



GÖTEBORGS UNIVERSITET
GSLT

28 januari 2004

GÖTEBORG UNIVERSITY Graduate School of Language Technology

Kulturminister Marita Ulvskog,
Näringsminister Leif Pagrotsky,
Utbildningsminister Thomas Östros

103 33 STOCKHOLM

Behovet av en sammanhållen strategi för svensk språkteknologi och ökade forsknings- och utvecklingsinsatser har påtalats i flera sammanhang både av det språkteknologiska forskarsamhället och i de utredningar som berört språkteknologiområdet, såsom Nuteks (1999) om maskinöversättning och Mål i mun-utredningen (SOU 2002: 27). Av skäl som sammanfattas i bifogade skrivelse är det viktigt att en sådan nu etableras och att Sverige förvaltar och förstärker de framgångsrika resultat som uppnåtts. Vi anhåller med detta brev också om att få träffa representanter för berörda departement och diskutera behov och åtgärder för att utveckla svensk språkteknologi.

Lars Ahrenberg
Professor, Linköpings universitet
Biträdande föreståndare, GSLT

Robin Cooper,
Professor, Föreståndare,
Sveriges nationella forskarskola
i språkteknologi GSLT

Anna Sågvall Hein,
Professor, Ordförande i det tidigare
programrådet för VINNOVAs program
i språkteknologi

Olle Josephson
Föreståndare, Svenska Språknämnden

Bengt Waernulf
Styrelseordförande, Centrum för talteknologi, CTT

Sverige behöver en strategi för språkteknologi

Språkteknologi är ett tvärvetenskapligt forskningsområde av avgörande betydelse för utvecklingen av de informations- och kommunikationstjänster som utgör grunden för visionerna om 24-timmarsmyndigheten, ett informationssamhälle för alla, och den semantiska webben. Exempel på språkteknologiska insatsområden är interaktion mellan människan och datorn i tal och skrift, informationssökning och informationsfiltrering, hjälpmedel för kommunikationshandikappade, språkundervisning samt maskinöversättning. Det är också, som utredningen SOU 2002:27 Mål i mun visar, ett område med betydande kulturella och språkpolitiska implikationer eftersom de tjänster som kan erbjudas står i proportion till de resurser i form av kunskap, utbildade människor, databaser och programvara som en språkgemenskap förfogar över. Av dessa skäl är det oerhört väsentligt att Sverige har en långsiktig strategi för språkteknologi med stöd i Sveriges riksdag.

Behovet av en sammanhållen strategi för svensk språkteknologi har påtalats i flera sammanhang både av det språkteknologiska forskarsamhället och i de utredningar som berört språkteknologiområdet, såsom Nuteks (1999) om maskinöversättning och Mål i mun-utredningen. Den senare anvisar också förslag på hur en sådan strategi skulle kunna realiseras, bland annat genom inrättandet av ett statligt finansierat språkteknologiskt sekretariat med samordnande funktion. De senaste årens utveckling inom svensk forskningsfinansiering och internationell forskning och utveckling inom språkteknologin gör behovet än mer angeläget.

Svenska forskningsfinansiärer har tagit flera initiativ som främjat utvecklingen av svensk språkteknologi. Vi kan här nämna de olika forskningsprogram som genomfördes under 90-talet av HSFR och Nutek i samarbete. 1996 startades Centrum för talteknologi (CTT) som en av VINNOVAs 10-åriga satsningar på kompetenscentra. CTT, med KTH som värd utgör ett framgångsrikt initiativ för samarbete mellan akademisk forskning och industriell utveckling. Sveriges nationella forskarskola i språkteknologi, GSLT, startades 2001 som en av sexton nya forskarskolor med statlig finansiering. GSLT med Göteborgs universitet som värdhögskola har samlat i stort sett alla lärosäten i Sverige som har forskarutbildning inom språkteknologi, både humanistiska och tekniska, i en gemensam utbildningsinsats och därtill Växjö universitet och högskolorna i Skövde och Borås som tidigare inte haft forskarutbildning inom området.

Flera av de framgångsrika satsningarna på slutet av 90-talet är snart avslutade. VINNOVAs kompetenscentrum CTT avslutas 2006. VINNOVAs pågående språkteknologiska program upphör med utgången av 2004 och fortsättningen är högst oviss. VINNOVA har för femårsperioden 03-07 definierat ett antal tillväxtområden och kunskapsplattformar inom vilka satsningar skall ske. Språkteknologi omnämns som en del av tillväxtområdet Programvaruprodukter men väntas inte bli prioriterat. Vetenskapsrådet finansierar för närvarande ett mindre antal begränsade projekt. Vad gäller GSLT finns inga beslut för tiden efter 2007. Den fortsatta forskningsfinansieringen inger således oro och en strategi behövs för att fullfölja de viktiga satsningar som hittills gjorts.

