fire ulike relasjonstyper for å angi graden av samsvar mellom begrepene som mappes

Kildevokabularer

Humord og Realfagstermer

Målvokabular

Norsk WebDewey

Eksakt ekvivalens

Begrepet tilsvarer hele deweyklassen

Hester

599.6655 «Equus caballus (hester)»

Eksakt ekvivalens

Begrepet tilsvarer hele deweyklassen

Hester

636.1 «Hester - husdyrhold»

Relatert mapping

Begrepet er assosiert med deweyklassen

Hester

357.2 «Remonteskoler»

Relatert mapping

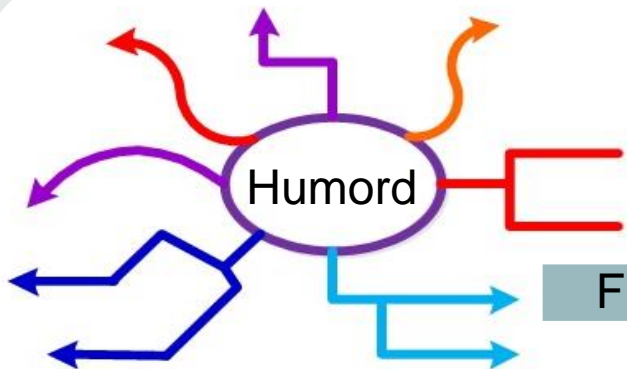
Begrepet er assosiert med deweyklassen

Hester

615.851581 «Hester - terapeutisk bruk»

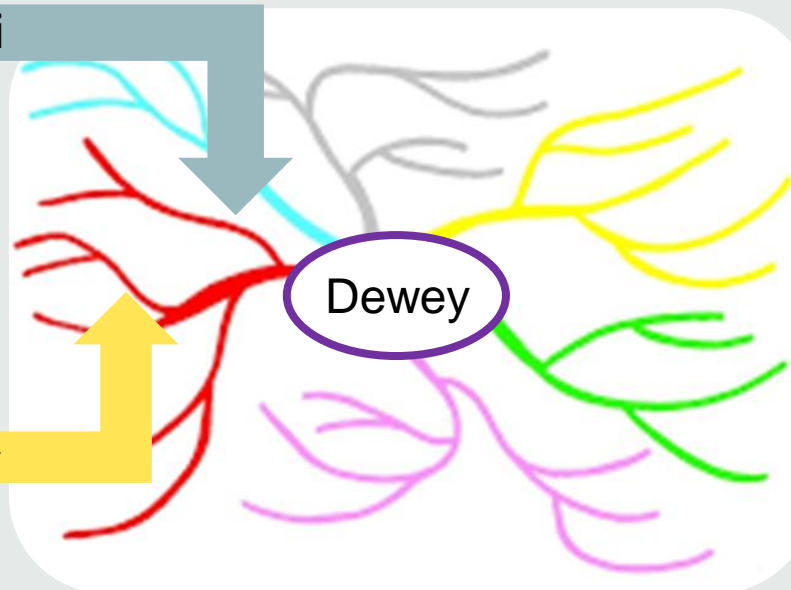
Hva – Hvorfor – Hvordan

Hva er mapping?



