

RISKS MŪSU VALODĀM UN IZAICINĀJUMS VALODU TEHNOLOĢIJĀM

Mēs esam liecinieki digitālajai revolūcijai, kas atstājusi milzu ietekmi uz saziņu un sabiedrību. Pēdējie digitālās informācijas un sakaru tehnoloģijas jaunumi nereti tiek salīdzināti ar Johana Gūtenberga izgudroto iespiedpresi. Ko šis salīdzinājums varētu mums pavēstīt par Eiropas informācijas sabiedrības nākotni un jo īpaši par mūsu valodām?

Pašlaik mēs esam liecinieki digitālajai revolūcijai, ko var salīdzināt ar Johana Gūtenberga izgudroto iespiedpresi.

Gūtenberga izgudrojums radīja iespēju īstenot patiesus jaunievedumus saziņas un zināšanu apmaiņas jomā, piemēram, Lutera Bībeles tulkojumu dažādās valodās. Nākamajos gadsimtos kultūras jomā notiekošā attīstība veicināja valodu apstrādi un zināšanu apmaiņu:

- lielāko valodu ortogrāfiskā un gramatiskā standartizācija veicināja jaunu zinātnisku un intelektuālu ideju ātru izplatīšanos;
- oficiālo valodu attīstība sniedza pilsoņiem iespēju sazināties noteiktās nozarēs (piemēram, politikā);
- valodu mācīšana un tulkošana radīja saziņas iespējas dažādu valodu lietotājiem;
- rediģēšanas un bibliogrāfisko vadlīniju izveide nodrošināja drukāto materiālu kvalitāti un pieejamību;
- dažādu masu saziņas līdzekļu, piemēram, grāmatu,

laikrakstu, radio, televīzijas, izveide apmierināja dažādas saziņas vajadzības.

Pēdējo divdesmit gadu laikā informācijas tehnoloģijas ir automatizējušas un atvieglojušas daudzus procesus:

- datorizdevniecības programmatūra ir aizstājusi mašīnrakstītāju un burtliču darbu;
- programma *Microsoft PowerPoint* ir nomainījusi projicēšanas iekārtu transparentus;
- e-pasts nosūta un saņem dokumentus daudz ātrāk nekā faksa aparāts;
- *Skype* piedāvā lētas tālsarunas internetā un virtuālas sanāksmes;
- audio un video kodēšanas formāti atvieglo apmaiņu ar multivides saturu;
- meklētājprogrammas nodrošina piekļuvi tīmekļa lapām, izmantojot atslēgvārdus;
- tiešsaistes pakalpojumi, piemēram, *Google tulkotājs*, piedāvā ātrus, aptuvenus tulkojumus;
- sociālās saziņas platformas, piemēram, *Facebook*, *Twitter* un *Google+*, atvieglo saziņu, sadarbību un informācijas koplietošanu.

Lai gan šādi rīki un lietojumprogrammas ir noderīgi, tomēr pagaidām tie vēl nespēj sniegt vajadzīgo atbalstu ilgtspējīgai daudzvalodu Eiropas sabiedrībai, kurai nepieciešama brīva informācijas un preču plūsma.

2.1 VALODU BARJERAS KĀ ŠĶĒRSLIS EIROPAS INFORMĀCIJAS SABIEDRĪBĀ

Mēs nespējam prognozēt, kāda tieši būs nākotnes informācijas sabiedrība. Tomēr pastāv liela varbūtība, ka revolūcija sakaru tehnoloģijā piedāvās kardināli jaunas metodes, lai cilvēki, kas runā dažādās valodās, varētu saprasties. Tas veicina jaunu valodu apguvi un izstrādātājus īpaši mudina izveidot jaunas lietojumprogrammas, kuras varētu nodrošināt savstarpēju saprašanu un piekļuvi kopīgajām zināšanām. Izmantojot jaunus multivides līdzekļus, globālajā ekonomikas un informācijas telpā mijiedarbība starp daudzām valodām, to lietotājiem un saturu notiek daudz ātrāk. Sociālās saziņas vietņu (*Vikipēdija, Facebook, Twitter, YouTube, Google+*) šā brīža popularitāte ir tikai aisberga virsotne.

Globālās ekonomikas un informācijas telpas izaicinājums ir daudzās valodas, to lietotāji un saturs.

