JEZIČNE TEHNOLOGIJE



Dosta vam je svađanja oko pravopisnih grešaka na internetu? Ne možete naći prave riječi za svoj online argument? Ovi jezični alati uvelike će vam olakšati pisanje

PIŠE: LUCIJA PILIĆ

Koliko vam se puta na raznim forumima i društvenim mrežama dogodilo da ste napisali elokventni, argumentirani, trotomni tekst o negativnom utjecaju krivog nanošenja termalne paste na procesor, samo da bi vas prvi protivnik u temi dočekao s podrugljivim komentarom: "Napisao si č umjesto ć u trećoj rečenici drugog odlomka."? I zatim gledali kako vaš argument i cijela rasprava odlaze u nepovratni off-topic i raspirivanje niskih strasti. Eh, što pravopis i gramatika čine od razumnih ljudi. Srećom, postoji rješenje za vaš problem. Točnije, nekoliko rješenja. Pretražili smo cijeli internet da pronađemo najkorisnije alate, usluge i stranice koje će vam pomoći u eliminaciji pravopisnih i gramatičkih grešaka te obogatiti vaš rječnik. Usporedili smo i testirali masovno korištene alate za provjeru pravopisa i gramatike na hrvatskom jeziku kako bismo otkrili koji je softver u toj nišnoj, ali vrlo bitnoj kategoriji najbolji. I uz sve to, na primjeru najstarijeg hrvatskog pravopisnog provjernika (spelling checkera), doznali kako se razvija(la) tehnologija u pozadini pravopisnih provjernika.



Rječnik stranih riječi
spelling checker, spell-checker – alat za provjeru pravopisa, pravopisni provjernik
grammar nazi – osoba koja živi za ispravljanje tuđih praovpisnih i gramatičkih grešaka

23 godine Hašeka

Prošli mjesec bili smo u jedinstvenoj prilici upoznati tvorca Hrvatskog akademskog spelling checkera. Za uslugu Ispravi.me / Hascheck / Hašek ste sigurno čuli. Ako ste student, novinar ili pisac, vjerojatno se ju i označili na internetskom pregledniku. Ali tko stoji iza te usluge bila je velika misterija oko koje su se raspredale brojne legende na studentskim kavama. Prvenstveno ona da se radi o studentu Fakulteta elektrotehnike i računarstva (FER) koji je iz zafrkancije napravio web za provjeru pravopisa. Je li ta legenda istinita? Žao nam je, ali morat ćemo vas ispraviti.

Idejni začetnik, tvorac i glavni developer ovog sustava je Šandor Dembitz, umirovljeni profesor FER-a. Ideja za Hašek pala mu je na pamet 70-ih godina prošlog stoljeća, kad je radio na korekturi englesko-hrvatskog leksikografskog korpusa pokojnog akademika Željka Bujasa. U vrijeme kad većina stanovništva uopće nije znala što su računala, jezikoslovac Bujas ih je koristio za obradu teksta. Englesko-hrvatski leksikografski korpus bio je prvi dvojezični milijunski korpus napravljen u Hrvatskoj, na flexowriter stroju. "Vi ste premladi da biste uopće znali što je flexowriter. To vam je pisaći stroj koji kao izlaz daje bušenu papirnatu traku. Ono što zatipkate ne može se korigirati na takvom mediju", objašnjava nam Dembitz. U takvim uvjetima rada pojavljivalo se bezbroj grešaka. To je potaknulo Dembitza na razmišljanje kako riješiti problem programski. U isto vrijeme, početkom 70-ih, u Americi su se pojavili prvi spelling checkeri (pravopisni provjernici).

Tijekom izrade doktorata "Automatizacija postupaka otkrivanja grešaka u tekstu u novim telekomunikacijskim službama" 1993. godine Dembitz je ponudio alat prvo FER-u, a zatim u proljeće 1994. i široj javnosti u obliku e-mail embedded usluge. Deset godina kasnije, Hašek je osvanuo na webu, na adresi hascheck.tel.fer.hr, a prije pola godine obukao se u novo ruho te je preimenovan u Ispravi.me.