Dagens språkteknologi kräver också stora satsningar, inte minst i uppbyggnaden av den nödvändiga infrastrukturen: text och tal som är uppmärkt med nödvändig språklig information, från det fonetiska till det semantiska, analysatorer för att åstadkomma detta och utvärderingsmetoder och verktyg för att bedöma systems prestanda ur olika perspektiv. Ett stort kliv i utvecklingen tas för de stora språken när man nu bygger upp sådana resurser. Den franska satsningen Technolanguage omfattar exempelvis drygt 20 miljoner euro per år i tre år. Även Norge genomför en stor satsning inom språkteknologi med ett program som omfattar 66 miljoner NOK under fem år.

Sverige har hittills saknat ett permanent sammanhållande organ för språkteknologi. Ett språkteknologiskt sekretariat kan öka samarbetet i Sverige mellan näringsliv, forskning, högre utbildning och myndigheter, agera för svensk språkteknologi och ge Sverige en högre profil på

området internationellt. Sverige har i EU-sammanhang varit framgångsrikt när det gäller deltagande i språkteknologiska forskningsprojekt men haft svårare att hävda sig i de större samverkansprogrammen som exempelvis MLIS-programmet (Multi-Lingual Information Society). Vi har alltså inte kunnat påverka programmen särskilt mycket eller rätt kunnat utnyttja de resurser som ställts till förfogande. Sverige riskerar i dag att hamna utanför pågående europeiska initiativ som bland annat gäller etablerandet av ett språkteknologiskt nätverk som ett ERA-nät (European Research Area net) för att stärka samarbete och samverkan mellan nationella europeiska forskningsfinansieringsorgan.

Element i en strategi för svensk språkteknologi

Enligt vår mening måste en språkteknologisk strategi eller plattform vila på fem hörnpelare:

Forskning och utveckling. Flera tidsbegränsade svenska satsningar har gjorts. Nu behövs en rejäl och långsiktig satsning för att ta vara på och fullt ut utnyttja den kompetens som byggts upp i Sverige och som ytterligare kommer att förstärkas med de doktorer som blir färdiga under de närmaste åren. Detta gäller såväl grundforskning som behovsmotiverad forskning.

Grund- och forskarutbildning. Sverige är i dag förhållandevis väl försett både vad gäller grundutbildning och forskarutbildning på området. Utan ett långsiktigt perspektiv riskerar dock det nationella engagemang och den dynamik som skapats i forskarutbildningen genom tillkomsten av GSLT att gå förlorad.

Kommersialisering och kunskapsöverföring. Kommersialisering av språkteknologiska forskningsresultat hör samman med utvecklingen av å ena sidan generella IKT-tjänster och å andra sidan förändring av språkligt inriktade arbeten som traditionellt utförs av personer med en humanistisk utbildning såsom översättning, informationssökning, terminologi, lexikografi och språkutbildning. Forskningen inom CTT har skapat ett produktivt kunskapsflöde mellan akademi och industri. Enligt en nyutkommen studie: "Benchmarking HLT progress in Europe" (HLT = Human Language Technologies) framtagen i ett projekt inom EU:s IST-program, har Sverige "the highest HLT Opportunity score in the EU, due to its highly competitive knowledge society infrastructure". Att utnyttja denna möjlighet kräver en fortsatt industriell och akademisk satsning.

Språkbanker, i första hand för svenska, men även för andra för Sverige viktiga språk, som är allmänt tillgängliga. En språkbank är en stor och växande samling av språkliga material, både text och tal, i ren och i bearbetad form som exempelvis lexikon, termbanker, trädbanker, språkligt uppmärka tal-, dialog- och textdata.

Ett språkteknologiskt sekretariat med en samordnande funktion som också kan agera för svensk språkteknologi i en nordisk och europeisk kontext.

Utformningen av en långsiktig svensk strategi för språkteknologi bör, menar vi, uppmärksammas i den forskningspolitiska propositionen liksom i den proposition som ska följa på utredningen Mål i mun. De åtgärder som nu framstår som de mest angelägna är följande:

- Väsentliga långsiktiga förstärkningar till forskning, forskarutbildning och utveckling inom språkteknologi, häri inräknat fortsatta och ökade infrastrukturella satsningar på språkbanker.
- Skapandet av ett nationellt samordnande organ i form av ett språkteknologiskt sekretariat.