

Máltækni í þágu samfélagsins

Eiríkur Rögnvaldsson
prófessor emeritus
landsfulltrúi CLARIN og ELG



Þótt nokkur verkefni á sviði máltækni hafi verið unnin hér á landi undir lok 20. aldar má segja að skipuleg uppbygging íslenskrar máltækni hafi hafist fyrir rúmum 20 árum með úttekt á ástandi og horfum í íslenskri tungutækni eins og sviðið var kallað þá. [Tungutækni – skýrsla starfshóps](#) kom út snemma árs 1999 en þar var áætlað að það kostaði u.þ.b. einn milljarð króna að gera íslenska máltækni sjálfbæra. Að teknu tilliti til verðlagsbreytinga er þetta reyndar svipuð upphæð og nú er gert ráð fyrir að verja til verkefnisins.

Í kjölfar skýrslunnar kom ríkið á fót sérstakri [tungutækniáætlun](#) sem fékk samtals 133 milljónir króna á fjárlögum árána 2000-2004. Fyrir það fé voru unnin eða sett af stað ýmis verkefni sem sum lognuðust út af, önnur lögðu drög að seinni tíma starfi á þessu sviði, og enn önnur eru enn í fullu gildi. Því fór þó fjarri að íslensk máltækni væri orðin sjálfbær þegar áætluninni lauk og næstu 10 ár var ekki sett neitt fé á fjárlögum í íslenska máltækni en með fáeinum styrkjum úr [Rannsóknasjóði](#), m.a. einum [öndvegisstyrk](#), tókst að halda lífi í rannsóknar- og þróunarstarfi.

Þáttaskil urðu með þátttöku Íslands í evrópska [META-NET](#) verkefninu árin 2011-2013. Meginmarkmið þess voru annars vegar að gera ítarlega úttekt á stöðu máltækni fyrir 30 evrópsk tungumál, og hins vegar að byggja upp og gera aðgengileg hvers kyns málleg gagnasöfn og hugbúnað fyrir þessi mál. Á Íslandi tókst að safna saman og byggja upp margvísleg gögn á þessum tíma. Þau voru gerð aðgengileg í varðveislusafninu [META-SHARE](#) en einnig var settur upp vefurinn [málföng.is](#) þar sem hægt er að nálgast þessi gögn.

Auk þess var skrifuð ítarleg skýrsla, ein af 30 í sama sniði: [Íslenska á stafrænni öld](#). Þegar niðurstöður skýrslanna voru bornar saman kom í ljós að íslenska stæði næstverst málanna 30 hvað varðar máltækni. Þessar niðurstöður vöktu töluverða athygli hér á landi og voru m.a. [ræddar á Alþingi](#). Það er óhætt að segja að skýrslan hafi verið einn helsti hvatinn að [þingsályktun](#) sem var samþykkt einróma 2014 um gerð áætlunar um uppbyggingu íslenskrar máltækni og leiddi að lokum til skýrslunnar [Máltækni fyrir íslensku 2018-2022 – Verkáætlun](#).

Þátttakan í META-NET sýndi okkur mikilvægi alþjóðlegs samstarfs á þessu sviði og í máltækniskýrslunni var lagt til að Ísland gerðist aðili að evrópska innviðaverkefninu [CLARIN ERIC](#). Það var sett á stofn til að staðla og gera aðgengileg gagnasöfn til nota í rannsóknum á sviði félags- og hugvísinda, en þessi gagnasöfn geta einnig nýst í máltækni og þess vegna er aðild að CLARIN mjög gagnleg. Ísland er þegar orðið [áheyrnaraðili](#) að samstarfinu og nú hefur íslenskum lögum verið breytt þannig að Ísland gæti sótt um fulla aðild.

Samstarfsnetið [European Language Resource Coordination](#) hefur verið í gangi undanfarin fjögur ár. Það gengur einkum út á að afla mállegra gagna frá opinberum stofnunum til nota í vélrænum þýðingum og nýta þau í opinberri þjónustu. Annað samstarfsnet er [European Language Grid](#) sem er nýtt og snýst aðallega um gögn sem nýtast fyrirtækjum í iðnaði og viðskiptalífi. Ísland tekur þátt í báðum þessum netum sem skiptir gífurlegu máli til að geta fylgst með þróuninni, fengið aðgang að gögnum og þekkingu og komið okkar eigin gögnum á framfæri.

Máltækniáætlunin sem nú er verið að setja af stað mun kosta á þriðja milljarð króna. Þótt það sé ekki stór hluti af heildarútgjöldum ríkissjóðs er þetta mikið fé, a.m.k. í samanburði við það sem við sem vinnum með íslenskt mál erum vön að sjá. Ég er samt sannfærður um að fáar fjárfestingar borga sig betur og þjóðin á eftir að fá þetta margfalt til baka. Í huga margra tengist þetta verkefni fyrst og fremst verndun og varðveislu íslenskunnar og hún verður vitanlega ekki metin til fjár. En máltækni er mikilvæg af fjölmörgum öðrum ástæðum og ég nefni örfá dæmi:

- Fyrir jafnrétti og mannréttindi – máltækni getur skipt sköpum fyrir fjölmarga sem búa við einhvers konar hömlun eða skerðingu, auðveldar þeim að ferðast um, bætir aðgengi þeirra að margs kyns þjónustu og lífsgæðum og getur gert þeim kleift að taka fullan þátt í daglegu lífi og starfi.
- Fyrir tækniþróun og nýsköpun – máltækni er alþjóðleg og innan máltækniverkefnisins gefst íslenskum fyrirtækjum tækifæri til að vinna með vísindamönnum að þróun nýrrar tækni og afurða sem geta nýst á alþjóðlegum markaði, auk þess sem aðferðirnar nýtast víðar en í máltækni.
- Fyrir hagkvæmni og hagræðingu í rekstri – með hjálp talgreiningar, talgervingar og gervi-greindar geta fyrirtæki sem reka þjónustuver t.d. látið tölvur sinna verulegum hluta aðstoðarbeiðna sem bæði lækkar kostnað og stytta biðtíma og eykur þannig ánægju viðskiptavina.
- Fyrir skilvirkni í opinberri þjónustu – með notkun máltækni má hraða samskiptum milli opinberra aðila innbyrðis en fyrst og fremst stórbæta og auðvelda aðgengi almennings að þjónustu opinberra stofnana og hraða afgreiðslu erinda og fyrirspurna.
- Fyrir hraða og öryggi í heilbrigðisþjónustu – með notkun máltækni við greiningu og lýsingu rannsóknargagna og hvers kyns skráningu og miðlun upplýsinga má hraða upplýsinga-streymi og þannig flýta fyrir sjúkdómsgreiningu og auka líkur á lækningu.
- Fyrir almannavarnir við náttúruvá – með sjálfvirkum textaskrifum og vélrænum þýðingum má t.d. koma upplýsingum um yfirvofandi eldgos eða flóð á fjölda tungumála til ferðamanna hvar sem þeir eru staddir og bæta samskipti milli viðbragðsaðila innbyrðis og við almenning.

Tölvutækni er nú orðin fléttuð inn í flestar daglegar athafnir okkar. Tungumálið er helsta samskiptatæki okkar á öllum sviðum mannlífsins og þess vegna þurfum við máltækni á öllum sviðum – íslenska máltækni. Uppbygging hennar og þróun er sannarlega í þágu samfélagsins.