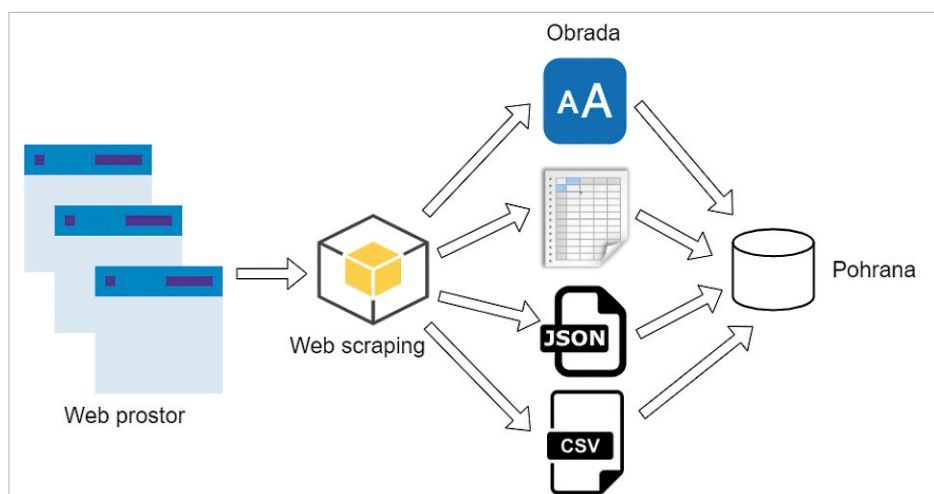


Znanje weba nadohvat ruke

Više puta pisali smo o problemu davanja kvalitetne korisničke podrške i kako nam tu virtualni asistenti mogu pomoći. Pomoću modela umjetne inteligencije i računalne obrade prirodnog jezika, danas je moguće automatizirati dobar dio korisničke podrške, bilo kad je riječ o automatizaciji razgovora s korisnikom, bilo o pronalaženju najrelevantnije informacije za korisnika

Domagoj Marić, Voditelj odjela umjetne inteligencije i znanosti o podacima, Megatrend Poslovna Rješenja



Dobivanje strukturiranih podataka web-scrapingom

Što ako želimo obogatiti našu bazu znanja vanjskim podacima s weba? Moramo li onda sav ciljani sadržaj preuzeti, pa prebaciti u neki oblik prikladan za pretraživanje? *Web-scraping* omogućuje nam da ne moramo.

Dva glavna pojma u području ekstrakcije web-sadržaja su *web-scraping* i *web-crawling*. *Web-scraping* jest sistematizirana ekstrakcija tekstualnog ili medijskog sadržaja s web-stranica, postignuta korištenjem alata zvanih *web-scraperi*. Koncept *web-scrapinga* temelji se na korištenju metoda *web-crawlinga*, automatiziranog sustavnog pretraživanja weba postignutog

praćenjem poveznica web-stranica pomoću *web-crawlera*.

Procesi *web-scrapinga* i *web-crawlinga* čine kontinuirani ciklus: *crawlingom* dolazimo do HTML dokumenata iza web-stranica, iz kojih izvlačimo željeni sadržaj i poveznice na ostale web-stranice pomoću *scrapinga*, te dalje vršimo *crawling* po prikupljenim poveznicama.

VIRTUALNI ASISTENTI

Virtualni asistenti danas su najučinkovitiji način za iskorištavanje modela umjetne inteligencije u svrhu poboljšanja i automatizacije korisničke podrške. U svojoj najprikladnijoj formi, obični *chatbotovi* simuliraju

razgovor s ljudskim agentom te služe kao interaktivna alternativa izlistavanja često postavljanih pitanja i odgovora (koje ljudi ionako ne čitaju) te eliminiraju veliku količinu postavljenih repetitivnih pitanja, a ako nemaju odgovor, preusmjeravaju korisnika na ljudskog agenta.

Pored te osnovne funkcionalnosti, u virtualne asistente moguće je ugraditi razne napredne funkcionalnosti, poput tražilica informacija baziranih na semantičkoj pretrazi dokumenata, analize sentimenta, odnosno tona govora korisnika, dodatnih klasifikacija upita i mnogo toga drugog. Najefektivnija kombinacija virtualnih asistenata je obični *chatbot* koji odgovara na pitanja o tematici na koju je naučen, uz dodatni mehanizam semantičke pretrage – ako asistent ne zna odgovor, traži informacije relevantne za korisnički upit u svojoj bazi znanja, odnosno dokumenata. Tek ako ta pretraga ne rezultira nekom relevantnom informacijom, virtualni asistent preusmjerava korisnika na ljudskog agenta. Time smo eliminirali velik broj upita, onih repetitivnih, jednostavnih, pa čak i neke složenije, specifične upite.

IBM WATSON

Kao vodeću svjetsku platformu za rješavanje NLP problema i integraciju NLP-a u vlastite poslovne procese valja istaknuti IBM Watson. On je, kao sustav za umjetnu inteligenciju, popularan među organizacijama u različitim industrijskim granama zbog jednostavnosti njegove primjene u vlastitom poslovnom okruženju. Vrlo jednostavno i brzo možete razviti vlastiti program za asistenciju u radu s podacima na kojem mogu sudjelovati stručnjaci iz različitih područja. Platforma IBM Watson sadrži raznolik izbor



Virtualni asistent na našoj web-stranici izrađen pomoću platforme IBM Watson Assistant

alata za obradu prirodnog jezika, s manjim ili većim opsegom njihove primjene.

