



## Razumiju li računala hrvatski jezik?

Tehnologija još uvijek nije toliko napredovala da bi računala mogla razumjeti hrvatski jezik na isti način kao i ljudi, ali s njima ipak možemo komunicirati. Grana jezikoslovlja koja se bavi razvojem programa za obradu prirodnoga jezika zove se računalno jezikoslovlje. Jezikoslovci i informatičari razvijaju sustave za prepoznavanje i sintezu govora, analizu rečenica i automatizirano strojno prevođenje. Više je takve tehnologije razvijeno za, na primjer, engleski i španjolski jezik, ali takvi se alati i programi razvijaju i za hrvatski jezik. Tako u programu Word paketa računalnih programa Microsoft Office možemo naći pravopisni provjernik koji na temelju ograničenoga popisa riječi za hrvatski jezik, u svim njihovim paradigmatiskim oblicima, obavještava korisnika koje su riječi neispravno napisane. Također postoje programi koji mogu riječima u rečenici na hrvatskome jeziku odrediti vrstu te padež, rod, broj ili glagolsko vrijeme, ovisno o tome o kojoj je vrsti riječi riječ, pa čak i sintaktičku ulogu određene riječi u rečenici. Ipak, pretpostavljamo da sve navedeno ne zvuči dovoljno zanimljivo prosječnomu govorniku hrvatskoga jezika.

Postavimo radije onda drukčija pitanja: Možemo li računalo pitati kako se zove ili kako se osjeća? Može li ono nama postaviti ista ta pitanja i zapamtiti odgovore? Odgovor je na ta pitanja – Da, naravno! Postoje programi koji su osmišljeni upravo za tu svrhu. Oni se na engleskome zovu *chatbotovima* (sljubljenica od *chat* + (ro)bot), a kao hrvatsku istovrijednicu toga engleskog naziva predlažemo naziv *govorni robot*. Za izradu tih programa obično se koristimo AIML-om (Artificial Intelligence Markup Language), što je metajezik za označavanje podataka. Koristeći se AIML-om radimo popis mogućih rečenica koje sugovornik može napisati u razgovoru s govornim robotima. Program zatim odgovara unaprijed određenom rečenicom ili nasumično bira jedan od više mogućih odgovora.

```
<category>
  <pattern>BOK</pattern>
  <template>
    <random>
      <li>Pozdrav.</li>
      <li>Bok i tebi.</li>
      <li>Bok... Tko si ti?</li>
      <li>Bok... U gužvi sam. Javi mi se poslije!</li>
    </random>
  </template>
</category>
```



Kao što se može vidjeti na slici, govorni robot na upit „Bok” odgovora jednim od četiriju odgovora: „Pozdrav.”, „Bok i tebi.”, „Bok... Tko si ti?” i „Bok... U gužvi sam. Javi mi se poslije!”. Možemo se koristiti i posebnim znakovima, kao što su zvjezdica (\*), da bi računalo ponovilo napisanu riječ, ili donja crta ( \_ ), koja zamjenjuje bilo koju riječ. Takav program također može zapamtiti određene podatke. Na primjer, ako nas govorni robot pita: „Kako se zoveš?”, on može pohraniti odgovor i upotrijebiti ga u daljnjemu razgovoru.

Tako sastavljen program zapravo ne analizira značenje pojedinih riječi ni rečeničnu sintaksu. On nas pokušava uvjeriti da nas razumije, a bez stvarnoga razumijevanja dijaloga. Izradu govornih robota potaknuo je slavni britanski matematičar Alan Turing. U svojem članku *Computing Machinery and Intelligence*, objavljenom 1950. godine, Turing je predstavio test koji ispituje inteligenciju računala ili računalnoga programa iz čovjekove perspektive. Što zapravo znači da strojevi razmišljaju jest, naravno, složeno pitanje, pa je Turing osmislio test koji pitanje *Mogu li računala razmišljati?* zamjenjuje pitanjem *Mogu li računala učiniti sve što i bića koja razmišljaju?* Drugim riječima, njegov test zapravo ispituje mogu li računala navesti sugovornike da pomisle da razgovaraju s čovjekom. Turing taj test naziva igrom oponašanja. Upravo se zato biografski film o tome matematičaru iz 2014. godine zove *Igra oponašanja* (eng. *The Imitation Game*). Svake se godine organizira nekoliko natjecanja u sklopu kojih se natječu brojni govorni roboti, a pobjednici su autori govornih robota koji uvjere najveći postotak ispitivača da su se dopisivali s ljudskim bićem. Najpoznatije je natjecanje te vrste *Loebner Prize*. Mediji nam često javljaju o uspjehu pojedinih programa na takvim natjecanjima, ali katkad ih, zbog nedovoljnoga razumijevanja takvih programa, zamjenjuju s razvojem prave umjetne inteligencije!

Na internetu su trenutačno dostupna dva govorna robota s kojima možete razgovarati na hrvatskome jeziku. Prvi je Marvin, nazvan prema liku iz romana *Vodič kroz galaksiju za autostopere* (<http://pandorabots.com/pandora/talk?botid=f98d56804e374cba>). Marvin je osmišljen kao razgovorni program čiji je zadatak odgovarati na pitanja studenata Odsjeka za informacijske i komunikacijske znanosti Filozofskoga fakulteta Sveučilišta u Zagrebu i promicati navedeni studij. Drugi je hrvatski govorni robot Ivek, razgovorni program koji su osmislili Vanja Štefanec i autor ovoga članka. S njim možete razgovarati koristeći se sljedećom poveznicom: <http://hrvatskijezik.eu/>. Oba su navedena programa napravljena s pomoću AIML-a i servisom Pandorabots.

Dugogodišnjim razvojem govornih robota i njihovim povezivanjem s drugim računalnojezikoslovnim tehnologijama kao što su sinteza i prepoznavanje govora dobili smo programe koji mijenjaju naš svakodnevni život. Primjer je takvoga programa Siri, virtualna pomoćnica koja je vjerojatno postavljena na vaš iPhone, a što nas sve čeka u budućnosti, tek ćemo vidjeti!