

## **Automatisk termekstraktion: En ny udfordring inden for terminologiarbejde i Island**

*Ágústa Þorbergsdóttir, Einar Freyr Sigurðsson & Steinþór Steingrímsson*

*Árni Magnússon-instituttet i islandske studier*

Vi præsenterer TermPortal, et arbejdsbord til terminologiarbejde, med det første ATE-system til islandsk. Værktøjet letter standardisering af terminologiarbejdet, da det gør det muligt at eksportere data i standardformater for at tilpasse indsamling og distribution af materialet. Vores udfordring er at prøve automatisk termekstraktion (ATE) fra elektroniske tekster for mere effektivt terminologiarbejde.

Indtil nu er terminologidatabaser i Island blevet opbygget manuelt af eksperter inden for deres fagområder. Termbaser inden for mange forskellige fagområder er blevet oprettet og gjort tilgængelige online i Íðorðabankinn. Der er flere ulemper ved en manuel indsamling af terminologi. Metoden er meget tidskrævende. Nogle af de indsamlede termer bliver muligvis ikke udbredt, før de bliver erstattet af nyere eller bedre kendte. De ældre termer bliver dog ikke slettet fra termbasen, hvilket øger risikoen for tvetydighed.

I vores præsentation fokuserer vi på økonomitermer for at kunne opfylde behov for et vigtigt og stort felt. Vi præsenterer en omfattende undersøgelse blandt de mest fremtrædende organisationer på området, hvis resultater understreger behovet for en god, opdateret og tilgængelig termbank og villigheden til at bruge termer på islandsk. Ydermere præsenterer vi ATE-værktøjet for islandsk, som anvender en række forskellige metoder og viser et stort potentiale med en tilbagekaldelsesrate på op til 95% og en høj C-værdi. Resultatet indikerer, at værktøjet er god til at finde de vigtigste termkandidater for inputteksten.

### Referencer

Ágústa Þorbergsdóttir. 2021. Staða íslensku á sviði fjármála. *Orð og tunga* 23:111–

136. <https://doi.org/10.33112/ordogtungu.23.5>

Íðorðabankinn, [www.arnastofnun.idord.is](http://www.arnastofnun.idord.is)

Nøgleord: termer; termekstraktion; korpus; domæner; islandsk