Å mappe er som å lage koblinger mellom ulike kart over kunnskapsuniverset

Fraseologi



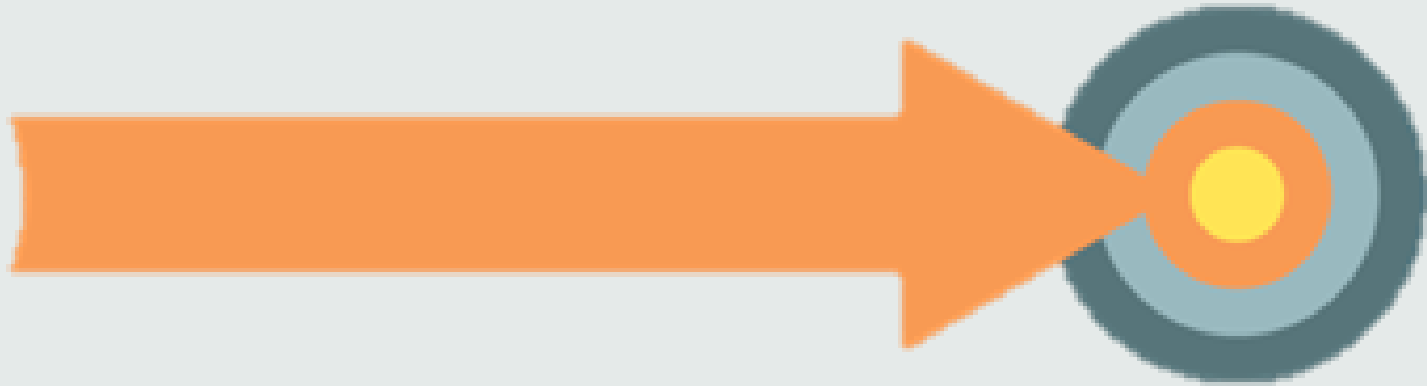
Realfagstermer



Solformørkelser

Hva – Hvorfor – Hvordan

Hva er mapping?



Interoperabilitet mellom vokabularer

Hvorfor er dette nyttig?

Hva – Hvorfor – Hvordan

Hvorfor mapper vi?

UBs målsetting:

Bibliotekets hovedoppgave er å støtte universitetets formål innen forskning, utdanning, formidling og innovasjon

Hva – Hvorfor – Hvordan



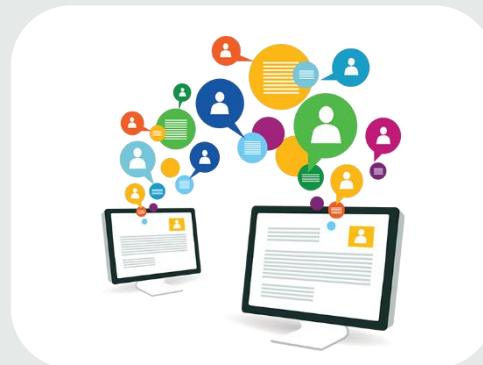
Tilgjengelighet



Vekst



Globalisering



Teknologi

Hva – Hvorfor – Hvordan



Tilgjengelighet

«Sørge for gode gjenfinningsmuligheter i en stadig voksende informasjonsmengde».

Sluttbrukerne:

- søk på tvers av samlinger og språk
- importerte poster blir søkbare med våre egne emneord

Bibliotekpersonalet:

- støtte i indekseringsarbeidet

Hva – Hvorfor – Hvordan

Søk for sluttbrukere: WebDeweySearch

WebDewey Search

Søk etter i DDK23-notasjoner Registeret Humord Realfagstermer

Deweys hovedklasser

Deweynummer	Klassebetegnelse	Treff i denne klassen og dens underordnede klasser
	Deweys hovedklasser	
000	Informatikk, informasjon og generelle verker	172663
100	Filosofi og psykologi	141833
200	Religion	186077
300	Samfunnsvitenskap	911315
400	Språk	146615
500	Naturvitenskap	209986
600	Teknologi	471858
700	Kunst og fritid	732962
800	Litteratur	615370
900	Historie og geografi	586935

Hva – Hvorfor – Hvordan

Støtte i indekseringsarbeidet for bibliotekpersonale

523.78 Formørkelser

- [500](#) Naturvitenskap
- [520](#) Astronomi
- [521-525](#) Astronomi
- [523](#) Bestemte himmellegemer og fenomener
- [523.3-523.7](#) Bestemte deler av solsystemet
- [523.7](#) Sola
- 523.78 Formørkelser**

Historikk

Kommentarer

Noter

Registertermer

[Formørkelser--sola](#)

Realfagstermer

[Solformørkelser](#) [bokmål]

[Solar eclipses](#) [engelsk]

[Solformørkingar](#) [nynorsk]

=EQ

Hva – Hvorfor – Hvordan

«Gi brukerne utvidet tilgang til fysiske og digitale samlinger».



