

Design af et sprogteknologisk værktøj for leksikografer

Christina Dideriksen

Dansk Sprognævn

Det Centrale Ordregister (COR) for dansk er et langsigtet projekt som lægger fundamentet til fremtidige udviklinger inden for sprogteknologien. Udviklingen af sprogværktøjer på dansk har indtil nu været afhængig af små og specifikke initiativer, men det nuværende, større projekt vil muliggøre en forenet indsats hvor alle sprogteknologiprojekter refererer tilbage til det samme sprogregister. Med COR klargør vi en fælles ramme for hele ordforrådet, hvor alle lemmaer får tildelt et indeks som kan bruges på tværs af opslagsværk. Dermed er det muligt at koble opslag fra flere forskellige ordbøger og leksika til det samme lemma og samle alt information om det samme ord under ét indeksnummer.

Det er dog problematisk hvis vi forventer at slutbrugerne alle er computerprogrammører. Målgruppen består snarere af fagpersoner med vidt forskellige erfaringer med programmering. Ud over en fælles interesse i at bygge sprogværktøjer vil nogle være erfarne programmører med meget lidt viden om lingvistiske strukturer, mens andre vil være erfarne leksikografer med kun sparsom erfaring med programmering. Hvordan kommunikerer vi bedst mulighederne med COR til en forskelligartet brugergruppe? Hvordan engagerer vi brugere fra forskellige fagretninger, firmaer, udviklere og forskere? Med andre ord, hvordan sikrer vi at manualen til COR ikke bliver så teknisk at mange potentielle brugere står af? I denne posterpræsentation gennemgår jeg de kommunikative beslutningsprocesser, samt præsenterer de designvalg vi har taget i udviklingen af en manual som skal ramme bredt. Projektet har relevans for forskere og udviklere som gerne vil udvikle en lignende resurse eksempelvis på norsk eller svensk.

Keywords: Sprogteknologi; sprogprogrammering; dansk