

148. löggjafarþing 2017–2018.

Þingskjal 306 — 219. mál.

Tillaga til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensks prentmáls.⁴

Takk fyrir að senda mér þetta til yfirlestrar, ég tel að höfundar hljóti að hafa alfarið um þetta að segja, þeir eiga þetta efni, ekki allir Íslendingar. Þegar og ef þeir gefa sitt leyfi þá er tímabært að fara að vinna að útbreiðslu fyrir hönd íslensku þjóðarinnar.

Með kveðju, Heiðrún D. Eyvindardóttir
forstöðumaður Bókasafns Árborgar
heidrun@arborg.is www.bokasofn.arborg.is

Nefndasvið Alþingis,
Allsherjar og menntamálanefnd,
b.t. Sigrún Helga Sigurjónsdóttir,
Austurstræti 8 – 10,
101 Reykjavík.

netfang. nefndasvid@althingi.is

Efni. Tillaga til þingsályktunar mál 219/2018

Vísað er til tölvupósts Allsherjar og menntamálanefndar Alþingis sem sent var umbjóðanda mínum Félagi Íslenskra bókaútgefenda þann 06.03.2018. Óskað var umsagnar félagsins um tillögu til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensk prentmáls. Hefur undirrituðum lögmanni félagsins verið falið að svara Allsherjar og menntamálanefnd fyrir hönd félagsins.

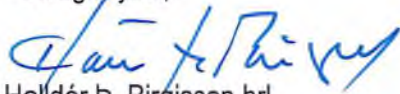
Fyrir liggur að Félag Íslenskra bókaútgefenda er jákvætt í málinu og styður hugmyndir um skipun nefndarinnar.

Félagið ítrekar sjónarmið sem það hefur áður haft í frammi opinberlega um mögulega stafræna endurgerð útgefinna verka eða að verkefnið sjálft og skipulagning þess verði í samráði við réttihafa þar með talið útgefendur og ekki verði farið af stað með verkefnið án slíks samráðs. Gera þurfi samninga áður en skönnun hefst, sem nái í senn til skönnunar efnis og mögulegar miðlunar þess og möguleika réttihafa til að koma í veg fyrir miðlun innskannaðs efnis sem enn er aðgengilegt og eða til sölu á almennum markaði. Leggur félagið því áherslu á að samningum við réttihafa um efnið og meðferð þess verði lokið áður en verkefnið hefst.

Um samninga um skönnun og dreifingu sem gætu verið leiðbeinandi má meðal annars benda til samninga Norska landsbókasafnsins við þarlenda réttihafa sem voru gerðir áður en til skönnunar prentaðs máls kom til í Noregi en þeirri skönnun er að verða lokið og hluti þess efnis sem skannaður hefur verið hefur nú þegar verið gerður aðgengilegur almenningi þar í landi.

Reykjavík, 26.03.2018.

Virðingarfyllst,



Halldór P. Birgisson hrl.



Hljóðbókasafn Íslands

Bókasafn fyrir blinda, sjónskerta
og prentleturshamlaða

Hljóðbókasafn Íslands, Kópavogi 8. mars 2018

148. löggjafarþing 2017-2018

Þingskjal 306 – 219. mál

Tillaga til þingsályktunar um gerð áætlunar un stafræna endurgerð íslensk prentmáls.

Hljóðbókasafn Íslands fékk til skoðunar tillögu til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensk prentmáls. Tillagan er tímabær og er henni fagnað innilega. Eins og tekið er fram er með stafvæðingunni unnt að stórauka aðgengi íslensks almennings að þjóðararfinum og eru Norðmenn til mikillar fyrirmyndar í þessu efni.

Evrópusambandið hefur lagt fram sérstök tilmæli um hvernig skuli staðið að slíkri stafvæðingu, hvernig skuli yfirfæra og varðveita efni. Hljóðbókasafn Íslands veitir prentleturshömluðum aðgang að prentuðu máli sem annars er óaðgengilegt fyrir þennan hóp, í formi hljóðbóka og stafrænni endurgerð á annars lokuðum texta.

Undirrituð vill ítreka sérstaklega að farið verði eftir tilmælum Evrópusambandsins í þaua hvað varðar aðgengi að stafrænu efni og að tryggt verði að stafvæðingin ná til allra þjóðfélagsþegna, að tryggt verði að prentleturshamlaðir einstaklingar, blindir, sjónskertir, lesblindir og fleiri sem eiga við slíka hömlun að stríða hafi jafnan aðgang að efninu og aðrir.