Vekst

Vekst i egen katalog: Importerte poster uten våre metadata, men med dewey-klassifikasjon

Våre mapper publiseres som lenkede data til bruk på den semantiske veven: Potensiale for utnyttelse utenom UB og utenfor biblioteksammenheng

Hva – Hvorfor – Hvordan

UBs kjerneverdier:
Kvalitet, åpenhet og tilgjengelighet



Kvalitetsheving av egne vokabularer (Humord, Realfagstermer og UBs emneregister) ved rettinger og forbedringer

Kvalitetsheving av WebDewey ved rettinger og nummerbygging (1/4 av all nummerbygging)

Kompetanseheving av personalet

Hva – Hvorfor – Hvordan

Prosjektet bringer oss i kontakt med nasjonalt og internasjonalt kompetansemiljø innenfor metadata

Nybrottsarbeid: anvendelse av ISO 25964-2

Utarbeiding av EDUG recommendations for mapping

Nasjonalbiblioteket/WebDewey-redaksjonen, Biblioteksentralen, EDUG, OCLC, Pansoft



Globalisering

«Samarbeide på tvers av institusjons- og landegrenser for å bli best mulig».

«Utvikle strategiske partnerskap gjennom internasjonalt og nasjonalt samarbeid».

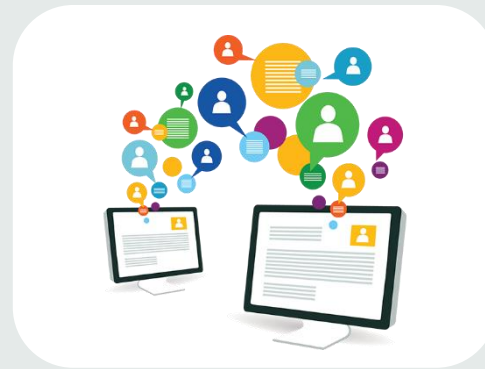
Hva – Hvorfor – Hvordan

Utvikling av mappingverktøy på samme plattform som WebDewey: ccmapper (Pansoft)

OCLC ønsker å ta i bruk ccmapper

Relevans for UB-prosjektet «Visuell navigasjon» – forbedre søk ved utnyttelse av norske emnevokabularer

«Levere kunnskapskilder og tjenester med effektive brukergrensesnitt på stadig flere plattformer»



Teknologi

Hva – Hvorfor – Hvordan

Hvordan jobber vi?



Utarbeiding av retningslinjer



Utføring av mappinger

CC  MAPPER

Utvikling av mappingverktøy

Hva – Hvorfor – Hvordan

Med et mappingverktøy tenker vi noe (ikke noen) som kan

- 1) foreslå mappingkandidater
- 2) tilby et effektivt grensesnitt som støtter opp om en arbeidsflyt for å jobbe seg gjennom foreslåtte mappingkandidater og legge til nye.
- 3) tilby lagring og gjenfinning av mappinger og metadata
- 4) fange opp endringer i kilde- eller målvokabular som kan påvirke eksisterende mappinger og varsle om dette.

Hvordan jobbe? Vi er i mars 2014

	A	B	C	D	E	F	G	H
1	ID	TERM	DDK	RELASJON	SIGN	DDK-treff	TORT-treff	KOMMENTAR
12	REAL000420	Daler	551.442	EQ	ml	1	0	
13	REAL000335	Knopper	575.486	EQ	ml	1	0	
14	REAL000717	Perturbasjoner	521.4	EQ	vk	1	0	
15	REAL013869	Mørk materie	523.1126	EQ	vk	1	0	
16	REAL000616	Vakuu	533.5	EQ	vk	1	0	
17	REAL013862	Nøytroner	539.7213	EQ	vk	1	1	
18	REAL000252	Bioluminescens	572.4358	EQ	ml	1	1	(vil også ha NM 572.435)
19	REAL000876	Analytisk topologi	514.7	EQ	vk	1	0	
20	REAL031037	Uranus	523.47	EQ	ml	1	1	
21	REAL000806	Fjelløkologi	577.53	EQ	vk	1	1	Også: NM: Alpin økologi 577.5/38
22	REAL013930	Skogøkologi	577.3	EQ	vk	1	2	En del strenger/bygde nummer i wd, NM 577.327-skogs
23	REAL000812	Protoner	539.72123	EQ	vk	1	1	også: Klassifiser protoner betraktet som kosmisk stråling
24	REAL000778	Erosjon	551.372	EQ	vk	1	1	Geo
25	REAL000593	Lineære operatører	515.7246	EQ	vk	1	1	
26	REAL013407	Biologi	570	EQ	vk	1	0	
27	REAL000862	Flislegging	516.132	EQ	vk	1	0	(geometri/matematikk)
28	REAL000197	Metamorfe bergarter	552.4	EQ	ml	1	1	(av inkluderer noten følger også NM til hhv Gneis, marm
29	REAL000530	Atomkjerner	539.7232	EQ	ml	1	0	Men også: ~EQ 539.723. Og hva med de i noten: Klassi
30	REAL013627	Solstråling	551.5271	EQ	vk	1	1	
31	REAL000000		510.00	EQ	ml	1	1	

