

En rekommenderad svensk språkteknologisk terminologi

Viggo Kann^α Lars Ahrenberg^β Rickard Domeij^γ Ola Karlsson^γ Jussi Karlgren^δ
Henrik Nilsson^ε Joakim Nivre^ζ

^αKTH ^βLinköpings universitet ^γSpråkrådet ^δGavagai ^εTerminologisentrum TNC ^ζUppsala universitet
viggo@kth.se, lars.ahrenberg@liu.se,
rickard.domeij@sprakochfolkminnen.se, ola.karlsson@sprakochfolkminnen.se,
jussi.karlgren@gavagai.se, henrik.nilsson@tnc.se, joakim.nivre@lingfil.uu.se

Abstract

In 2014 the Swedish Language Technology Terminology Group was created, with representatives from different parts of the language technology community, both higher education and research, industry and governmental agencies. In 2016 we have recommended Swedish terms for the 270 language technological concepts in the Bank of Finnish Terminology in Arts and Sciences. The language technology terms are published on folkets-lexikon.csc.kth.se/LTterminology, where anyone can lookup Swedish and English terms interactively and read the full list of terms. We also try to enter the most important Swedish terminology into the Swedish Wikipedia. We encourage use of these Swedish terms and welcome suggestions for improvements of the Swedish terminology.

1. Bakgrund

Att en väl definierad och gemensam terminologi är viktigt i forskning och utbildning är väl känt. Även om engelska används mycket i den språkteknologiska forskningen och utbildningen i Sverige är det viktigt att det också finns svenska termer. Det finns flera skäl till detta:

- Den svenska språklagen (Språklag, 2009) säger att svenska är huvudspråk i Sverige och ska kunna användas inom alla samhällsområden. Svenska folket har rätt att hålla sig informerade
- Svenska vetenskapsmän använder ofta svenska när de kommunicerar med varandra, även om publicering sker på engelska (Salö, 2016).
- Svenska studenter är ofta aktivare och får ut mer av undervisningen om den är på svenska än på engelska (Airey, 2009).
- Bättre begreppsförståelse kan nås om både svenska och engelska termer ges för samma begrepp i undervisningen (se figur 1).

En gemensam svensk terminologi för språkteknologiområdet har saknats. Universitet (som är myndigheter) har enligt 12 § i den svenska språklagen (Språklag, 2009) ett tydligt terminologiansvar: *Myndigheter har ett särskilt ansvar för att svensk terminologi inom deras olika fackområden finns tillgänglig, används och utvecklas.*

Inom ett par andra ämnesområden finns sedan flera år svenska termgrupper, se tabell 1. Kimmo Koskeniemi har lagt in engelsk och finsk språkteknologisk terminologi i den webbaserade Vetenskapstermbanken i Finland (Tieteen termipankki, tieteen.termipankki.fi), sammanlagt 270 begrepp.

2. Metod

Idén att bilda en svensk språkteknologitermgroup väcktes vid NoDaLiDa 2013. Gruppen bildades under 2014 och

Svenska biotermgruppen

www.biotermgruppen.se

Svenska datatermgruppen

www.datatermgruppen.se

Svenska optiktermgruppen

www.photonicsweden.com/node/35

Tabell 1: Svenska termgrupper

vara självförklarande
inte vara onödigt komplicerade
likna engelska termen
kunna böjas naturligt på svenska
vara i bruk redan

Tabell 2: Kriterier vid val av svenska termer

började sitt arbete i början av 2015. Författarna till detta bidrag utgör Svenska språkteknologitermgruppen, som alltså har representanter från tre svenska universitet med forskning och undervisning i språkteknologi, Terminologisentrum TNC, Språkrådet i Sverige och språkteknologiindustrin (representerad av Gavagai). Viggo Kann fungerar som sammankallande.