Gild ástæða er fyrir því að taka þetta sérstaklega fram því skort hefur meðvitund um þessi máli í íslensku samfélagi þegar kemur að stafvæðingu efnis. Sem dæmi má nefna að vefurinn timarit.is sem er sannarlega mikil góð upplýsingaveita er alveg lokaður fyrir blinda og aðra textahamlaða.

Efnið þar hefur verið skannað inn og hægt er að fara inn í textann en það vantar alla uppsetningu (navigation) svo ekki er hægt að lesa textann í réttri röð. Þetta þýðir að efnið er ónothæft fyrir prentleturshamlaða. Svona vinnubrögð þýða einfaldlega að endurgera þarf efni af vefnum, sem blindir og sjónskertir óska eftir að fá aðgang að, sem er mikil sóun á tíma og peningum hins opinbera.

Stafræn endurgerð á texta með ásættanlegu aðgengi fyrir blinda krefst ekki mikillar aukavinnu og til eru leiðbeiningar og sjálfvirk tól til þess að tryggja að stafrænn texti verði aðgengilegur fyrir alla. Efalaust felast slíkar leiðbeiningar í tilmælum Evrópusambandsins um stafvæðingu. Alþjóðlegu aðgengissamtökin DAISY hafa það að markmiði að allt efni verði gert „born accessible“ (aðgengilegt frá upphafi) og vonin er sú að með tímanum verði sérstök bókasöfn fyrir prentleturshamlaða óþörf en ráðgjafar er þörf. Samkvæmt rannsóknum glímir að lágmarki 5% þjóðar við prentleturshömlun af einhverju tagi, en talið er að hlutfallið sé nær 10%.

Því vil ég leggja þunga áherslu á að við skipun starfshóps um stafvæðingu verði tilnefndur aðili sem hefur hagsmuni prentleturshamlaðra að leiðarljósi svo koma megi í veg fyrir að þúsundir íslendinga komi að lokuðum dyrum þegar þessu mikilvæga verkefni verður hleypt af stokkunum.

Með virðingu og fyrirfram þökkum,

Þóra Sigríður Ingólfsdóttir, forstöðumaður Hljóðbókasafns Íslands.



LANDSBÓKASAFN ÍSLANDS – HÁSKÓLABÓKASAFN

Allsherjar- og menntamálanefnd Alþingis
Austurstræti 8–10
101 Reykjavík

Reykjavík, 22. mars 2018
Tilvísun: 201803-0005 / ISS/öb/ks

Efni: Umsögn um tillögu til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensks prentmáls – 148. löggjafarþing 2017–2018. Þingskjal 306 – 219. mál

Tekið er undir markmið tillögunnar, að gerð verði kostnaðar- og tímaáætlun um stafvæðingu íslensks prentmáls og að kannaðir verði möguleikar á að ná samningum við rétthafa um afnot þessara ritverka.

Verkið felur í sér tvo stóra kostnaðarþætti sem hægt er að meta og tímasetja.

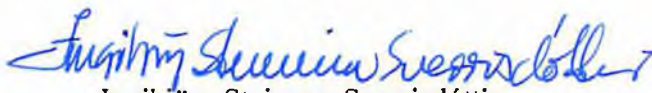
Annars vegar er það kostnaður við stafræna endurgerð, s.s. við starfsmannahald og tækjabúnað en hann er mishár eftir þeim tíma sem er ætlaður í verkið, auk kostnaðar við langtímavarðveislu þeirra stafrænu gagna sem verða til.

Hins vegar eru greiðslur til rétthafa eða samtaka þeirra sem þarf að semja um. Í því samhengi er bent á mikilvægi þess að stafvædda efnið verði opið öllum og að hægt verði að nýta textana til rannsókna og eflingar íslenskri tungu.

Reynslan sýnir að ávinningurinn af slíkum verkefnum er stórbætt aðgengi alls almennings að menningararfinum þegar hverjum og einum hentar og að með þessu móti má vekja efni sem ekki er á markaði og jafnvel vandfundið á bókasöfnum.

Landsbókasafn Íslands – Háskólabókasafn er reiðubúið að taka þátt í verkefni sem þessu og tilnefna fulltrúa í starfshóp.

