



NHH

WP7 – Terminology integration **CLARINO oppstartsmøte 4. juni 2012**

Gisle Andersen

Marita Kristiansen

18.06.2012



NHHs bidrag: en nasjonal termportal

- With respect to terminology, there is at present no national **infrastructure**, despite the fact that an increasing number of scholarly domains face the threat of domain loss.
- This effort will take an important step towards overcoming the **fragmentation** problem in the terminology field and the need for coordination of terminological efforts at international and national levels.

NHHs bidrag: en nasjonal termportal (2)

- In CLARINO, NHH wants to establish a national infrastructure to **harmonize terminological language resources** and technologies, in response to the national responsibility now given to higher education to develop and disseminate Norwegian terminology.
- This will be a natural extension of already existing resources originating from the Norwegian Term Bank which in recent years have been further developed in projects such as KB-N, Mikroøkonomen, Termportalen and CLARA.
- It is not the aim to develop new terminologies per se, but to provide the **technical architecture needed for their integration and accessibility**, and to vouch for their **interoperability with the CLARIN and ISO/TBX standards**.

Terminology Services ('CLARINTerm')

- CLARINO will include a technical architecture that integrates a variety of **structured, mono- and multilingual terminology** bases via a **common interface**.
- The infrastructure will provide **public and controlled unified access** to **distributed resources** in a national grid-based architecture, with **links to European terminology resources via CLARIN**. Both distributed and centralised solutions will be offered.

Terminology Services ('CLARINTerm')

- CLARINTerm will cater for **autonomous bases** with a **common interface and search system**, as well as **integrate a variety of existing termbases into a central knowledge base** that contains entries for concepts and conceptual descriptions such as main terms and synonyms, definitions and concept relations.
- It will be the decision of each data provider to either have the data fully integrated into the central knowledge base or made accessible as a distributed base.
- Moreover, **hybrid solutions** may be offered, in which an **instantiation of CLARINTerm draws from both a local source and the central knowledge base**, but is presented uniformly as one resource.

Bakgrunn – tidligere arbeid

- Utredninger
 - Fagterminologi
 - Norsk i hundre
 - Mål og mening
 - Utredningsgruppe
- Pilotprosjekt
 - Termportalen
- Prosjektsøknad
 - Terminor (=CLARINTerm)
- Forskningsprosjekter/relaterte prosjekter
 - KB-N
 - CLARA
 - MetaNord

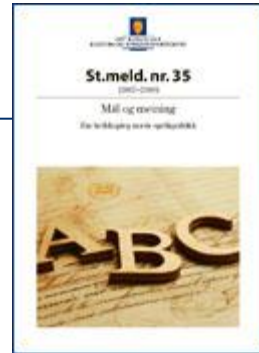


Språkrådet

Terminologiarbeid i
universitets- og høyskolesektoren
Rapport fra en utredningsgruppe



SPRÅR
13. mai 2012



Utfordringer og målsettinger

- Utfordringer
 - "å utvikle nye norske termar som får brei aksept i dei respektive fagmiljøa, og å gjere eksisterande og nye termar tilgjengelige på ein slik måte at dei også blir tekne i bruk i alle aktuelle samanhengar" ... Elektroniske termbankar er suverene som medium for å utvikla, lagra og spreia terminologi. (Mål og mening, s. 104)
 - "Dei terminologiske metodane varierte mykje, ofte gjekk arbeidet føre seg utan dokumenterte retningslinjer, og terminologien vart ofte ikkje gjord tilgjengeleg eksternt. Dei mange ulike dataformata som var i bruk, hemma effektiv søking og nettbasert presentasjon"
- Mange ulike terminologiresurser, ulike løsnings- og presentasjonsmåte
- Behov for koordinert tilgang til terminologi; det trengs en overbygning dvs. en nasjonal termportal
- Betrakte terminologi som en infrastruktur for forskning, utvikling og formidling

Termportalen – fakta

- Forarbeidet til prosjektet *Termportalen* startet ved Unifob AKSIS i 2007
- I 2008 fikk prosjektet delfinansiering fra Språkrådet
- Prosjektet formelt avsluttet 22. april 2009
- Termportalen er en **pilotversjon** av en nasjonal termportal
- System for **søking på tvers av termbaser**
- Oversikt over **terminologiske ressurser** (lenkesamling)

