

Anteckningar från referensgruppsmöte tis 2 september 2008

Närvarande: Henrik, Jussi, Lars, Ola (t.o.m. punkt 2), Rickard, Viggo

1. Språkpolitisk omvärldsbevakning

- Ny norsk språkpolitisk utredning och Språkbanken i Norge (Rickard)

Rickard redogjorde för en norsk språkpolitisk utredning med långtgående förslag för språkteknologins del. Utredningen föreslår att en norsk språkbank etableras med resurser för norsk språkteknologi. Språkbanken blir verklighet 1 januari 2009.

Utredningen: <http://www.regjeringen.no/nb/dep/kkd/dok/regpubl/stmeld/2007-2008/stmeld-nr-35-2007-2008-.html?id=519923>

Om språkbanken: <http://www.spraakbanken.uib.no/>.

Det borde göras en sammanställning av språkpolitiska dokument om språkteknologi. *Rickard kollar om det är något som kan tas fram i arbetsgruppen om språkvård och språkteknologi i Norden.*

Henrik berättade att han sett en prototyp till ISO:s termbank, som han inte fann så imponerande

2. Språkteknologi.se

- Lägga upp arbetsdokument (Ola, Rickard)

Ola har skapat en intern mapp på språkteknologi.se/referensgruppen där mötesanteckningar och andra dokument för referensgruppen läggs upp. Inloggningsuppgifter delades ut.

- Avdelning för SNK/Blark (Ola)

Ola har också skapat en publik avdelning för arbetet med SNK/Blark där information om projektet och dess resultat ska läggas upp.

- Talsyntes (Ola)

Vi har nu ett avtal om talsyntes till Språkrådets webbplatser så att text kan läsas upp. Det fungerar för www.sprakradet.se men inte ännu för Språkteknologi.se.

Problemet är att vi inte har någon som arbetar med det tekniska. Robert och Peter i Göteborg kan bara göra saker i mån av tid. Det är ett problem som måste lösas på sikt.

3. Terminologi

- Norddoknet-termer (Henrik)

Vi hade fortsatta diskussioner om språkteknologiska grundbegrepp, t.ex. språkteknologi kontra data-/datorlingvistik, talteknologi kontra nlp (textlingvistik flertydigt).

Som ett steg i arbetet med språkteknologisk terminologi översätter Jussi Filts finska definitioner av språkteknologiska begrepp till svenska.

Rickard har kontaktat Kimmo Koskienniemi som vill samarbeta om tvärspråklig terminologi i det experimentella terminologisystem i wiki-form som han skapat.

Se: <https://kitwiki.csc.fi/twiki/bin/view/Terms/WebHome>

Henrik kollar terminologisystemet till nästa gång, gärna också vi andra.

- Terminologin på Språkteknologi.se (Ola)

Terminologin på Språkteknologi.se bygger på LT-worlds termer och är harmoniserade med Norddoktermerna. Ola visade hur norddoktermerna kan användas för tvärspråklig sökning på <http://www.ling.helsinki.fi/filt/info/search/search.php?sv>

- Termer för språkteknologisk infrastruktur (Rickard)

Rickard presenterade ett underlag för vilka termer som används för språkteknologisk infrastruktur och hur de definieras, t.ex. språkresurser-teknologiska resurser m.m. *Vi beslutade att ta med dessa termer i vårt prioriterade terminologiarbete.*

4. Ansökningar

- SFS-ansökan (Viggo)

SFS-ansökan är ett samarbetsprojekt mellan många språkteknologiska forskningsmiljöer (inkl Språkrådet) i Sverige som löper över 5 år. Om flerspråkig informationsbehandling inom myndigheter. De ansökta medlen är på ca 7 milj. per år. Svar kommer under hösten.

- Språk och internet (Rickard)

Rickard och Jussi har lämnat in en ansökan till .SE om medel för att utreda situationen vad gäller språklig Internetinfrastruktur för språken i Sverige. En workshop med syfte att ta fram aktuella användningsfall ingår. Besked i nov.

- Folkets engelsksvenska lexikon (Viggo)

Viggo har fått 200000 från .SE för att konstruera Folkets engelsksvenska lexikon som bygger på eng-sv Lexin som KTH fått rättigheter till att använda för projektet. Viggo framförde en önskan att inte Lexinmaterialet görs tillgängligt (utom för forskare) förrän efter det att Folkets lexikon har släppts fritt någon gång under nästa år. Det skulle annars kunna minska intresset för att bidra till Folkets lexikon.

5. Tillgängliggörande av Lexin

Lars har talat med Sven-Göran Malmgren som inte tror att Språkbanken har några upphovsrättsliga anspråk. Han tvivlar på att Språkdata har publiceringsrättigheterna. Men han förstår inte varför man ska göra materialet fritt annat än för forskning. *Referensgruppens förslag: Vi gör materialet fritt och ser vad som händer och drar tillbaka det om det tar hus i helvete.*

Andra ordböcker än eng-sv: enbart för forskning men se genom fingrarna vid dold användning i kommersiella tillämpningar

Språkrådet har också bett om juridisk hjälp för att reda ut problemen kring upphovsrätt och andra juridiska frågor.

6. Konferenser och seminarier

- Nordiskt seminarium om översättningsteknologi i Oslo 23-24 okt. Alla i referensgruppen är välkomna att delta.
- SLTC 20-21 nov. Rickard är med och arrangerar en workshop om språkkontroll.
- Nodalida09 15-16 maj i Odense. Språkrådet planerar att arrangera en workshop om nordiskt deltagande i Clarin (deadline för workshop 1 dec).

7. Informationsarbete

- Hur presentera gruppen utåt?

Henrik tog på förra mötet upp frågan om hur referensgruppen ska presentera sig utåt. Vi bestämde att vänta med ett sådant utåtriktat arbete tills vi kommit igång med verksamheten mer, särskilt den terminologiska verksamheten som är ett utåtriktat arbete.

Vi diskuterade också vikten av att profilera den språkteknologiska verksamheten vid Språkrådet i allmänhet. Det gäller både utåt och inåt. Sedan Språknämnden blivit Språkrådet och utvidgat sin verksamhet har det kommit nya verksamheter som gör att språkteknologi har hamnat i bakgrunden som en svårutskiljbar del av den svenska språkvårdens verksamhet. Rickard har talat med Olle om att också profilera språkteknologin mer självständigt eftersom den inte bara gäller svenska utan också andra språk i Sverige.

I likhet med Klarspraksgruppen och andra grupper på Språkrådet skulle det vara bra om vi hade något departement som kramar oss och kräver att vi får plats.

- Bok om språkteknologi klar i höst. (Rickard)

Titeln på antologin om språk, it och språkteknologi blir "Tekniken bakom språket". Undertitel ej klar men "Språkteknologi för språken i Sverige" är ett förslag. Norstedts nöjd med huvudtiteln men inte med undertiteln, så den kanske blir en annan. Boken ska finnas i diskarna inför vårterminen.

8. Övrigt

- Nästa möte

Det är osäkert om vi kan ha ytterligare ett möte i år pga tillfälligt dålig ekonomi.

Kan vi inte ha det skulle ett möte över nätet vara ett alternativ. Rickard undersöker möjligheterna.

- Testning av rikstermbankens gränssnitt

Henrik skickar snart ut information om testning av rikstermbankens gränssnitt. Gränssnittet ska vara klart i slutet av nov.