Virðingarfyllst,


Ingibjörg Steinunn Sverrisdóttir
landsbókavörður

ARNGRÍMSGÖTU 3 · 107 REYKJAVÍK · SÍMI: 525 5600 · BRÉFASÍMI: 525 5615

NETFANG: upplys@landsbokasafn.is VEFFANG: www.landsbokasafn.is

Umsögn

Mennta- og menningarmálaráðuneyti, 4. apríl 2018.

219. mál:

Tillaga til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensks prentmáls.

Allsherjar- og menntamálanefnd Alþingis sendi mennta- og menningarmálaráðuneyti með tölvupósti 6. mars sl. til umsagnar tillögu til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensks prentmáls, 219. mál. Var óskað eftir að í umsögninni kæmi jafnframt fram hugsanlegt mat á kostnaði. – Loks var þess óskað að undirrituð umsögn bærist eigi síðar en 26. mars nk.

Umrædd tillaga hljóðar svo:

„Alþingi ályktar að fela mennta- og menningarmálaráðherra að skipa starfshóp sem geri kostnaðar- og tímaáætlun um stafræna endurgerð íslensks prentmáls sem verði aðgengileg til lestrar á nettengdum búnaði og kanni möguleika þess að ná samningum við höfundaréttarhafa um slík afnot stafrænnar endurgerðar ritverka. Starfshópurinn geri grein fyrir kostnaði og möguleikum á að ljúka stafrænni endurgerð íslensks prentmáls á a) 5 árum, b) 10 árum og c) 15 árum. Starfshópurinn ljúki störfum í tæka tíð til þess að ráðherra verði unnt að flytja Alþingi skýrslu um málið og kynna niðurstöður hans fyrir 1. desember 2018.“

Segja má að þegar hafi verið unninn á vegum mennta- og menningarmálaráðuneytis umtalsverður hluti þess verkefnis, sem fyrirhuguðum starfshópi er ætlað að sinna, og er lítillega vísað til þeirrar vinnu í greinargerð þingsályktunartillögunnar. Þessi vinna skiptist í meginatriðum í tvo hluta:

A.

Á grundvelli ákvæðis til bráðabirgða sem sett var í lög um bókmenntir nr. 91/2007 og tók gildi 1. janúar 2013 skipaði þáv. mennta- og menningarmálaráðherra samráðsnefnd, sem skyldi skila ráðherra skýrslu um framtíð íslenskrar bókaútgáfu. Þar skyldi m.a. leggja fram tillögur og marka skýra stefnu um aukna rafbókavæðingu á Íslandi og stafrænan aðgang almennings að höfundarréttarvörðu efni. Nefndin skilaði skýrslu sinni 10. janúar 2014, og meðal tillagna hennar var að höfundaréttvarinn prentarfur, sem ekki væri á markaði, yrði aðgengilegur í gegnum vefgáttina bækur.is, þ.e. að efnið yrði ljósmyndað og birt á vefgáttinni, sem Landsbókasafn Íslands – Háskólabókasafn hefur umsjón með. Samkomulag um greiðslur yrði gert af Fjölís fyrir hönd höfundaréttshafa við Landsbókasafn Íslands - Háskólabókasafn.

Mennta- og menningarmálaráðuneyti óskaði í kjölfar skýrslunnar eftir að Landsbókasafn Íslands – Háskólabókasafn gerði drög að áætlun sem gæfi til kynna umfang og kostnað við það verkefni sem samráðsnefndin lagði til að yrði unnið við að gera prentarfinn aðgengilegan í stafrænu formi. Í svari safnsins kom fram að á tímabilinu 1850 – 2009 hafi verið gefnar út rúmlega 78.300 bækur (íslenskar bækur sem gefnar voru út fyrir 1850 eru um 2.600 talsins, en árið 2014 var búið að færa um 70% þeirra á stafrænt form og fyrir lágu áætlanir um að færa þau 30% sem eftir voru á stafrænt form. Því voru bækur sem gefnar voru út fyrir 1850 ekki taldar með í þeim tölum sem hér um ræðir). Ef miðað væri við ákveðið meðaltal blaðsíðufjölda hvernar bókar (105 bls.) væri um að ræða rúmlega 8,2

milljónir blaðsíðna. Safnið gaf sér síðan ákveðnar forsendur um afköst tækjabúnaðar við skönnun (300 bls. á klst. fyrir hverja myndavél), nauðsynlegan mannafla (28 ársverk), tækjakaup, varðveislukostnað í 10 ár o.fl. og taldi að miðað við þær forsendur mætti vænta þess að kostnaður við verkefnið yrði um 196 m.kr. á verðlagi ársins 2014.