KB-N	Økonomi/administrasjon	8467
EØS-basen	70 fagområder, EØS-rettsakter	36776
NOT-basen	38 fagområder, Norsk termbank	30521
RTT-materialet		48314
		124078

Termportalen

[Forside](#) | [Ressurser](#) | [Termsøk](#) | [Om prosjektet](#) | [Aktuelt](#) | [Kontakt oss](#)

Du er her: [Forside](#) → [Ressurser](#)

Ressurser

Begrepsanalyse og terminologiarbeid

[Termlosen - kort innføring i begrepsanalyse og terminologiarbeid](#)

Gass, olje og energi

[NOT - Norsk terminologisk database](#)

[Oljedirektoratet: Begrep og definisjoner](#)

Geofag

[GeoLeksi](#)

[Meteorologileksikon](#)

[Statens kartverk](#)

Helse

[Helsenett.no](#)

Contents

1. [Begrepsanalyse og terminologiarbeid](#)
2. [Gass, olje og energi](#)
3. [Geofag](#)
4. [Helse](#)
5. [Informasjonsteknologi](#)
6. [Juridiske fag](#)
7. [Maritime fag](#)
8. [Miljø](#)
9. [Offentlig forvaltning](#)
10. [Organisasjoner](#)
11. [Samisk språk og kultur](#)
12. [Større terminologidatabaser](#)
13. [Tegnspråk](#)
14. [Økonomi/administrasjon](#)

Termportalen

[Forside](#) | [Ressurser](#) | [Termsøk](#) | [Om prosjektet](#) | [Aktuelt](#) | [Kontakt oss](#)Du er her: [Forside](#) → [Termsøk](#)

Termsøk

I prototypen av Termsøk er det mulig å søke både i enkeltressurser og på tvers av flere ressurser. Søkene kan spesifiseres etter domene. Res prototype består av fire store og flerspråklige terminologidatabaser. Termsøk er passordbeskyttet så lenge Termportalen er i valideringsfase være fritt tilgjengelig for alle eksterne brukere.

Søk

Søkestreng: (oblig.)	<input type="text" value="olje"/>
Høyretrunkert:	<input checked="" type="checkbox"/>
Termbase:	<input type="text" value="Velg termbase"/>
Språk:	<input type="text" value="Velg termbase"/>
Felt:	<input type="text" value="KB-N"/>

(Totalt 513 treff)

<u>Base</u>	<u>Språk</u>	<u>Felt</u>	<u>Tekst</u>
KBN	nob	hvd	oljekorrigert overskudd

Terminor

- Mål: å etablere en **nasjonal portal for norsk terminologi** og fagspråk
- NFR-søknad, Forskningsrådets program for nasjonal satsing på forskningsinfrastruktur i 2009
- NHH, UiB, Uni Computing
- Integrering av termbaser; søk på tvers av termbaser
- Gjennomføring av deler gjennom prosjektet CLARINO
- I gang med øk./adm. arbeidspakke



(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
NHH Termbase (NHH-T)	Terminology database	1600 termbase entries	no, en	NHH	Public and free	Public and free	NHH	a) 4 PW b) 4 PW	Fully updated termbase covering central concepts in microeconomics and other business and finance domains	regular use for educational purposes at NHH; formal public launch autumn 2010	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)
KB-N Kunnskapsbank for norsk økonomisk-administrativt domene	Terminology database	8467 termbase entries	no, en	NHH/Uni Digital	Public and free	Public and free	NHH	a) 8 PW b) 8 PW	Wide scope of concepts in business & economics, need for updating/quality check	high potential user value; in current use for research purposes only	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)