+ ☰ 2 Ark1 ▾ TORT ▾ Liste A ▾ Liste B ▾ 1 Liste B.2 ▾ Liste C ▾ Liste D ▾

Skjerm bilde fra arbeid i Google Docs i avslutningen av TORT-prosjektet

April–mai 2014: μ mapper blir født

```
5d94712 - Show mapping abbreviations used in ISO 25964 (2 years, 8 months ago) <Dan Michael O. Heggø>
ef0b292 - Icons for each activity (2 years, 8 months ago) <Dan Michael O. Heggø>
a6db5a9 - Increase default number of relationships to show per page from 10 to 20
753e16f - New relationship form: should not be possible to add relationships with the same name
afc26b4 - Renamed 'tag' → 'list' (2 years, 8 months ago) <Dan Michael O. Heggø>
d521227 - Adding tags views that weren't added initially (2 years, 8 months ago) <Dan Michael O. Heggø>
fc9b67a - Fix: reviewed_by has been removed (2 years, 8 months ago) <Dan Michael O. Heggø>
97ded4a - Relationships form: clarify and translate (2 years, 8 months ago) <Dan Michael O. Heggø>
c53374d - More flexible filter by review state (2 years, 8 months ago) <Dan Michael O. Heggø>
57fd76f - DB: remove revision reviewed_by (covered by created_by) (2 years, 8 months ago) <Dan Michael O. Heggø>
21a36ac - Removing contributing.md from framework (2 years, 8 months ago) <Dan Michael O. Heggø>
ace407c - Add .bowerrc (2 years, 8 months ago) <Dan Michael O. Heggø>
6b07cee - (tag: v0.1.0) Update composer.json (2 years, 8 months ago) <Dan Michael O. Heggø>
a84f820 - Add README (2 years, 8 months ago) <Dan Michael O. Heggø>
5ec2ee6 - First usable version (2 years, 8 months ago) <Dan Michael O. Heggø>
cbfe45a - Initial commit (2 years, 10 months ago) <Dan Michael O. Heggø>
```

(END)

Notater

Kommentarer

Fce2b20

µmapper

Relationship #8683

Nåværende arbeidsliste: [Hopp til neste](#) [rt-ddk-katkor](#) Vekting: 1.05

Source concept:



→ Mapping relation: →

Target concept:

RT: «Solceller»

Solar cells (*en*)
Fotovoltaiske celler
Solceller (*nn*)

Other relationships








-  [suggested](#) to DDK23:
621.31244 «Produksjon av
elektrisitet fra solstråler»
-  [exact equivalence \(=EQ\)](#) to
TEK: 621.383 «Solceller»
- [+ New mapping](#)

broader (BM)

Godkjent av Heidi

[Godkjenn](#)
DDK23: 621.381542
«Fotoelektriske og
fotoelektroniske komponenter»

Fotoceller
Fotoelektriske celler
Fotoelektriske enheter
Fotoelektriske komponenter
Fotoelektroniske enheter
Fotoelektroniske komponenter
Fotoemisjonsceller
Fotokonduktive celler
Fotomultiplikatorer

Aktivitet [Heidi](#) kommenterte for 2 years ago Inkluderer-notene er ikke likestilte, mens her-notene er det. Derfor ikke EQ. [Heidi](#) godkjente relasjonen for 2 years ago [Dan Michael](#) kommenterte for 2 years ago eller EQ?? Solceller (photovoltaic cells) ligger under "Inkluderer" i WD [Dan Michael](#) endret status for relasjonen fra **suggested** til **broader (BM)** for 2 years ago [BiblioBot](#) kommenterte for 2 years ago [rt-ddk-katkor] DDK-nummeret er brukt på 1 av 1 dokumenter med dette RT-emnet.