Landsbókasafn Íslands – Háskólabókasafn benti í svari sínu til ráðuneytisins á að rétt væri að miða við að bækur sem gefnar hafa verið út eftir 1930 nytu enn höfundaréttar, en um 70% af þeim bókum sem gefnar hafa verið út á Íslandi falla í þann hóp. Í samningum sem samtök höfundaréttarhafa í Noregi, Kopinor, hafa gert við norska þjóðarbókasafnið um stafræna endurgerð prentaðra bóka á norsku sem hafa verið gerðar aðgengilegar almenningi (og er vísað til í greinargerð fyrirbyggjandi tillögu til þingsályktunar) er gert ráð fyrir að safnið greiði árlega 33 – 36 norska aura, eða u.þ.b. 4,2 – 4,6 íslenskar krónur fyrir hverja blaðsíðu.

- Með sambærilegum samningum hér á landi mundu árlegar greiðslur vegna höfundaréttar af 70% ofangreinds blaðsíðufjölda því nema um 24,1 – 26,4 m.kr.

Rétt er að hafa fyrirvara á ofangreindum áætlunum, t.d. hvað varðar meðalblaðsíðufjölda bóka, afköst við skönnun, hlutfall bóka sem enn njóta höfundaréttar o.s.frv., en þessi drög að áætlun á þessu sviði ættu engu að síður að gefa nokkra hugmynd um þann kostnað sem fælist í verkefninu.

B.

Alþingi samþykkti 16. maí 2014 þingsályktun (nr. 36/143) um „að fela mennta- og menningarmálaráðherra að setja fram markvissa, heildstæða og metnaðarfulla stefnu um varðveislu íslenskrar menningararfleifðar á stafrænu formi og leggja drög að því að verkið verði unnið á næstu 10–20 árum eftir því hversu umfangsmikið það verður talið“.

Í greinargerð þeirri tillögu til þingsályktunar sem Alþingi samþykkti með þessum hætti sagði m.a. að markmið hennar væri annars vegar að koma í veg fyrir að menningarverðmæti glötuðust eða eyðilegðust og hins vegar að auðvelda fólki aðgang að menningarverðmætum sem væri að finna á stafrænu formi. Þar sagði einnig að til að takast á við þetta viðfangsefni þyrfti að móta markvissa, heildstæða og metnaðarfulla stefnu sem tryggði að bókmenntaverk, tónverk, myndverk og önnur menningarverðmæti fyrri alda, áratuga og ára verði alltaf tiltæk almenningi á stafrænu formi. Var bent á að slíkt menningarverkefni yrði umfangsmikið að vöxtum og krefðist verulegra fjármuna, en með því að setja raunhæf markmið til nokkurra ára ætti verkefnið að geta orðið fjárhagslega framkvæmanlegt. Minnt var á að í náinni framtíð væru tvö stórafmæli á Íslandi, þ.e. 100 ára afmæli fullveldis 1. desember 2018 og 1100 ára afmæli Alþingis 2030, og af því tilefni gæti þjóðargjöfin verið sú að hefjast handa við þetta verkefni sem fyrst með það í huga að því yrði lokið árið 2030.

Eins og kom fram í greinargerð tillögunnar voru bókmenntaverk – íslenskt prentmál – meðal þeirrar menningararfleifðar sem ályktunin nær til, og hefur verið unnið að verkefninu á þeim grundvelli.

Vinna við undirbúning stefnumótunar ráðherra á grundvelli ályktunarinnar hófst 2015, og fólst fyrsti áfangi þeirrar vinnu í að kortleggja stöðu stafrænnar endurgerðar íslenskrar menningararfleifðar hjá söfnum og menningarstofnunum landsins. Var spurningakönnun um stöðu varðveislu menningarefnis á stafrænu formi send til 105 safna og

menningarstofnana sem heyra m.a. undir lög um menningarminjar og lög um opinber skjalasöfn sem og til Landsbókasafns Íslands – Háskólabókasafns og annarra safna og stofnana sem geyma menningararfleifð þjóðarinnar. Alls svöruðu 75 þeirra sem leitað var til öllum þáttum könnunarinnar, en 5 svöruðu könnuninni að hluta.