Dembitz kontinuirano radi na unaprjeđenju jezgrenog dijela usluge, a u

TE(K)ST ZA PROVJERU PRAVOPISNIH PROVJERNIKA

Ovo je tekst za provijeru pravopisa. Ima puno riječi. Rječi se sastoje od slogova. Tekst neče biti dugaćak. Jedino što je važno jest da je ovo tekst za testiranje spell-checkera na hrvatskom jeziku. Hrvatski jezik nema puno spell-checkera. Mnogi programi za provjeru urednosti na engleskom jeziku, osim provjere pravopisa i gramatike, imaju uključene mogućnosti poput provjere originalnosti teksta, čitljivosti, korištenja pasiva, upotrebe priloga, i dr. Svaki spell-checker mora se podvrgnuti detaljnom testiranju grešaka, da provjerimo njegovu brzinu, preciznost i mogućnosti. Pisanje teksta za provjeru pravopisa, umjesto provjeravanja istog, dozlaboga je dosadan posao, koji ne bih poželjela ni najgorem neprijatelju. Posebno je teško osmišljavati proširene rečenice bez pravoga smisla... Npr., da je Ivo Ivić znao kako ga žena vara sa susjedon Markom Smithom, nikad ga ne bi počastio finom čokoladnom tortom, koju je spekao za 18. obljetnicu braka. Tog kobnog dana, kad se vraćao kući s posla da iznenadi i razveseli ženu poklonom za Valentinovo, na radiju je svirala pjesma It Wasn't Me jamajkanskog pjevača Shaggyja. Kamo je otišao nakon što ih je uhvatio in flagranti, nitko ne zna. Javio se ženi telefonom tek nakon šest mjeseci prjetećim tonom tonom: „Judita, ukoliko ne potpišeš papire za razvod, otkrit ću strašnu tajnu o tvojoj prošlosti." Judita je bila užasnuta. Slijedeća stvar koju je napravila zauvijek je promijenila način na koji koristimo programe za provjeru pravopisa.

tome mu, koliko mu obaveze dopuštaju, pomaže i Goran Gledec, FER-ov prodekan za znanost. Usto, više studenata FER-a je tijekom godina pridonijelo razvoju ovog projekta. Korisničko sučelje novog portala osmislio je Ivan Srdić, a dizajn Tomislav Korman. "Sada je u fazi štrikanje koda da se usluga inkorporira kao add-in u Word, odnosno Outlook, tj. da izravno otvorite Word/Outlook, u njemu pozovete Hašek i da se editiranje obavi u dokumentu, što će ubrzati posao uređivanja teksta. Tako mi moji studenti stalno pomažu", ponosno ističe Dembitz.

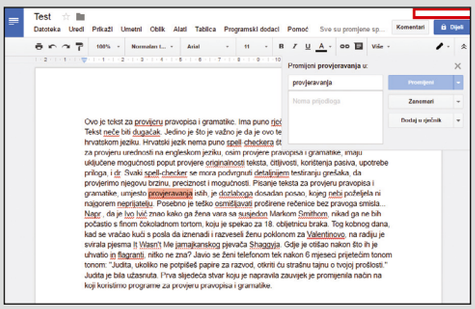
Strojno učenje i ispravljanje

Strojno učenje danas je jedna od top pet tehnoloških tema među direktorima ICT firmi. No, prije 23 godine, kad je još "internet" bila egzotična riječ među inženjerima, strojno učenje, kaže Dembitz, bilo zamišljeno u samom idejnom dijelu sustava: "Dok je jezik živ, alat koji na bilo koji način prati jezik mora pratiti život, biti osposobljen da uči ono što se u jeziku događa." U praksi, sustav dnevno izdvoji oko 300 potencijalno valjanih riječi na koje loše reagira, a Dembitz ih zatim osobno pregledava kako bi sve legitimne postale dijelom leksičke baze sustava. Tu se najčešće radi o ispravno napisanim posvojnim pridjevima izvedenim iz stranih imena (McCormick - McCormickov).

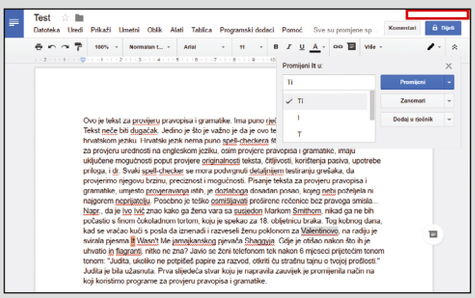
Hašek ne radi samo provjeru pravopisa nego i kontekstualnu provjeru teksta. Bez problema upozorava na pravilnu uporabu prijedloga „s/sa" ovisno o riječi koja za prijedlogom slijedi, označit će pleonazme, upozorit će vas na upotrebu

USPOREDNI TEST PRAVOPISNIH PROVJERNIKA ZA HRVATSKI JEZIK

Uz već navedeni pravopisni provjernik (spelling checker) Ispravi.me (Hascheck / Hašek), za test i usporedbu spelling checkera uzeli smo tri masovno korištena i cjenovno pristupačna programa na ovim područjima: Googleove dokumente, Microsoftov Word i LibreOfficeov Writer. Svi navedeni programi, osim web aplikacije Ispravi.me, ujedno su i alati za uređivanje teksta, a provjera pravopisa samo je jedna od njihovih mogućnosti. Dapače, Microsoftov Word i LibreOfficeov Writer hvale se mogućnošću provjere pravopisa i gramatike (u nazivu modula) pa smo za prvu provjeru ukrasili tekst gramatičkim i pravopisnim greškama. Nažalost, ispostavilo se da je provjera gramatike hrvatskog jezika u oba programa, unatoč imenu alata, mrtvo slovo na digitalnom papiru. Zato je provjera pravopisa (tekst za provjeru u okviru) funkcionirala puno bolje.



