



Nefndasvið Alþingis  
b.t. allsherjar- og menntamálanefndar  
Austurstræti 8-10  
150 Reykjavík

Reykjavík, 25. mars 2018

**Umsögn um 36. mál á 148. löggjafarþingi,  
frumvarp um breytingu á höfundalögum nr. 73/1972**

Ég undirritaður er framkvæmdastjóri Miðeindar ehf. sem unnið hefur sl. tæp þrjú ár að gerð hugbúnaðar og veflausna á sviði máltækni fyrir íslensku. Meðal annars heldur fyrirtækið úti vefnum **greynir.is** þar sem unnt er að málgreina (þátta) íslenskan texta, skoða þáttaðar málgreinar myndrænt, svara fyrirspurnum úr textagrunni o.m.fl. Fyrirtækið hefur á grundvelli umsókna hlotið þrjá styrki úr Máltæknisjóði, árin 2016 og 2017, til verkefna sinna. Allur útgefinn hugbúnaður Miðeindar ehf. er opinn og aðgengilegur á netinu undir GNU GPLv3 hugbúnaðarleyfi (e. *free open-source software*).

Fyrirtækið vinnur nú m.a. að verkefni þar sem djúpt tauganet (e. *deep neural network*) er þjálfað til að málgreina íslenskan texta. Vonir standa til að það verkefni muni einnig nýtast til að bæta sjálfvirka þýðingu úr íslensku yfir á önnur tungumál með djúpum tauganetum. Djúp tauganet hafa rutt sér mjög til rúms á sviði máltækni á undanförunum árum og er því spáð að miklar áframhaldandi framfarir geti orðið á grundvelli þeirra.

Djúp tauganet byggja á þjálfun með miklu magni gagna. Vandí íslenskunnar í því sambandi er sá að framboð af tölvutækum textum og öðrum gögnum til slíkra nota er takmarkað í samanburði við tungumál sem fleiri nota. Íslendingar þurfa því að hafa allar klær úti við að afla texta og gagna til nota í rannsóknum og hagnýtum verkefnum á sviði máltækni, ef möguleikar eiga að vera á að standa nokkurn veginn jafnfætis öðrum tungum í þeirri byltingu sem er rétt að hefjast á því sviði og tengdum sviðum málskilnings og gervigreindar.

Í því sambandi má benda á að undanþágur varðandi texta- og gagnavinnslu í höfundarréttar-regluverki sem nú er í undirbúningi á vegum Evrópusambandsins ná einungis til rannsóknastofnana. Að mínu mati, með hliðsjón af íslenskum aðstæðum og þá einkum smæðar markaðarins, þyrftu slíkar undanþágur einnig að ná til verkefna sem ætluð eru til hagnýtingar í atvinnuskyni eða eru a.m.k. unnin á vegum einkaaðila, eins og gert er ráð fyrir í því frumvarpi sem hér er til umræðu.

Ljóst er að ef átak í máltækni fyrir íslensku á að ná tilgangi sínum þurfa afurðir þess að vera nýtanlegar fyrir almenning í raunverulegum vörum og þjónustu, svo sem í farsímum, raddstýringu tækja og bíla, í fyrirspurnakerfum á vef o.s.frv. Slíkar lausnir verða (vonandi) á

endanum þróaðar á vegum einkaaðila, innlendra og erlendra, eftir atvikum í samstarfi við rannsóknastofnanir. En til þess þarf sem hindrunarlausastan aðgang að eins stórum gagnasöfnum og mögulegt er, í þágu þess markmiðs sem að er stefnt.

Ekki verður séð að umrætt frumvarp breyti neinu um réttarstöðu höfundarréttarhafa hvað varðar birtingu upphaflegs höfundarverks, sem verður áfram undir almennum skilmálum laganna og vernduð af þeim. Hins vegar ætti að geta verið beinn og óbeinn hagur íslenskra höfundarréttarhafa að stuðla að framgangi íslenskrar máltækni og viðhaldi og eflingu tungumálsins að því leyti.

Ég undirritaður mæli eindregið með því að fyrirliggjandi frumvarp verði samþykkt, enda yrði það að mínu mati framsýnt heillaspor fyrir íslenska máltækni og framtíð íslenskunnar.

*F.h. Miðeindar ehf.*

*Vilhjálmur Þorsteinsson,  
framkvæmdastjóri*



