

Anteckningar från möte med referensgruppen i språkteknologi 5 dec 2008

Närvarande: Lars Borin, Rickard Domeij, Viggo Kann, Jussi Karlgren, Ola Karlsson, Henrik Nilsson, Peter Svanberg, Åsa Wengelin

Rapportering

Alla rapporterade kort om vad som händer i världen vad gäller språkteknologi:

- Rickard berättade att Lena Ekberg blir ny chef på Språkrådet fr.o.m. 1 feb 2009
- Åsa berättade att hon fått nytt jobb som lektor i handikappvetenskap på Högskolan i Halmstad fr.o.m. sommaren 2009.
- Rickard berättade att Antologin Tekniken bakom språket är klar för utgivning före jul.
Rickard skickar ut erbjudande till alla.
- Rickard berättade att det Nordiska seminariet om översättning i Oslo blev lyckat trots att inbjudan var sent ute och att det krockade med GSLT-veckan.
- Viggo, Lars och Rickard berättade om SLTC 2008 med workshoppar om språkteknologisk infrastruktur och om skrivstöd. Där presenterades också en ny språkteknologisk tidskrift Nejl (<http://www.nejl.ep.liu.se/>) och ett förslag på ev. magisterutbildning i språkteknologi.
- Lars berättade också kort om ett symposium om infrastruktur för humanvetenskaperna på KB arrangerat i samarbete mellan KB, VR och RJ. Se <http://www.kb.se/bibliotek/kurser-konferenser/konferens-seminarier/Infrastruktur-i-samverkan/> och kommentar från RJ:s forskningsdirektör Mats Rolén här: <http://www.rj.se/svenska/1023/var/newsID/99>.
- Lars berättade att rapporten om en svensk nationell korpus är klar och att ansökan om medel för fortsatt arbete med Blark/SNK är inlämnad till VR. **Lars uppmanade alla att lämna förslag på namn på personer som är lämpliga att hjälpa VR värdera ansökan.**
- Lars berättade kort om den svenska trädbanken som nu går att ladda ner från Språkbankens webbplats. Dessutom kommer Saldo i en ny, mera korrekt version 1.1.
- Viggo berättade att SFS-ansökan inte beviljades men fick bra betyg. Hercules Dalianis har däremot fått pengar från Vinnova för ett delprojekt i SFS-ansökan om automatisk flerspråkig e-posthantering där också Viggo är med på ett hörn.
- Rickard och Jussi berättade att de fått pengar för att skriva en rapport om språk och internet under 2009 och också arrangera en workshop under våren.
- Henrik och Peter berättade att arbetet med Termbanken är något försenat. Instruktioner och enkät för testare kommer före jul. Invigning i mars.
- Viggo berättar att arbetet med folkets engelsksvenska lexikon är något försenat på grund av mer arbete med kodning än beräknat. Inga problem dock.
- Rickard berättade att Svenskt avstavningslexikon nu finns inskannat i elektronisk form. Språkrådet har planer på att göra det tillgängligt för allmänheten via ett webbgränssnitt. Mikroverkstedet har gjort upp med förlaget om att använda det i sina skrivstödsprodukter. Mikrov. ger Språkrådet tillgång till lexikonet i expanderad form.
- Rickard berättade att han fått inbjudan till the Language Technology Days organised by Unit E1 - Language Technologies and Machine Translation, on 14-15 January 2009 in Luxembourg. Där presenteras närmare call 4 inom fp7 i EU:s nya forskningsprogram med satsningar på maskinöversättning och flerspråkighet på webben. Se http://cordis.europa.eu/fp7/ict/language-technologies/fp7-call4_en.html

- Henrik berättade om att han besökt 2 terminologikonferenser i Canada, bl.a. Tama som hade termbanker som tema.
- Henrik berättade att det finns planer på att göra terminologi till en egen modul på språkteknologiprogrammet i Uppsala. Han visade förslag på examensarbeten om terminologi. Flera förslag efterfrågades. Två uppsatser om automatisk termextraktion av Mirza Škornja och Henrik Nyh är klara.
- Jussi berättade att ett lektorat i språkteknologi/språkvetenskap utlysts på SU.
- Viggo berättade att ett lektorat med möjlig inriktning på språkteknologi utlysts på KTH.
- Jussi berättade att han numera sitter med i KB:s styrgrupp.
- Peter demonstrerade Rikstermbanken kort under pausen.

Uppföljning från förra mötet

- Rickard har talat med de nordiska språknämndernas arbetsgrupp om språkteknologi om att göra en sammanställning av språkpolitiska dokument som berör språkteknologi. Gensvaret var positivt men inget är bestämt om när det ska göras. Men Rickard tror att han själv kommer att behöva göra en sådan sammanställning under 2009 inom utredningsarbetet om språk och Internet.
- Jussi har påbörjat översättning av Filt-definitionerna till svenska.
- Henrik har tittat lite på Kimmos webbaserade terminologisystem och tycker att det kan vara värt att testa.