Þessi könnun og úrvinnsla gagna sem bárust, sem og rannsóknir á ýmsum tæknipáttum verkefnis af þessu tagi tóku lengri tíma en vænst var í upphafi. Í mars 2017 var voru niðurstöður þessa fyrsta áfanga gefnar út í skýrsluformi, og þannig skapaður grundvöllur fyrir stjórnvöld til að taka ákvarðanir um næstu skref með vísan til þeirra upplýsingar sem aflað hefur verið.

Frá því skýrslan kom út hefur verið unnið í mennta- og menningarmálaráðuneyti að mótun heildstæðrar og metnaðarfullrar stefnu á þessu sviði sem ráðherra geti lagt fyrir Alþingi með vísan til ofangreindrar þingsályktunar. Er þess vænst að ráðherra geti lagt tillögu þessa efnis fyrir Alþingi á haustþingi 2018.

Það er álit mennta- og menningarmálaráðuneytis að fyrirbyggjandi tillaga til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensks prentmáls auki ekki við þá áherslu sem felst í þingsályktun nr. 36/143 um að mennta- og menningarmálaráðherra setji fram markvissa, heildstæða og metnaðarfulla stefnu um varðveislu íslenskrar menningararfleifðar á stafrænu formi og leggi drög að því að verkið verði unnið tilteknum árafjölda.

Verði fyrirbyggjandi tillaga til þingsályktunar samþykkt kann hún að skapa íslensku prentmáli nokkra sérstöðu og jafnvel forgang á aðra þætti íslenskrar menningararfleifðar þegar kemur að gerð áætlana um stafræna endurgerð og ákvörðunar um fjárveitingar til slíkra verkefna.

- Um leið þyrfti mennta- og menningarmálaráðuneyti að endurskoða viðeigandi hluta þeirrar stefnumótunar um varðveislu íslenskrar menningararfleifðar á stafrænu formi sem nú er unnið að á grundvelli þingsályktunar nr. 36/143.

Allsherjar- og menntamálanefnd Alþingis
Austurstræti 8 – 10
150 Reykjavík

Efni: Umsögn um þingsályktunartillögu um gerð áætlunar um stafræna endurgerð íslensks prentmáls, 148. löggjafarþing, þskj 306, 219. mál

Undirrituð fyrir hönd Rithöfundasambands Íslands hefur kynnt sér efni þingsályktunartillögu um gerð áætlunar um stafræna endurgerð íslensks prentmáls, 148. löggjafarþing, þskj 306, 219. mál og óskar að koma á framfæri eftirfarandi athugasemdum við tillöguna.

Tillagan lýtur ekki að efnislegum breytingum á íslenskum lögum en gengur út á að gerð verði kostnaðar- og tímaáætlun um tiltekna stafræna endurgerð íslensks prentmáls sem verði aðgengileg til lestrar á nettengdum búnaði og kanni möguleika þess að ná samningum við höfundaréttarhafa um slík afnot. Í sjálfu sér er jákvætt að gerðar séu kostnaðar- og tímaáætlanir og hefur undirrituð engar athugasemdir við slíkt. Hins vegar verður ekki fram hjá því litið að tillagan af afskaplega rúmt orðuð og þar með nokkuð óljóst hvað um sé að ræða, í fleiri skilningi en einum.

Sem dæmi má nefna að heppilegra væri að a.m.k. í greinargerð væri að finna útskýringu eða skilgreiningu á því hvað felist í stafvæðingu íslensks prentmáls, hvernig slíkar skrár verði gerðar og/eða varðveittar. Það era ð auki óljóst hvað átt er við með tilvísun greinargerðar til þess að bók hafi „náð þeim aldri að [hún sé] ekki lengur söluvara á markaði“. Hvenær bók hefur náð þeim aldri er mismunandi eftir því hvað miðað er við. Er hér miðað við heildarsölu á gildandi útgáfusamningstímabili? Er miðað við árssölu? Er miðað við forsendur ákvörðunar um hvort farga megi lager með vísan til útgáfusamnings aðila? Er miðað við hversu langur tími þarf að líða án þess að nokkuð eintak bókarinnar seljist? Eða er einfaldlega miðað við tiltekinn árafjölda frá fyrstu útgáfu? Þetta þarf augljóslega að skilgreina betur.