GOOGLE DOKUMENTI ne mogu prepoznati obične hrvatske riječi poput "provjeravanja"



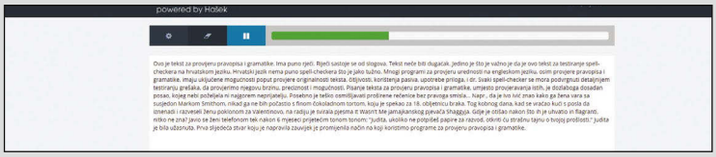
DOSTUPNE su osnovne opcije za zamjenu / ispravak riječi (pa i prijedlozi)

Google dokumenti

Imate li Android mobitel, sigurno ste ih otvorili barem jedanput, da na brzaka pregledate neki tekst i po potrebi ga izmijenite. Google dokumenti su besplatni, možete im pristupiti s bilo koje lokacije, s (gotovo) bilo kojeg uređaja, samo trebate imati Gmail račun. Ipak, cijena koju plaćate za ovaj servis mogla bi biti visoka, ako se odlučite osloniti na isti za provjeru pravopisa. Unatoč tome što Googleovi dokumenti mogu prepoznati elementarne greške (korištenje č i ć, ije i je, tipfelere), zažmirit će na neke očite propuste (zarez viška, ponavljanje riječi, pogrešnu uporabu riječi "ukoliko" i "slijedeća"). Međutim, ono što nas je najviše živciralo u upotrebi ovog alata za provjeru pravopisa je broj pogrešno označenih, "krivo napisanih" riječi. Općenito, alati nisu puno griješili u označavanju krivo napisanih riječi nego su češće viđali greške tamo gdje ih nema. Google dokumenti prvi su među njima. Ovaj alat pogrešno je označio čak 17 ispravno napisanih riječi. Još nam je razumljivo krivo označavanje stranih imena i izraza (spell-checker, Smithom, Shaggyja, in flagranti), ali u pogledu normalnih, hrvatskih riječi poput "originalnosti", "neprijatelju" ili "Valentinovo" alat ne bi smio griješiti. Možete ispraviti grešku Googleovih dokumenata te navedene riječi dodati u rječnik ili ignorirati, a ponekad će vam program dati prijedlog za zamjenu. Radi se o bazičnim mogućnostima koje malo popravljaju, ali ne mogu izbrisati loš dojam koji program ostavlja u pogledu provjere pravopisa.

USPOREDNA TABLICA

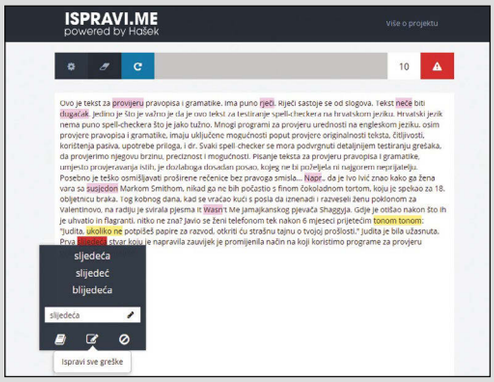
	Google Docs	Ispravi.me	Microsoft Word	LibreOffice Writer
BROJ UOČENIH GREŠAKA	9 / 13	11 / 13	10 / 13	9 / 13
BROJ KRIVO PROCESUIRANIH RIJEČI KAO GREŠKE	17 / 17	1 / 17	6 / 17	14 / 17
+	besplatno, dostupno na svim uređajima	najbolji rezultati na testiranju, besplatno, optimizirano za sve uređaje	ne radi previše grešaka, rječnik sinonima	besplatno, open-source
–	najveći broj grešaka, potrebna registracija za Googleove usluge (ako ne želite samo pregledavati i mijenjati dijeljene datoteke)	morate imati internet vezu kao biste se njime služili	cijena	velik broj grešaka, nema aplikacije za mobilne uređaje
DOJAM	Googleove dokumente zaobilazite u širokom luku kad je u pitanju provjera pravopisa	Najbolji alat za provjeru pravopisa na hrvatskom jeziku	Microsoftov program ima najbolju provjeru pravopisa u konkurenciji softvera za uređivanje teksta	Solidan, besplatan uređivač teksta, ali loš izbor za provjeru pravopisa u istom