Mūsdienās mēs varam dažās sekundēs izsūtīt teksta gigabaitus pa visu pasauli un tikai tad saprast, ka tie ir citiem nesaprotamā valodā. Saskaņā ar Eiropas Komisijas neseno ziņojumu 57% interneta lietotāju Eiropā iegādājas preces un pakalpojumus valodās, kas nav viņu dzimtā valoda (visbiežāk izmantotā valoda ir angļu valoda, tai seko franču, vācu un spāņu valoda). 55% lietotāju prot lasīt tekstu kādā svešvalodā, bet tikai 35% lietotāju izmanto citu valodu, lai rakstītu e-pasta ziņojumus vai nosūtītu komentārus tīmeklī [7]. Pirms pāris gadiem angļu valoda kalpoja par saziņas valodu tīmeklī, jo lielākā daļa saturs tajā bija angļu valodā, toties tagad situācija ir būtiski mainījusies. Tīmekļa saturs citās Eiropas valodās (kā arī Āzijas un Tuvo Austrumu valodās) ir strauji pieaudzis.

Pārsteidzoši, bet visaptverošā digitālā plaisa valodu barjeru dēļ nav guvusi lielu sabiedrības uzmanību, toties tā uzdod kādu svarīgu jautājumu: kuras no Eiropas valodām uzplauks informācijas un zināšanu sabiedrībā, bet kurām būs lemts izzust?

2.2 MŪSU APDRAUDĒTĀS VALODAS

Iespiedprese palīdzēja pāriet uz citu informācijas apmaiņas līmeni Eiropā, bet vienlaikus izraisīja daudzu Eiropas valodu izzušanu. Drukā darbi reti tika iespiesti reģionālajās un mazākumtautību valodās. Tādas valodas kā kornvoliešu un dalmāciešu valoda tika lietotas galvenokārt mutiski, tā samazinot to lietošanas jomu. Vai internetam būs tāda pati ietekme uz mūsu valodām?

Eiropas valodu daudzveidība ir viena no tās vērtīgākajām un svarīgākajām kultūras bagātībām.

Apmēram 60 Eiropā runātās valodas ir viena no tās vērtīgākajām un nozīmīgākajām kultūras bagātībām, kā arī īpaša sabiedrības modeļa neatņemama sastāvdaļa [8]. Tādas valodas kā angļu un spāņu, visticamāk, noturēs savas pozīcijas augošajā digitālajā tirgū, bet daudzas Eiropas valodas tīmekļa sabiedrībā var kļūt nevajadzīgas. Tas novājinātu Eiropas globālo stāvokli un būtu pretrunā ar stratēģisko mērķi nodrošināt ikviena Eiropas pilsoņa līdzdalību sabiedriskajā dzīvē neatkarīgi no valodas. Saskaņā ar UNESCO ziņojumu par daudzvalodību valodas ir svarīgs priekšnoteikums, lai tiktu īstenotas tādas cilvēka pamattiesības kā izteikšanās brīvība, tiesības uz izglītību un līdzdalību sabiedriskajā dzīvē [9].

2.3 VALODU TEHNOLOĢIJAS – KRITISKI SVARĪGAS TEHNOLOĢIJAS

Agrāk galvenie centieni valodu saglabāšanā tika vērsti uz valodu mācīšanu un tulkošanu. Saskaņā ar aprēķiniem rakstiskās un mutiskās tulkošanas, programmatūras lokalizācijas un tīmekļa vietņu globalizācijas izmaksas Eiropas tirgū 2008. gadā bija EUR 8,4 miljardi, un tiek lēsts, ka šī summa ik gadu pieaugs par 10% [10]. Tomēr šis skaitlis ir tikai maza daļa no tā ieguldījuma, kas gan pašlaik, gan nākotnē būs nepieciešams saziņai starp valodām. Vislabākais risinājums valodu lietojuma paplašināšanai nākotnes Eiropā ir atbilstošu tehnoloģiju izmantošana — tieši tāpat, kā mēs izmantojam tehnoloģijas, lai risinātu, piemēram, ar transportu, siltumapgādi vai invaliditāti saistītās vajadzības.

Valodu tehnoloģijas palīdz cilvēkiem sadarboties un dalīties zināšanās dažādās valodās.