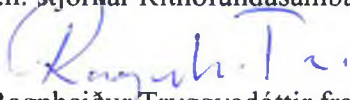
Í greinargerð eru t.a.m. sérstaklega nefndar til sögunnar stofnanir og embætti í Evrópu og á Norðurlöndum, sem hafið hafa yfirfærslu efnis á stafrænt form en betur færi ef skýrar væri kveðið á um það hvort markmið og tilgangur starfshópsins sé að líta sérstaklega til þeirra aðferða sem þeir aðilar nota eða hvort bera eigi þær aðferðir saman og jafnvel kostnaðargreina. Þetta er einnig óljóst.

Þá er starfshópnum skv. tillögunni ætla að kanna möguleika þess „að ná samningum við höfundaréttarhafa um *slík afnot* stafrænnar endurgerðar ritverka“ (skáletrun undirritaðrar). Hér er komið að kjarna athugasemda undirritaðrar. Þetta orðalag er alls ófullnægjandi

enda alveg á huldu hvað átt er við með „slík afnot“. Hér væri nauðsynlegt að skýra hugtakið og útlista nákvæmlega hvað átt sé við. Erfitt er að sjá fyrir sér að samtal við rétthafasamtök um stafræn afnot verka, beri nokkurn árangur þegar aðilum sem í slíku samtali eiga er í raun ekki ljóst hvað verið er að fjalla um. Það er grunnforsenda slíks samtals að nákvæmlega sé útskýrt hvað átt sé við með þeim afnotum sem vísað er til í tillögunni, til að tillagan megi ná markmiði sínu.

Virðingafyllst

f.h. stjórnar Rithöfundasambands Íslands


Ragnheiður Tryggvadóttir framkvæmdastjóri

Stofnun Árna Magnússonar í íslenskum fræðum
Reykjavík, 26. mars 2018

148. löggjafarþing 2017-2018

Umsögn um þingskjal 306 — 219. mál: Gerð áætlunar um stafræna endurgerð íslensks prentmáls.

Stofnun Árna Magnússonar í íslenskum fræðum fékk til skoðunar þingsályktunartillögu um gerð áætlunar um stafræna endurgerð íslensks prentmáls. Undirrituð hafa kynnt sér tillöguna.

Í greinargerð með þingsályktunartillögunni er fjallað um mikilvægi stafvæðingar prentefnis. Þar er vísað til ávinnings með tilliti til varðveislu efnisins, stóraukins aðgengis að því og því að með stafvæðingunni sé stutt við markmið um að efla stöðu íslensks máls í stafrænum heimi. Þá er vísað til sambærilegs verkefnis í Noregi, Bokhylla.no, sem lokið var við nú í upphafi árs eftir tæplega 10 ára vinnu. Í samningi norska landsbókasafnsins, Nasjonalbiblioteket, og samtaka rétthafa í Noregi, Kopinor, kemur markmið stafvæðingar efnisins skýrt fram. Efnioð skal vera aðgengilegt í nettengdum tækjum með norskum IP-tölur til lestrar á skjá í gegnum vefsíður norska landsbókasafnsins og það skal vera leitarbært í leitarvélum. Þar að auki, og þá kemur að þeim lið sem undirrituð láta sig helst varða, getur norska landsbókasafnið veitt aðgang að efninu með öðrum hætti í sérstökum tilgangi, til rannsókna eða menntunar, að uppfylltum ákveðnum skilyrðum.¹

Stór textasöfn eru afar mikilvæg verkfæri í mörgum greinum hugvísinda, í félagsvísindum og einnig í tæknigreinum þegar unnið er með tungumál, þ.e. í greinum sem tengjast máltækni. Við orðabókagerð er mikilvægt að hægt sé að finna sem fjölbreytilegust dæmi um orðnotkun og úr sem fjölbreytilegustum textum til að unnt sé að greina sem flest blæbrigði sem kunna að finnast í málinu, hvort sem það eru merkingarblæbrigði, breytileiki í setningargerð eða beygingum, yfirfærsla á notkunarviði orða eða orðmynda í gegnum tíðina eða önnur einkenni málsins sem rannsóknir í orðfræði láta sig varða. Á undanförunum árum hefur aðgengi að stafrænu efni aukist mjög sem hefur opnað nýja möguleika í rannsóknum á þessu sviði. Áþreifanlegur skortur á útgefnum bókum sem leitarbærar eru með þessum hætti, bæði á fræðiritum og skáldritum, getur þó skekkt myndina því í þeim er

¹ Sjá: <https://nffo.no/ViewFile.aspx?id=989>

texti sem oft er annars eðlis en það sem aðgengilegt er. Hann getur verið ígrundaðri, orðaforðinn getur verið annar og sjaldgæfari eða sértækari orðnotkun er stundum hvergi að finna nema þar. Stafræn endurgerð útgefna bóka myndi því styðja mikið við rannsóknir á þessu sviði.