Fra μ mapper til CCMapper

H14/V15:

- Are kommer inn. Plan: Bygge videre på ideer fra μ mapper, men (1) bytte ut all teknologi til å basere seg fullstendig på RDF og lenka data, og (2) designe brukergrensesnitt på nytt fra scratch i samråd med brukerne.
 - Faresignal: altfor mye uprøvd og ukjent teknologi på en gang, samtidig som det er uklart hva den løser. Ingen reell vurdering av alternativer.
- Samtidig i prosjektet: Innser at det er fryktelig mye vi ikke visste om Dewey (og forsåvidt om hierarkiet i Humord). Mappingretningslinjene blir til.

Fra μ mapper til CCMapper

Mellomspill sommeren 2015:

- April 2015: møte med Pansoft (som utvikler WebDewey). Avtaler å bruke Pansoft som konsulenter for å generere mappingkandidater med Solr, siden de har mye erfaring med teknologien.
 - ... men ikke til denne bruken.
 - juni–august 2015: Kommunikasjonen går altfor tregt frem og tilbake og det blir ingen dynamikk som fører til nye ideer. Noen mener resultatene virker lovende, andre ikke.

Fra μ mapper til CCMapper

Høst 2015:

- Fremdeles ingen realistisk plan for hvordan CCMapper-prototypen vår skal komme opp.
- Pansoft får i oppdrag å utvikle og drifte CCMapper.
 - (De ønsker ikke å jobbe med RDF og lenka data, men MARC21 og den samme databaseteknologien brukt i μ mapper)
 - Kravspekk må skrives
 - ... men ikke så fryktelig detaljert, for vi blir enige om å bruke smidig utvikling.
- Setter opp og oppgraderer μ mapper mens vi venter på CCMapper

Fra μ mapper til CCMapper

- Smidig utvikling (agile) baserer seg på tre (fornuftige) antakelser:
 - Du vil aldri klare å samle alle krav til en løsning før du begynner.
 - Kravene vil garantert komme til å endre seg underveis.
 - Det vil alltid være mer å gjøre enn man har tid til.
- Hvorfra det følger at:
 - Man må utvikle noe som fungerer så raskt som mulig, sånn at man kan prøve det, vise det til andre, og gjennom det oppdage veien videre.
 - Man må utvikle systemer *på en måte som gjør at de kan endres.*
 - osv...

Fra μ mapper til CCMapper

- Januar 2016: Pansoft starter utvikling av CCMapper
- Resten av året: Forsøk på smidig utvikling, men i praksis ikke så veldig smidig. Men fremover går det ihvertfall!
- Desember 2016: CCMapper tas i bruk!

Hva – Hvorfor – Hvordan

Hvordan mapper vi rent konkret?

Datastøttet, intellektuell prosess

Demonstrasjon av verktøyet

The logo for CC BY-NC-MAPPER is displayed in white text on a black rounded rectangular background. It consists of the letters 'CC' on the left, a central icon of a person with a crossed-out figure (representing 'BY-NC'), and the word 'MAPPER' on the right.

CC BY-NC MAPPER

ccmapper screenshot – mapping

CC MAPPER
Språk ▾ Hjelp Grete Seland Logg av

Mapping Kontroll Søk

Humord

Filosofi > Begreper > Filosofiske begreper > Etsiske begreper > Dyder > Kardinaldyder

Lingvistikk > Grammatikk > Semantikk > Begreper > Filosofiske begreper > Etsiske begreper > Dyder > Kardinaldyder

Foretrukken term
Kardinaldyder Neste

Bytt begrep

Tilmappede begreper
(ingen)

Aktiviteter
(ingen)

Status
? Ikke påbegynt

Søk etter valgt term

Emnesøk NO WebDewey EN WebDewey NO WebDeweySearch Skosmos Norart	Bibbi websøk Wikipedia NO Wikipedia EN Wiktionary SNL Bokmålsordboka
---	---