NOT-basen (Norsk termbank)	Terminology database	30521 termbase entries	no, en	Uni Digital	Public and free	Public and free	NHH	a) 8 PW b) 16 PW	Very wide coverage of domains and concepts including petroleum sector, not updated	high potential user value; in current use for research purposes only	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)
UHR-basen – Universitets- og høyskolerådets termbase	Terminology database	1000 termbase entries	no, en	UHR	Public and free	unknown	UHR	a) 4 PW b) 4 PW	Fully updated termbase covering central concepts in administrative domains	regular use for administrative purposes, many users	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)
RTT-materialet (Rådet for teknisk terminologi)	Terminology database	48314 termbase entries	no, en, fr +	Fagbokforlaget	unknown	unknown	unknown	a) 8 PW b) 16 PW	Wide coverage of technical domains, not updated	high potential user value; in current use for research purposes only	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)
EØS-EU-basen (EØS-sekretariatets terminologidatabase og norske oversettelser av rettsakter innlemmet i EØS-avtalen)	Terminology database	36776 termbase entries	no, en, fr	EØS-sekretariatet (EØS-/EFTA-seksjonen i Europaavdelingen, UD)	unknown	unknown	EØS-sekretariatet, en del av EØS-/EFTA-seksjonen i Europaavdelingen, UD	a) 8 PW b) 8 PW	Continuously updated, high-quality termbase; inclusion must be negotiated with EØS-sekr.	regular use for translation in UD; many external users; continuously updated	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)
SNORRE - Standard Norges termbase	Terminology database	? termbase entries	no, en	Standard Norge	unknown	unknown	Standard Norge	a) 16 PW b) 8 PW	Termbase being updated in ongoing project	new terminology resource; launch planned in autumn 2010	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)
KRIPOS-basen	Terminology database	? termbase entries	no, en, fr	KRIPOS' language section	unknown	unknown	KRIPOS' language section	a) 16 PW b) 8 PW	Continuously updated, high-quality termbase; inclusion must be negotiated with KRIPOS	regular internal use for policing purposes	EuroTerm-Bank (LV), Rikstermbanken (SE), Dan-Term (DK), IATE (EU)



Gass, olje og energi

[NOT – Norsk terminologisk database](#)

[Oljedirektoratet: Begrep og definisjoner](#)

Geofag

[GeoLeksi](#)

[Meteorologileksikon](#)

[Statens kartverk](#)

Helse

[Helsenett.no](#)

[Kompetansesenter for IKT i helse- og sosialsektoren](#)

[Norges Kyokushin Karate Organisasjon: Kyokushinkai terminologi](#)

[NKKN – Norsk Klassifisering, Koding & Nomenklatur](#)

[Pasienthåndboken: Medisinsk ordbok](#)

Informasjonsteknologi

[Begrep og definisjoner IT og helse](#)

[Dataspråk \(Språkrådet\)](#)

[Flerspråklig tabell med termer innen språk og tale](#)

[Norske oversettelser av termer fra Language Technology World](#)

[Språkteknologisk terminologi med definisjoner](#)

[UNINETT ABC: Ordliste](#)

Juridiske fag

[Sámi Láhkagiella](#)



Språkressurser Termportalen (forts.)

Maritime fag

[Faguttrykk i shipping](#)

[Fiskenavn på fem språk](#)

[Islandsk flerspråklig artsliste \(fisk\)](#)

[Kystverket: Begreper](#)

[Kystverket: Ord og uttrykk](#)

[MaritimStart](#)

[Norges Sildesalgslag](#)

[Shipping norsk/engelsk](#)

[Ship circular](#)

Miljø

[Eionet – GEMET Thesaurus](#)

[GRIP – Grønt i Praksis: Begreper](#)

Offentlig forvaltning

[Arbeids- og inkluderingsdepartementet: Sentrale begreper](#)

[Jernbaneverket: Definisjoner og begreper](#)

[KBT – Kollegiet for brannfaglig terminologi](#)