Í máltækni eru stór textasöfn greind málfræðilega og notuð til að búa til tölfræðileg líkön af málinu. Þessi líkön eru notuð til að auka nákvæmni í talgervlum, þýðingarvélum og við ritvilluleit, svo eitthvað sé nefnt.² Hér gildir það sama og við orðfræðirannsóknirnar, því fleiri og fjölbreytilegri sem textarnir eru sem hægt er að nota í þessa þróun, því betri verður niðurstaðan. Textar á borð við þá sem stafvæddir yrðu í því verkefni sem lýst er í þingsályktunartillögunni myndu því koma að miklu gagni á fleiri sviðum við að efla stöðu íslensku í stafrænum heimi en eingöngu því að gera fjölbreytilegt prentefni aðgengilegt.

Undanfarin ár hefur stafrænum hugvísindum vaxið fiskur um hrygg, ekki síst í nágrannalöndum okkar. Stafræn hugvísindi er rannsóknarsvið sem notar tölfræðilegar aðferðir og önnur verkfæri tölvunarfræðinnar til að skoða viðfangsefni hugvísinda, svo sem bókmenntafræði, sögu eða tungumál. Meginstoð slíkra rannsókna er aðgangur að stafrænum gögnum, sem hægt er að vinna með og greina. Tímarit eru gefin út og ráðstefnur haldnar sem sérhæfa sig í stafrænum hugvísindum, meðal annars á Norðurlöndum en þriðja norræna ráðstefnan í stafrænum hugvísindum fór fram í Helsinki nú í mars (2018). Þar var meðal annars fjallað um rannsóknir á breytingu á merkingu orða eins og þau birtast í prentuðu efni á löngum tíma, leit að höfundum gamalla texta með stílmælingum, tölfræðilega greiningu á því hvernig rit sem teljast til ákveðinnar bókmenntastefnu greina sig frá öðrum ritum, þróun á umfjöllun um kyn í bókmenntum á 75 ára tímabili og flutninga Norðmanna til Vesturheims á níttjándu öld og í upphafi þeirrar tuttugustu. Allar þessar rannsóknir byggðu á miklu magni prentefnis sem hefur verið stafvætt, sumar á *Bokhylla.no*. Án stafvæðingarinnar væru þessir nýju rannsóknarmöguleikar ekki fyrir hendi.

Í félagsvísindum eru stór textasöfn notuð í auknum mæli til að greina samfélagið og samfélagsbreytingar. Þar má nefna rannsóknir á fátækt, hatursorðræðu og viðhorfum til loftslagsbreytinga.³

Í Noregi hefur efnið í *Bokhylla.no* þegar verið nýtt til ýmiss konar rannsókna. Þar má nefna orðfræðirannsóknir þar sem skyldleiki orða er skoðaður⁴ og stílmælingar sem urðu til þess að finna höfund sem skrifaði á 6. áratugnum undir dulnefni umdeilda bók, *Blokk 7. Daglig liv i „Kjærlighetsredet“ i en tysk konsentrasjonsleir*.⁵ Þá hefur verið settur upp n-stæðuskoðari sem gefur almenningi kost á að skoða tíðni orða og orðasambanda í textum eftir útgáfuárum og þannig átta sig á tísku í orðnotkun eða vinsældum fólks eða menningarfyribæra.⁶

² Sjá t.d. viðeigandi kafla í skýrslunni Máltækni fyrir íslensku 2018-2022: verkáætlun.

(<https://www.stjornarradid.is/lisalib/getfile.aspx?itemid=56f6368e-54f0-11e7-941a-005056bc530c>)

³ Frekari umfjöllun og dæmi um notkun textasafna í félagsvísindum má t.d. finna á vef ESRC Centre for Corpus Approaches to Social Science við Lancaster-háskóla í Bretlandi: http://cass.lancs.ac.uk/?page_id=43

⁴ Sjá útdrátt í ráðstefnuriti DHN 2016: <http://www.hf.uio.no/iln/english/research/networks/digital-humanities/news-and-events/events/2016/pdf/bofab.pdf>

⁵ Sjá sama rit.