Planer för 2009

Rickard delade ut en lista på de uppgifter arbetsgruppen sagt sig vilja arbeta med och presenterade Språkrådets planer för 2009:

- Utredning om språk och internet med arrangerande av workshop under våren. Detta arbete kommer att pågå på knapp halvtid under 2009.
- Utveckling av gemensam taxonomi för de nordiska språknämndernas frågedatabaser (nordiskt harmoniseringsarbete gällande lagring och hantering av språkfrågor och svar). Danska språknämnden gör huvudarbetet. Språknämndernas arbetsgrupp är styrgrupp.
- Workshoppar på Nodalida och på Nordterm under våren. Rickard ansvarar för den första. Torbjörg, Gisle och Henrik är inblandade i den andra (se Terminologiarbete och planering av workshop nedan).
- Lobbning för satsning på språkteknologisk infrastruktur? **Viktigt men oklart hur detta ska göras. Vi får lägga upp en strategi. Rickard ansvarar.**
- Terminologiarbete (arbete med språkteknologiska termer, definitioner och samarbetsverktyg i anslutning till de nordiska webbplatserna om språkteknologi). Henrik ansvarar.
- Språkteknologi.se och webb om språk och Internet. Språkteknologi.se kräver löpande underhåll och behöver kanske också omarbetas med tanke på målgrupp vilket är något som behöver diskuteras mer. **Mycket av webbarbetet 2009 kommer att göras i anslutning till arbetet med utredningen om språk och Internet. Den informationen kommer sannolikt att ligga på Språkrådets webb. Ola ansvarar.**
- Fortsatt arbete med tillgängliggörande av Språkrådets resurser för språkteknologisk användning. Vi hann inte prata om det särskilt mycket men **men vi har ännu inte fått diskutera frågan med jurist som vi begärt. Rickard tar tag i det.**

- Examensarbete om automatisk nyordsexcerpering på KTH ska göras under året. **Viggo är handledare.**

Planering av infrastrukturworkshop på Nodalida

Rickard har skickat in ett förslag på workshop tillsammans med nordiska kolleger: Nordic perspectives on the CLARIN infrastructure of common language resources. Om den antas kommer den att behöva planeras under våren. Workshopen ska handla om nordiskt/baltiskt samarbete kring Clarinprojektet. Se:

<https://kitwiki.csc.fi/twiki/bin/view/Nealt/SigInfraNodalida2009WorkshpProposal>.

Terminologiarbete och planering av workshop om språkteknologisk terminologi på Nordterm 10-11 juni 2009

Workshop

Gruppen lade fram olika förslag på idéer till workshop i anslutning till Nordterm. Temat är ontologier och taxonomier, vilket passar väl till det vi vill göra och som omfattar frågor om samarbete, taxonomi/ontologi, terminologiverktyg och standarder kring språkteknologisk och språkvetenskaplig terminologi kring norddoknetwebbplatserna och de nordiska språknämndernas frågedatabaser. Förslag och frågor:

- Tidpunkt och längd: Vid sidan av eller köra parallellt? Minst en halvdag, mer om vi också ska ha arbetsgrupper som diskuterar termval, definitioner m.m. Ska det vara mer än en halvdag måste workshopen ligga vid sidan av konferensen.
- Utställning med olika terminologiverktyg ger workshopdeltagarna möjlighet att orientera sig i utbudet utan att det behöver vara en del av programmet. Det hindrar inte att erfarenheter kring arbete med enskilda verktyg presenteras på workshopen.
- **Rickard tar upp frågan på mötet med språknämndernas arbetsgrupp** och föreslår att Torbjörg Breivik (Språkrådet i Norge) och Gisle Andersen (ansvarig för termverktygsgruppen i Nordterm) diskuterar saken och kontaktar Henrik.
- **Henrik skickar ut "call" till Nordterm** så fort det finns.

Se http://www.cbs.dk/forskning_viden/konferencer/nordterm_2009

Terminologiarbete

Henrik ansvarar för att arbetet går vidare:

- **Jussi fortsätter arbetet med att översätta Filt-termerna** med fokus på centrala termer för användning på Språkteknologi.se och de nordiska dokumentationscentrumen.
- Vi börjar **pröva Kimmos webbaserade terminologisystem** utifrån Jussis översättningar.

Planering av utredning om språk och internet

- Styrgrupp och expertgrupp ska tillsättas. Referensgruppen uppmanas komma med förslag.
- Gruppen hade en kort diskussion om datainsamlingsmetoder. Experter inom olika områden med relevans för utredningen ska intervjuas och/eller ombedas att skriva ett utlåtande/delrapport. Möjligen ska också enkät skickas ut, men det är tveksamt om det ger mer än det kostar. **Rickard gör en lista på relevanta områden med tänkbara experter och skickar ut till andra i referensgruppen för synpunkter och förslag.**

- Jussi berättade om en workshop om användningsfall som planeras under våren som ett sätt att ta fram användarorienterade scenarier som sen ska ligga till grund för prioriterade förslag på framtida satsningar på teknikutveckling och andra projekt. Både tekniker, pedagoger, myndighetspersoner, tillgänglighetsfolk, användare inom relevanta områden m.m. ska bjudas in. Workshoppen ska hållas i mars-april.

Övrigt

- **Rickard uppmuntrade alla som är intresserade av nordisk/baltisk språkteknologisk infrastruktur att söka medlemskap i Sig-infra** som också är forum för planeringen av infrastrukturworkshoppen på Nodalida. Instruktioner skickas till alla.
- **Gruppen beslutade att använda Doodle** som verktyg för mötesbokning i fortsättningen.
- Rickard berättade att han kontaktats av företaget Tricom Solutions som vill diskutera samarbete kring utvecklingen av språkteknologiska resurser i Sverige.
- Henrik meddelade att han och andra i arbetsgruppen (Ola, Peter?) slutfört arbetet med att ge synpunkter på Vervas val av termer för begrepp på myndigheternas webbplatser.

Nästa möte

Rickard skickar ut förslag på datum.