Mapping

Mappingkandidater (10) Søk i WebDewey

Mapping fullført Lagre

Sorteringsrekkefølge Type Relevans Dewey

Mappinger med kommentarer

179.9 Dyder (1.016)

-EQ	-EQ	BM	RM	RJ
-----	-----	----	----	----

241.4 Dyder (1.016)

-EQ	-EQ	BM	RM	RJ
-----	-----	----	----	----

+ 704.9491799 Dyder--kunstmotiver (498)

-EQ	-EQ	BM	RM	RJ
-----	-----	----	----	----

Dewey +

Filosofi og psykologi > Etik > Anvendt etik > Andre etiske normer

179.9 Dyder
Som ikke har annen plassering
Inkluderer: Vennlighet, snillhet, taknemlighet, ærlighet, ydmykhet, beskjedenhet, tålmodighet, forsiktighet, respekt, selvkontroll, selvstendighet, toleranse
Her: Dyd
Se også [171.3](#) for dydsetikk som et etisk system

Registertermer

- Vennlighet
- Snillhet
- Taknemlighet
- Ærlighet
- Ydmykhet
- Beskjedenhet
- Tålmodighet
- Forsiktighet
- Respekt
- Selvkontroll
- Selvstendighet
- Toleranse
- Dyd
- Dyder

Mappet fra (27)

ccmapper screenshot – kontroll

CC MAPPER
Språk ▾ Hjelp Grete Seland Logg av

Mapping Kontroll Søk

Humord

Foretrukken term
Språk Neste

Underordnede begreper
Kunstige språk Bytt begrep
Naturlig språk

Tilmappede begreper
REAL: =EQ Språk ✔ kontroll fullført
Brukt for: Språk

Aktiviteter
 Vibeke Stockinger Lundetræ at 07. mai 2016 13:03
✔ Vibeke Stockinger Lundetræ at 07. mai 2016 13:09

Status
✔ kontroll fullført

Søk etter valgt term

Emnesøk	Bibbi websøk
NO WebDewey	Wikipedia NO
EN WebDewey	Wikipedia EN
NO WebDeweySearch	Wiktionary
Skosmos	SNL
Norart	Bokmålsordboka

Søk

UBO emneregister til Dewey

Forrige side
Neste side

Mapping

Mappingkandidater (58) Søk i WebDewey

Kontroll fullført Lagre

Sorteringsrekkefølge Type Relevans Dewey

Mappinger med kommentarer

400 Språk (2.894.440)

=EQ	~EQ	BM	RM	RJ
-----	-----	----	----	----

Vibeke Stockinger Lundetræ at 07. mai 2016 13:03 : endret fra Open til =EQ

Mapping Generator at 29. nov 2016 02:13 : Mappingforslag =EQ

H1--014 Kommunikasjon (2.590.830)

=EQ	~EQ	BM	RM	RJ
-----	-----	----	----	----

Vibeke Stockinger Lundetræ at 07. mai 2016 13:03 : endret fra Open til ~EQ

Mapping Generator at 29. nov 2016 02:13 : Mappingforslag ~EQ

372.6 *Språkerfardigheter (kommunikasjonsferdigheter) (1.538.300)

=EQ	~EQ	BM	RM	RJ
-----	-----	----	----	----

Dewey

Språk > Språk

400 Språk
Her: Tverrfaglige verker om filologi (språk og litteratur)
Språkpoltikk, se [306.449](#)
Litteratur, se [800](#)
Litterær komposisjon, se [808](#)
Språket innen et bestemt emne, se emnet, tilføyd notasjon [H1--014](#) fra Hjelpetabell 1, f.eks. naturvitenskapelig språk (terminologi) [501.4](#)
(Tillatt løsning A: For å legge hovedvekt på det lokale eller gi en kortere notasjon til et bestemt språk, klassifiser i [410](#) , hvor det finnes fullstendige anvisninger
(Tillatt løsning B: For å legge hovedvekt på det lokale eller gi en kortere notasjon til et bestemt språk, plasser det foran [420](#) ved å bruke bokstav eller annet symbol slik at det kommer først i klassen. Fullstendige anvisninger finnes under [420-490](#))

Registertermer
Språk
Filologi

Mappet fra (551)

Noter med hierarkisk gjennomslagskraft (0)	Underordnede klasser (0)
--	--------------------------