⁶ https://www.nb.no/sp_tjenester/beta/ngram_1/

Að auki má nefna að nú vinna tveir doktorsnemar að doktorsverkefnum við Háskólann í Osló sem snúast um greiningu efninu í *Bokhylla.no* í verkefni sem heitir Data-Mining the Digital Bookshelf.⁷

Af því sem hér hefur komið fram má ráða að fjölmargir nýir möguleikar myndu opnast í rannsóknum á ýmsum sviðum með stafvæðingu alls íslensks prentmáls, ekki síst á rannsóknarsviðum Stofnunar Árna Magnússonar í íslenskum fræðum en einnig á öðrum sviðum. Að auki myndi greining á stafrænt endurgerðum textum gagnast við að bæta þau máltækniverkfæri sem nú stendur til að þróa í máltækniáætlun sem ríkisstjórnin ákvað í stjórnarsáttmála að ráðast í og stendur til að hefjist á þessu ári. Við undirrituð lýsum yfir stuðningi við tillöguna en viljum árétta nauðsyn þess að tryggja möguleika á aðgangi að efninu til rannsókna, rétt eins og gert er í norska *bokhylla*-verkefninu í 4. grein samnings norska landsbókasafnsins og rétthafasamtakanna. Við skipan starfshópsins sem fjallað er um í þingsályktunartillögunni er því mikilvægt að fulltrúi rannsóknasamfélagsins yrði skipaður í hópinn, með þekkingu á textarannsóknum, gagnamálfræði eða stafrænum hugvísindum.

Virðingarfyllt

f.h. Stofnunar Árna Magnússonar í íslenskum fræðum

Steinþór Steingrímsson, verkefnisstjóri í máltækni

Ásta Svavarsdóttir, stofustjóri á orðfræðisviði

Þórdís Úlfarsdóttir, orðabókaritstjóri

Jón Hilmar Jónsson, orðabókahöfundur

⁷ <http://www.hf.uio.no/iln/english/research/networks/digital-humanities/news-and-events/events/2016/digital-humanities-forum/data-mining-the-digital-bookshelf.html>



ÞJÓÐMINJASAFN ÍSLANDS
National Museum of Iceland

Nefndasvið Alþingis
Austurstræti 8-10
150 REYKJAVÍK

Reykjavík, 26. mars 2018
Tilvísun: 201803-0008/00.03

Efni: Umsögn Þjóðminjasafns Íslands um tillögu til þingsályktunar um gerð áætlunar um stafræna endurgerð íslensks prentmáls, 219. mál

Þjóðminjasafn Íslands fagnar því að tekin séu skref til áframhaldandi stafrænnar endurgerðar menningararfsins og miðlunar hans til gagns fyrir þjóðina. Skýrsla mennta- og menningarmálaráðuneytis *Varðveisla menningararfleifðar á stafrænu formi. Staða hjá söfnum og menningarstofnunum landsins* frá 2017 var mikilvægt skref í að öðlast yfirsýn yfir stöðu mála á þessu sviði. Með fyrirliggjandi þingsályktunartillögu er lagt til að stafræn endurgerð íslensks prentmáls verði sett í forgang umfram annað efni sem fjallað er um í skýrslunni. Þó kemur þar fram að ástand annarra tegunda safneigna, eins og t.d. ljósmyndir, kvikmyndir og hljóðupptökur sem sumar liggja undir skemmdum, gefi tilefni til að forgangsörðun þurfi að vera önnur en hér er lagt til. Ekki er verið að draga úr mikilvægi þess að styðja við íslenskt mál þegar hvatt er til þess að næstu skref í stafrænni endurgerð menningararfsins verði tekin með fleiri sjónarmið að leiðarljósi. Þá er það jákvætt skref að kanna eigi möguleika á að ná samningum við höfundaréttarhafa um slík afnot stafrænnar endurgerðar ritverka. Það er mikilvægt skref, en um leið skal bent á brýna þörf fyrir að kanna möguleika á að ná einnig samningum við höfundarrétthafa myndefnis.

Virðingarfyllst,

Margrét Hallgrímsdóttir
þjóðminjavörður