IBM Watson Natural Language Understanding servis je za razne NLP operacije na podacima, te je njihova primjena "najšira". Sadrži mnoge funkcionalnosti, a uključuje i razne nekad odvojene IBM Watsonove servise – Watson Natural Language Classifier (alat za jednostavnu klasifikaciju tekstualnih podataka), Watson Personality Insights (alat za analizu dimenzija osobina ličnosti te osobnih potreba i vrijednosti) te Watson Tone Analyzer (analiza sentimenta, odnosno tona govora).

IBM Watson Studio još je jedan IBM-ov alat koji se primjenjuje za korištenje benefita umjetne inteligencije, točnije strojnog učenja na vlastitom skupu podataka, ali njegova najveća prednost je mogućnost kombiniranja različitih IBM Watsonovih servisa te vanjskog programskog koda u jednom AI okruženju.

IBM WATSON ASSISTANT + DISCOVERY

IBM Watson Assistant zaokružena je platforma za izradu konverzijskih sučelja, odnosno *chatbotova*, koji simuliraju ponašanje čovjeka u razgovoru. Watson Assistant intuitivan je alat pomoću kojeg je moguće u samo sat vremena napraviti svoj vlastiti *chatbot* i postaviti ga u razne komunikacijske kanale, poput Messengera ili Slacka. Sve što je potrebno, primjeri su korisničkih upita koji se klasificiraju u određene dijelove tematike, tzv. *intente*, a tijek razgovora modelira se pomoću grafičkog dijaloznog stabla.

IBM Watson Discovery alat je za inteligentno semantičko pretraživanje podataka pomoću umjetne inteligencije. Također, predstavlja i platformu za analizu teksta koja koristi NLP, kako bi otkrila korisne podatke iz složenih poslovnih dokumenata, web-stranica ili velikih skupova podataka, te

PODRUČJA PRIMJENE WEB-SCRAPINGA

Kompetitivnost cijena

- motrenje konkurencije
- ekstrakcija cijena radi ostvarivanja prednosti na tržištu

Istraživanje tržišnih scenarija (trendova)

- *brand monitoring*
- planiranje lansiranja proizvoda ili usluge na tržište
- *trademark infringement detection*

Analiza sentimenta

- *online reputation management*
- praćenje reakcija potrošača ekstrakcijom ocjena, recenzija i povratnih informacija na forumima i društvenim mrežama

Ljudski resursi

- selekcija i regrutiranje kandidata
- ekstrakcija oglasa za poslove

Agregacija sadržaja

- ekstrakcija i daljnja obrada u organizirane baze znanja

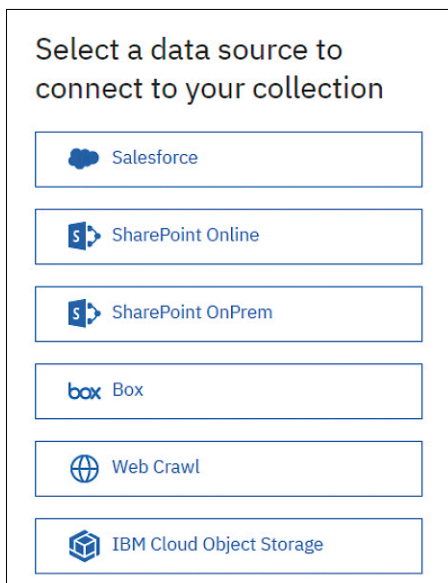
tako skratila vrijeme samog pretraživanja. Discovery omogućuje korisnicima dodavanje vlastitih skupova dokumenata te nad njima primjenjuje algoritme umjetne inteligencije koji obogaćuju umetnute podatke, izvlačeći ključne pojmove i entitete (poput lokacija, organizacija, osoba itd.) te provodi semantičku analizu nad dokumentima.

Nakon prijenosa podataka i njihova obogaćivanja, moguće je graditi upite te in-

tegrirati Discovery u vlastita rješenja ili s drugim IBM-ovim alatima, poput servisa Watson Natural Language Understanding, ili Watson Assistanta. Watson Discovery implementiran je u Watson Assistant preko Search Skills (u samo nekoliko klikova moguće je spojiti ta dva alata) te omogućuje virtualnom asistentu da odgovara na složena pitanja, na koja nema odgovor, pregledavajući veliku bazu dokumenata.

Pored opcije učitavanja vlastitih datoteka u bazu znanja, u alatu IBM Watson Discovery postoji i opcija ugrađivanja Web Crawl kolekcije, koja redovno automatizirano pretražuje web prema parametrima koje specificiramo. Tako u našeg virtualnog asistenta možemo ugraditi dodatni mehanizam pretraživanja sadržaja unutar web-stranica, u slučaju da naš asistent nije naučen na tematiku povezanu s korisničkim upitom, a odgovori će uvijek biti konzistentni i ažurni.

Umjetna inteligencija i računalna obrada prirodnog jezika područja su pomoću kojih je, uz dodatnu pomoć procesa web-scrapinga za dobivanje dodatnih podataka, danas moguće poboljšati razne poslovne procese, a korisnička podrška, koja je svakim danom sve kompleksniji zadatak, savršen je primjer njihove efektivne primjene. Ako želite saznati više o web-scrapingu, virtualnim asistentima, semantičkoj pretrazi dokumenata, IBM Watson platformi ili drugim primjenama umjetne inteligencije u poslovnim procesima, slobodno se obratite našim stručnjacima na poslovna.rjesenja@megatrend.com.



Opcije spajanja na vanjske izvore podataka u servisu IBM Watson Discovery