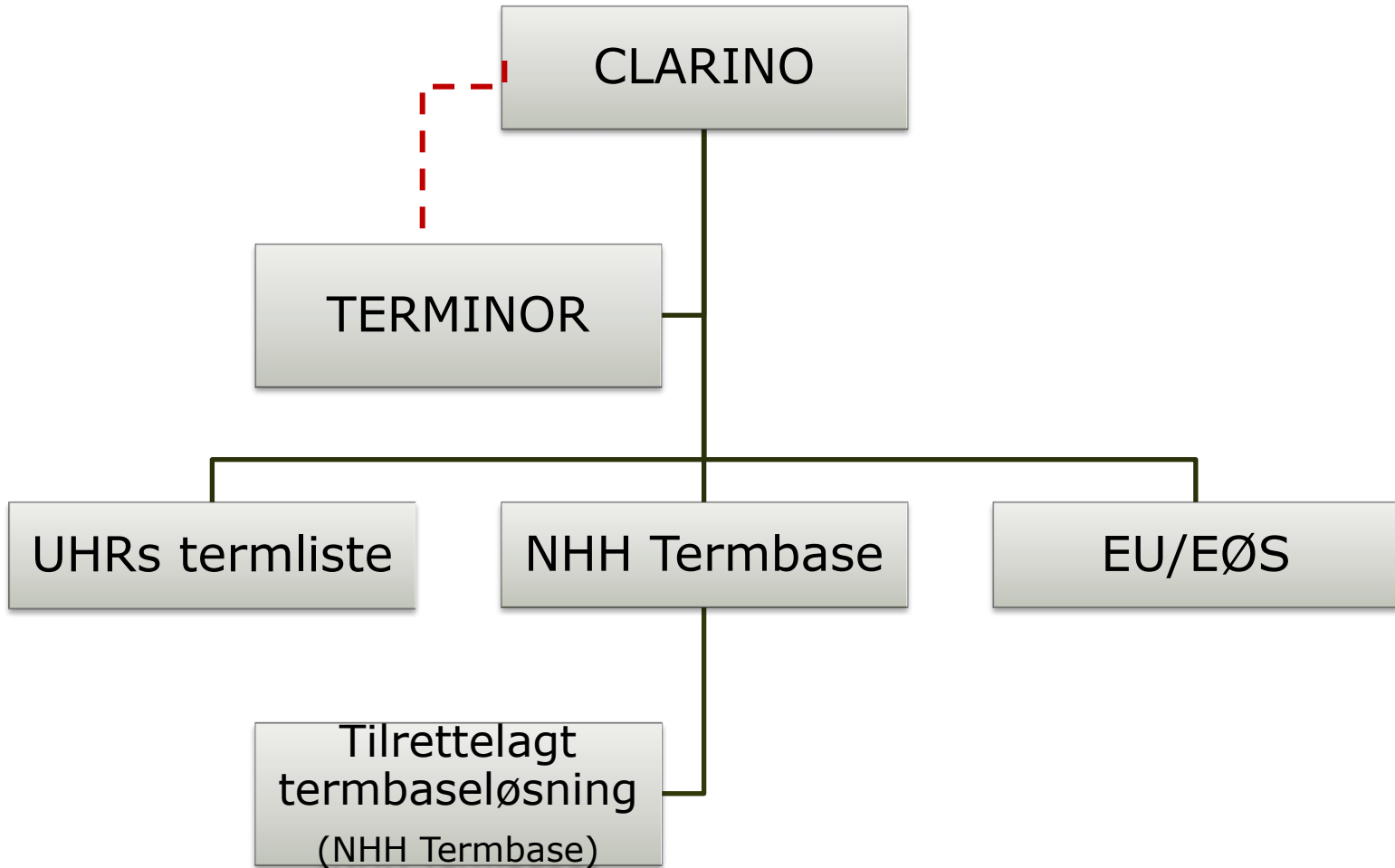
[KHiB-ordliste/-terminologi norsk/engelsk](#)

Termportalens funksjonalitet, tjenester og innhold (1/2)

- Én felles søkerute – mange lokale termbaser/-ressurser
- Søk i alle ressursene som er koblet sammen i portalen eller i utvalgte ressurser
- Søk i alle eller enkelte felt
 - term/uttrykk
 - domene (f.eks. IKT, biologi eller økonomi- og administrasjon)
 - subdomene (underinndeling innen økonomi og administrasjon: bedriftsøkonomi vs. samfunnsøkonomi)
 - språk (f.eks. norsk vs. engelsk)
 - kilde
- Kobling til begrepssystemer, tekstkorpus, nettsider, multimedieinnh.
 - «En nasjonal termportal må ha et system for å integrere semantisk strukturert informasjon i basene, for å navigere langs begrepsrelasjoner og for å visualisere generiske og partitive begrepssystem, helst i form av dynamiske begrepskart og ikke statiske bilder.» (Utredning 13. mars 2012: 9)



Termportalens funksjonalitet, tjenester og innhold - mulig på ulike nivå (2/2)





Nordiske initiativer – Sverige



RIKSTERMBANKEN
Alla termer på ett ställe

[Om Rikstermbanken](#) | [Kontakt](#) | [English](#) | [svenska](#)

- Ingångssidan
- Enkel sökning
- Utökad sökning
- Källor
- Senaste nytt
- Frågor och svar
- Hjälp

Välkommen till Rikstermbanken

Sveriges nationella termbank

Sök i Rikstermbanken


enbart bland termer även i all annan text

Hittar du inte vad du söker eller har du termproblem du vill ha hjälp med? Gå med i Terminologifrämjandet och använd dig av TNCs termtjänst. [Läs mer.](#)