Digitālās valodu tehnoloģijas (attiecas uz visiem rakstiskās un runātās valodas veidiem) palīdz cilvēkiem sadarboties, kārtot darījumus, dalīties zināšanās un piedalīties sociālās un politiskās diskusijās — neatkarīgi no valodu barjerām un datorprasmēm. Bieži tās ir nemanāmi iekļautas sarežģītās programmatūras sistēmās, lai mums palīdzētu:

- atrast informāciju, izmantojot interneta meklētājprogrammas;
- tekstastrādes programmā pārbaudīt pareizrakstību un gramatiku;
- skatīt ieteikumus par precēm tiešsaistes veikalā;
- noklausīties automašīnas navigācijas sistēmas mutiskās norādes;
- tulkot tīmekļa lapas, izmantojot tiešsaistes pakalpojumus.

Valodu tehnoloģijas ietver vairākas pamata lietojumprogrammas, kas nepieciešamas visdažādākajām sistēmām un lietojumiem. META-NET valodai veltīto balto grāmatu mērķis ir pievērst uzmanību šo pamattehnoloģiju gatavības līmenim katrā Eiropas valodā.

Eiropai ir vajadzīgas stabilas un pieejamas valodu tehnoloģijas visām Eiropas valodām.

Lai Eiropa spētu saglabāt savu vadošo lomu globālajā pasaulē, tai būs nepieciešamas visām Eiropas valodām pielāgotas valodu tehnoloģijas, kas ir stabilas, viegli pieejamas un iekļautas plaši lietotās ierīcēs un sistēmās. Lai viedierīču un datoru lietotājiem nodrošinātu patiesi efektīvas, interaktīvas multivides un daudzvalodu iespējas, jau tuvākā nākotnē neiztikt bez valodu tehnoloģijām.

2.4 VALODU TEHNOLOĢIJU IESPĒJAS

Grāmatu iespiešanas jomā liels tehnoloģiju sasniegums bija teksta (lapas) attēla ātra pavairošana, izmantojot piemērotu iespiedpresi. Taču vēl aizvien melno darbu veica cilvēki, kas uzmeklēja, lasīja, tulkoja un apkopoja zināšanas. Lai varētu ierakstīt runāto valodu, nācās gaidīt Tomasa Edisona izgudrojumu, bet arī viņa ieviestā tehnoloģija tikai veidoja analogas kopijas.

Digitālās valodu tehnoloģijas tagad var automatizēt visu Eiropas valodu tulkošanas, satura radīšanas un zināšanu pārvaldības procesu. Tās arī nodrošina iespēju sadzīves elektronikā, iekārtās, transportlīdzekļos, datoros un robotos lietot intuitīvi saprotamu interfeisu, kurā par pamatu izmantota valoda/runa. Tirdzniecībā un rūpniecībā izmantotās lietojumprogrammas joprojām ir to atbilstības pirmsākumos, tomēr pētniecības un izstrādes sasniegumi paver ceļu nebijušām iespējām. Piemēram, atsevišķās nozarēs mašīntulkošanas programmu rezultāti

ir jau diezgan precīzi, un eksperimentālās lietojumprogrammas nodrošina iespēju pārvaldīt informāciju un zināšanas vairākās valodās, kā arī veido saturu daudzās Eiropas valodās.

Kā jau tas parasti notiek, arī valodu tehnoloģiju attīstības pirmsākumos lietojumprogrammas (piemēram, lietotāja interfeiss un dialogu sistēmas, kuru pamatā izmantota balss) bija paredzētas ļoti specifiskām nozarēm, un to iespējas nereti bija ierobežotas. Tomēr izglītības un izklaides jomā valodu tehnoloģijām ir visplašākās tirgus iespējas. Tās var izmantot spēlēs, kultūras mantojuma vietnēs, izglītojošās izklaides materiālos, bibliotēkās, simulācijas vidēs un mācību programmās. Mobilie informācijas pakalpojumi, datorizēta valodu apguves programmatūra, e-mācību vide, pašnovērtēšanas rīki un plaši atklāšanas programmatūra ir tikai dažas no jomām, kurās valodu tehnoloģijām var būt ļoti liela nozīme. Sociālās multivides lietojumprogrammu, piemēram, *Twitter* un *Facebook*, popularitāte rosina domāt par vajadzību pēc gudrām valodu tehnoloģijām. Tās varētu izmantot, lai pārraudzītu publicēto informāciju, apkopotu diskusijas un populārākos viedokļus, analizētu emocionālos vērtējumus, identificētu autortiesību pārkāpumus un ļaunprātīgas izmantošanas gadījumus.