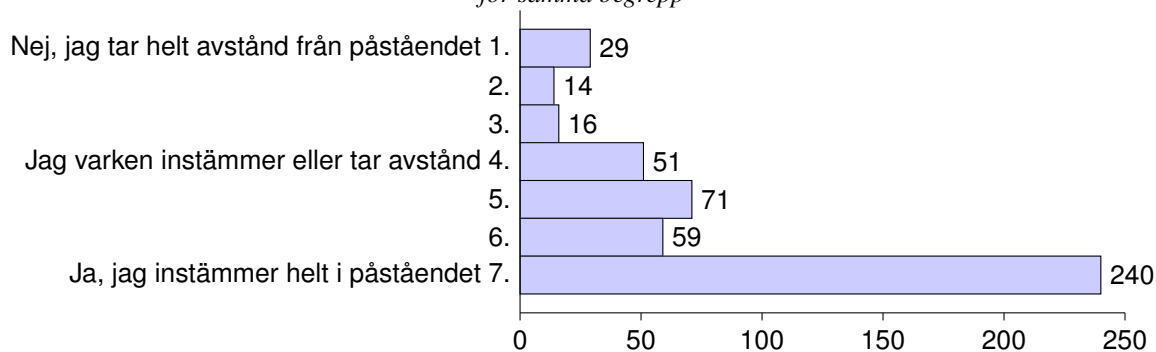
Gruppen har valt att börja med de 270 språkteknologiska begrepp som finns i Vetenskapstermbanken i Finland och även använda det systemet för det interna arbetet. För många begrepp har vi kunnat lägga in existerande vedertagna svenska termer. För andra har vi valt den bästa av existerande termer eller konstruerat nya termer. I några fall har vi lagt in flera svenska termer, då det inte varit tydligt att en viss term är lämpligast.

När vi har valt mellan alternativa svenska termer har vi använt kriterierna i tabell 2. Ofta har vi naturligtvis behövt väga kriterierna mot varandra och ge avkall på något eller några av kriterierna.

engelsk term	svensk rekommendation
alignment	länkning, linjering
computational linguistics	datalingvistik, datorlingvistik
named entity	namngiven entitet
outlier	utliggare
parsing	parsning
part of speech tagging	ordklasstagning, ordklassigenkänning
pushdown automaton	begränsad stackautomat
question answering system	frågebesvarande system
stemming	trunkering, stemming
tokenize, tokenise	ordsegmentera
word token	ordförekomst

Tabell 3: Exempel på terminologi

Påstående: Jag anser att det kan vara värdefullt för begreppsförståelsen att få lära sig både svenska och engelska termer för samma begrepp



Figur 1: Fråga till alla studenter i årskurs 1-3 på civilingenjörsprogrammet i datateknik på KTH våren 2016, $N = 480$.

3. Resultat

Exempel på termerna kan ses i tabell 3. Vi tillgängliggör nu våra termer på webben på adressen folkets-lexikon.csc.kth.se/LTterminology där termerna dels kan slås upp interaktivt och dels laddas ner i sin helhet.

För att ytterligare sprida de svenska termerna har vi börjat lägga in dem i Wikipedia, se exempelvis uppslagsordet *ordklasstagning*.

4. Diskussion

Vi hoppas att svenska språkteknologer ska använda de rekommenderade termerna. Dessutom rekommenderar vi att både svenska och engelska termer för språkteknologiska begrepp förmedlas till studenterna i språkteknologikurser. Ett enkelt sätt att göra detta är att för varje kurs tillhandahålla en svensk-engelsk ordlista över de viktigaste begreppen i kursen.

Arbetet med den språkteknologiska terminologin är långt ifrån färdigt. Det finns många begrepp kvar att täcka in och vissa rekommenderade svenska termer kan behöva diskuteras och revideras.

Vi uppmuntrar fler att engagera sig i terminologiarbetet och tar gärna emot förslag på utökningar och förbättringar av termrekommendationerna genom att ta kontakt med någon av gruppens medlemmar.

Bibliografi

- J. Airey. 2009. *Science, Language, and Literacy: Case Studies of Learning in Swedish University Physics*. Doktorsavhandling, Uppsala universitet.
- L. Salö. 2016. *Languages and Linguistic Exchanges in Swedish Academia*. Doktorsavhandling, Stockholms universitet.
- Språklag (2009:600). *Svensk författningssamling*