 Terminologicentrum TNC förvaltar, utvecklar och underhåller Rikstermbanken. Senaste uppdatering: 19 mars
Systemversion 0.97.41J © Terminologicentrum TNC



Nordiske initiativer – Finland

 Kirjaudu sisään tai luo tunnus



TIETEEN TERMIPANKKI

VETENSKAPSTERMBANKEN I FINLAND
BANK OF FINNISH TERMINOLOGY IN ARTS AND SCIENCES

Projektisivu

Keskustelu

Lue

Lähdekoodi

Näytä historia

Etsi



Etusivu

▼ Navigointi

[Ohjesivuille](#)

[Keskustelualue](#)

[Aihealueet](#)

[Lähdeluettelo](#)

[Tuoreet muutokset](#)

[Satunnainen sivu](#)

► [Termipankki](#)

► [Työkalut](#)

Bank of Finnish Terminology in Arts and Sciences

[Etusivu suomeksi](#)

This is the Pilot Wiki of the [Bank of Finnish Terminology in Arts and Sciences](#).


The Bank of Finnish Terminology in Arts and Sciences (BFT) is a multidisciplinary project which aims to gather a permanent terminological database for all fields of research in Finland. The project has created this Semantic MediaWiki platform, which offers a collaborative environment. This means that anyone can freely use it and also participate in the discussion about terms.

The project is being carried out in three pilot projects: botany, jurisprudence, and linguistics.

Fields

- [Botany](#) (3 304 concepts records)
- [Linguistics](#) (3 275 concept records)
- [Jurisprudence](#) (15 concept records)
- [Language Technology](#) (173 concept records)

Nordiske initiativer – Danmark


Copenhagen Business School
HANDELSHØJSKOLEN

DANTERM BANK

Search

- Front page
- About the project
- Knowledge acquisition
 - Automatic knowledge extraction
 - Automatic merging and validation of data
- Knowledge validation
- Knowledge dissemination
- Project funding
- Project participants
- Publications
- Media archive
- Contact

DanTermBank

Establishing a Danish Terminology and Knowledge Base

DanTermBank
 Phase 1: Establishing a Danish Terminology and Knowledge Base The aim of the project is to develop methods for automatic knowledge extraction, automatic construction and updating of ontologies. In the project methods will be developed for automatic merging of terminological data from various existing sources, as well as methods for target group oriented knowledge dissemination.


The research carried out in the current project is a prerequisite for establishing a national Danish term bank which can ensure development and quality of Danish LSP. When the term bank has been established, it will form the basis for various other research projects. This phase of the project is supported by the VELUX FOUNDATION.

The DanTermBank project consists of three subprojects, which you can learn more about by navigating through the left menu, or clicking the links below:

Subproject 1: Knowledge Aquisition

Subproject 2: Knowledge Validation

Subproject 3: Knowledge Dissemination



News

DTB Project team at TKE conference in Madrid 2012

DanTermBank project team participates in Terminology and Knowledge Engineering (TKE) Conference in Madrid 2012.

[See all news](#)

Establishing a national Danish term bank

The research carried out in the current project is a prerequisite for establishing a national Danish term bank which can ensure development and quality of Danish LSP

New website

The DanTermBank project launches a



Samkjøring og kvalitetssikring av terminologiske data

For å kunne etablere en termportal med et tilstrekkelig omfang og med ressurser av en tilstrekkelig høy kvalitet må det utvikles metoder for automatisk ekstraksjon av kunnskapselementer

- utfordringer ved samkjøring av eksisterende terminologiske data
 - a) data vil ha ulik struktur, ulike lagringsformater og være av ulik kvalitet

UHRs termliste vs. NHHs begrepsbaserte termbase

- b) data basert på ulike kilder kan føre til 'falske dubletter' og 'falske ekvivalenter'

(Lassen, T., B. Nistrup Madsen & H. Erdman Thomsen 2011. *Automatisk samkjøring og kvalitetssikring af data i en term- og vidensbank*)

Begreps- og termendringer

STORE NORSKE LEKSIKON (snl.no)

samfunnsøkonomi vitenskapen om den økonomiske virksomhet i samfunnet.

Samfunnsøkonomi kaltes i Norge tidligere sosialøkonomi og før det statsøkonomi. I andre skandinaviske land brukes betegnelsen nasjonaløkonomi. Det engelske navn på faget er *economics* ...