Valodu tehnoloģijas palīdz pārvarēt valodu daudzveidības radītos šķēršļus.

Valodu tehnoloģiju izmantošana var būt lielisks risinājums ar Eiropas Savienības daudzvalodību saistītajām sarežģītajām problēmām. Eiropas uzņēmumos, organizācijās un skolās līdzās pastāv dažādas valodas. Raugoties Eiropas Savienības kartē, varam iztēloties, ka tā ir krustām šķērsām pārdalīta neredzamām valodu robežām. Taču, lai pilsoņi varētu savā starpā sazināties, šīs robežas ir jāpārvar, un valodu tehnoloģijas var mums palīdzēt, vienlaikus arī atbalstot atsevišķu valodu neierobežotas lietošanas iespējas. Raugoties vēl tālākā nākotnē,

Eiropas novatoriskās daudzvalodu tehnoloģijas būs labs paraugs mūsu globālajiem partneriem, kad viņi sāks veidot savas daudzvalodu kopienas. Valodu tehnoloģijas ir neaizstājams atbalsts valodiskās daudzveidības šķēršļu pārvarēšanā un valodu kopienu savstarpējā saziņā.

Visbeidzot — šobrīd tiek aktīvi pētītas iespējas, kā valodu tehnoloģijas izmantot glābšanas darbos nelaimes piemeklētos apgabalos, kad ātra darbība ir dzīvības un nāves jautājums. Ļoti iespējams, ka nākotnes intelligentie roboti ar starpvalodu prasmi varēs glābt cilvēku dzīvības.

2.5 VALODU TEHNOLOĢIJU IESPĒJAMIE IZAICINĀJUMI

Kaut gan dažos pēdējos gados valodu tehnoloģijas ir ievērojami progresējušas, pašreizējais tehnoloģiskās attīstības un novatorisku produktu izstrādes temps ir pārāk gauss. Tādas plaši izmantotas tehnoloģijas kā pareizrakstības un gramatikas pārbaude tekstastrādes programmās parasti ir vienā valodā un ir pieejamas tikai dažām valodām. Tiesīsaistes mašintulkošanas pakalpojumi piedāvā ātru un pieņemamas kvalitātes aptuveno tulkojumu, taču šos pakalpojumus nevar izmantot, ja nepieciešams precīzs un pilnīgs tulkojums. Tā kā cilvēku valoda ir sarežģīta, tās atveide programmatūrā un pārbaude praksē prasa daudz laika un līdzekļu. Tāpēc Eiropai jāsaģlabā vadošā loma daudzvalodu kopienu tehnoloģijas izaicinājumu pārvarēšanā, ieviešot jaunas metodes, lai paātrinātu tehnoloģisko attīstību visā pasaulē. Tās var būt gan progresīvas metodes skaitļošanā, gan dažādi paņēmieni, piemēram, interneta lietotāju iesaiste datu apstrādē.

Pašreizējais tehnoloģiskā progress ir pārāk gauss.

2.6 VEIDS, KĀ VALODAS APGŪST CILVĒKI UN MAŠĪNAS

Lai parādītu, kā datori apstrādā valodu un kāpēc datorus ir grūti ieprogrammēt apstrādāt atšķirīgas valodas, aplūkosim veidu, kā pirmo un otro valodu apgūst cilvēki, un pēc tam aplūkosim, kā darbojas valodu tehnoloģiju sistēmas. Cilvēki valodu apgūst divējādi. Mazi bērni valodu iemācās, klausoties, kā sarunājas viņu vecāki, māssas, brāļi un citi ģimenes locekļi. Apmēram divu gadu vecumā bērni sāk izrunāt pirmos vārdus un īsas frāzes. To nodrošina cilvēku ģenētiskā spēja dzirdēto atdarināt un pēc tam to pārvērst jēdzienos.

Cilvēki valodu prasmes apgūst divējādi: mācoties piemērus un valodas pamatā esošos likumus.