«**Sosialøkonomi er blitt samfunnsøkonomi**

Navnet skjemmer ingen, sier gammel visdom. Men akkurat *det* er åpenbart ikke universitetets sosialøkonomer enige i. De har nemlig gått hen og døpt faget sitt *samfunnsøkonomi*. Og det tradisjonsrike Sosialøkonomisk institutt, som har fostret to nobelprisvinnere, skal nå kort og godt hete *Økonomisk institutt*.» (Uniform nr. 7/2000, <http://www.uniform.uio.no/>)

ISO-standarder

Standarder for terminologiens metode og grunnleggende begreper

- ISO 704 – Terminology work – Principles and methods
- ISO 1087 – Terminology work – Vocabulary

Standarder for utveksling av terminologi

- ISO 12620 – datakategorier for lagring av terminologiske data
- ISO 12200 – definerer MARTIF; SGML-basert standard for utveksling av terminologiske data
- ISO 16642 – retningslinjer for tekstmerking av terminologi

→ utvekslingsformatet TBX (TermBase eXchange), som også er kompatibelt med rammeverket TMF (Terminological Markup Framework)

(Madsen & Thomsen 2009. «Terminological Concept Modelling and UML Diarams.» I *International Journal of Metadata, Semantics and Ontologies (IJMSO)*, Vol. 4, No. 4, 2009.)

TBX

- TBX = TermBase eXchange
- Åpen, XML-basert standard
- Utviklet av LISA (Localization Industry Standards Organization)
- Vedtatt som internasjonal standard (ISO 30042) for merking og utveksling av terminologiresurser
- Kompatibelt med rammeverket TMF (Terminological Markup Framework; ISO/TC37: ISO 12620/12200/16642)
- Relevante ressurser som RNG-skjema, DTD og XCS-filer, samt valideringsprogram (TBX-sjekker) er tilgjengelig fra ressursiden <http://www.ttt.org/tbx/>

Aktuelle teknologileverandører for termportal

- EuroTermBank (<http://www.eurotermbank.com/>)
- IATE – InterActive Terminology for Europe (<http://iate.europa.eu/>)
- i-Term – system for terminologihåndtering og begrepsmodellering utviklet ved DANTERMcentret ved CBS
- Norsk termbank (NOT) / KB-N / NHHs termbase
- Rikstermbanken (<http://www.rikstermbanken.se/>)
- TEPA – finske terminologisentralen / Helsinki universitet
- TERMIUM – Canadas nasjonale termbank

Kommersielle teknologileverandører

- across (Ahead Software AG, Tyskland)
- CATS – Computer-Aided Terminology System (CATS, Tyskland)
- GFT DataTerm (GFT GmbH, Tyskland)
- TermExt (termekstraksjonsteknologi, IBM, USA)
- Lingo (Julia Emily Software, Frankrike)
- LingTools (Sietec Systemtechnik, Tyskland)
- LogiTerm (Terminotix Inc., Canada)
- MoBiDic (MorphoLogic, Ungarn)
- MTX™ (LinguaTech, USA)
- MultiTerm (TRADOS GmbH, Tyskland)
- SDL TermBase (SDL, Storbritannia)
- TEMIS (termekstraksjon, kunnskapsorganisering, TEMIS, Tyskland)
- TermStar (STAR AG, Sveits)
- Termwatch (ATRIL Software SL, Spania)
- UniTerm (Acolada GmbH, Tyskland)
- Xerox Terminology Suite (XTS) (Xerox Multilingual Knowledge Mgmt, FR)

Igangsatt arbeid

- Konvertering av pilotressurs i MetaNord-prosjektet til TBX (KdS)
- Lokalt oppstartmøte
- Kontakt/avtaler med (nye) sentrale terminologileverandører
 - Universitets- og høyskolerådet
 - Norges Bank
 - Norges forskningsråd
 - ++
- Igangsatt gjennomgang av ressursoversikt
- Rekruttering

Målsettinger for Y1

Administrativt

- Rekruttering av teknisk medarbeider (1 august 2012)
- Igangsette henvendelser til ressursleverandører
- Avtaleverk og lisenser – alternativer i CLARINO
- Arrangere teknisk seminar om terminologihåndteringssystem (okt.)

Språkressurser

- Oppdatert ressursoversikt med prioriteringer



Teknisk

- Tekniske spesifikasjoner for terminologiresurser i CLARIN(O)
- Avklaring av ressursintegrering i CLARIN(O)
- Beslutning vedr. terminologihåndteringssystem
- Pilotversjon av nasjonal portal for terminologi