Vēlāk, apgūstot otru valodu, jau ir nepieciešama lielāka piepūle, jo bērns vairs neatrodas tādā valodas kopienā, kurai apgūstamā valoda būtu dzimtā. Skolā svešvalodu apguve parasti notiek, iegaumējot gramatikas struktūru, vārdus un pareizrakstību, izmantojot vingrinājumus, kas valodu zināšanas parāda abstraktu likumu, tabulu un piemēru veidā. Jo vecāki kļūstam, jo grūtāk ir apgūt kādu svešvalodu.

Līdzīgi valodas apgūst abu kategoriju valodu tehnoloģiju sistēmas. Statistiskā (jeb datus balstītā) metode iegūst lingvistiskās zināšanas no apjomīgas konkrētu tekstu paraugu kolekcijas. Lai iemācītu, piemēram, veikt pareizrakstības pārbaudi, pietiek ar tekstu vienā valodā, bet, lai trenētu mašīntulkošanas sistēmu, ir nepieciešami paralēli teksti divās (vai vairākās) valodās. Tad mašīnmācīšanās algoritms “apgūst” modeļus, pēc kuriem tiek tulkoti vārdi, īsas frāzes un veseli teikumi.

Līdzīgi valodas apgūst arī abu kategoriju valodu tehnoloģiju sistēmas.

Šai statistiskajai metodei var būt nepieciešams miljoņiem teikumu, un kvalitāte pieaug atbilstoši analizētā

teksta apjomam. Tas ir viens no iemesliem, kāpēc meklētājprogrammu izstrādātāji vēlas apkopot pēc iespējas vairāk rakstīta materiāla. Pareizrakstības pārbaude tekstapstrādes programmās un tādi pakalpojumi kā *Google meklēšana* un *Google tulkotājs* balstās uz statistisko metodi. Statistikai ir būtiska priekšrocība: mašīnu var ātri trenēt secīgās apmācības ciklu sērijās, pat ja kvalitāte būtiski atšķiras.

Otra metode, ko izmanto valodu tehnoloģijās un īpaši mašīntulkošanā, ir likumos balstītas sistēmas. Lingvistikas, datorlingvistikas un datorzinātņu jomu eksperti vispirms apraksta valodas likumus ar formāliem līdzekļiem un izveido leksikas krājumus (leksikonus). Tas prasa daudz laika un pūļu. Dažas vadošās likumos balstītās mašīntulkošanas sistēmas atrodas pastāvīgā izstrādes stadijā jau vairāk nekā divdesmit gadus. Likumos balstītajām sistēmām ir būtiska priekšrocība: ekspertiem ir daudz lielākas iespējas kontrolēt valodas apstrādi. Tas ļauj sistemātiski labot kļūdas programmatūrā un sniegt detalizētus skaidrojumus lietotājiem, jo īpaši, ja likumos balstītās sistēmas tiek izmantotas valodas mācīšanā. Lielo izmaksu dēļ likumos balstītās valodu tehnoloģijas līdz šim ir attīstītas tikai lielākajām valodām. Tā kā statistisko un likumos balstīto sistēmu priekšrocības un trūkumi ir komplementāri, pašlaik pētījumi tiek koncentrēti uz hibrīdpieejām, kas apvieno abas šīs metodoloģijas. Tomēr līdz šim rūpnieciskajā izmantojumā šīs pieejas nav bijušas tik veiksmīgas kā pētniecības laboratorijās.

Kā izklāstīts šajā sadaļā, mūsdienu informācijas sabiedrībā daudzu lietojumprogrammu darbība lielā mērā ir atkarīga no valodu tehnoloģijām. Jo īpaši tas vērojams Eiropas ekonomiskajā un informācijas telpā, kurai raksturīga daudzvalodu kopiena. Kaut gan dažos pēdējos gados valodu tehnoloģijas ir ievērojami progresējušas, joprojām pastāv daudz iespēju uzlabot valodu tehnoloģiju sistēmu kvalitāti. Tālāk aplūkosim latviešu valodas lomu Eiropas informācijas sabiedrībā un novērtēsim pašreizējo valodu tehnoloģiju stāvokli latviešu valodai.

The Latvian Language in the Digital Age

Rehm, G.; Uszkoreit, H. (Eds.)

2012, VI, 91 p. 36 illus. in color., Softcover

ISBN: 978-3-642-30875-8