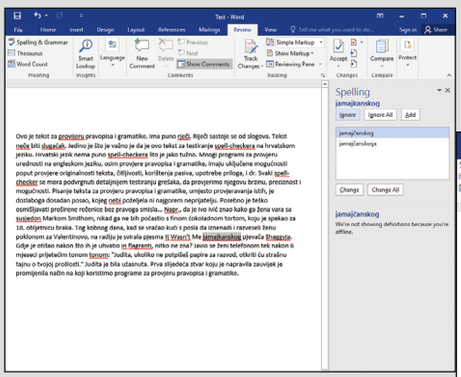
ISPRAVI.ME (Hašek) će za svega nekoliko sekundi skenirati i ispraviti tekst s 30 tisuća znakova

Ispravi.me (Hašek)

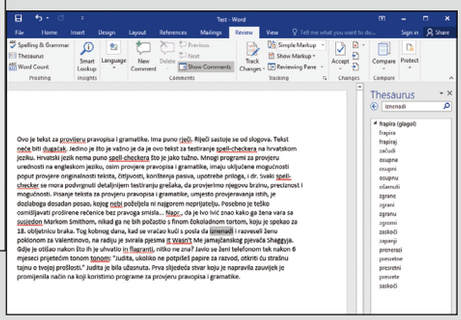
Web aplikacija Ispravi.me (poznata i kao Hašek) koncentrira se isključivo na provjeru pravopisa. I kao takva, potpuno uspijeva u svojoj namjeni. Radi se o najboljem alatu za provjeru pravopisa, pokazali su rezultati. Hašek nije uspio prepoznati dvije greške (zajedno napisan "nebih" i zarez viška), a krivo je označio tek jednu riječ ("Wasn't"). Jedini je alat koji je prepoznao krivu uporabu riječi "ukoliko" i "slijedeća" te dao točne prijedloge za zamjenu. Hašek za svaku označenu riječ nudi zamjenu, ali isto tako riječ možete dodati u rječnik, zanemariti ili ispraviti grešku. Iako ne postoji aplikacija za mobilne uređaje, sama web stranica je optimizirana za prikaz na svim veličinama i rezolucijama ekrana (responsive dizajn). Zanimljiva opcija koju možete podesiti je da Hašek radi automatsku provjeru teksta čim ga zalijepite u okvir i da omogućuje automatsko kopiranje ispravljenog teksta koji onda samo trebate natrag zalijepiti u uređivač teksta.



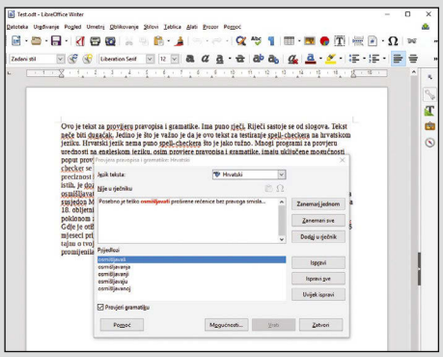
HAŠEK Ne ispravlja samo obične pravopisne greške, nego i označuje riječi koje su u danom kontekstu krivo upotrijebljene



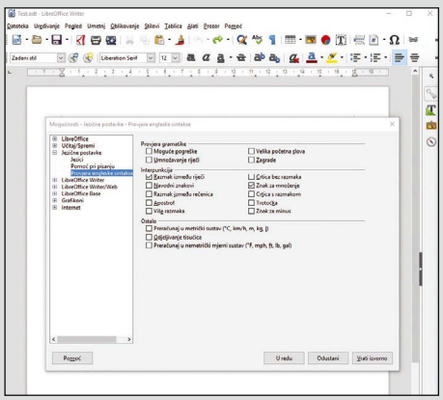
WORD Uglavnom nudi logične prijedloge za zamjenu riječi



RJEČNIK SINONIMA Zgodan i nadasve koristan dodatak ovog uređivača teksta



LIBREOFFICE WRITER Načelno, na papiru nudi mogućnost provjeravanja gramatike te podešavanje naprednih mogućnosti, ali u praksi to malo znači



S DRUGE STRANE Writer za engleski jezik nudi puno robusnije opcije provjere te rječnik sinonima. Usto, program je open-source pa ako mislite da možete napraviti bolji spelling checker, slobodno se pridružite zajednici

Microsoft Word

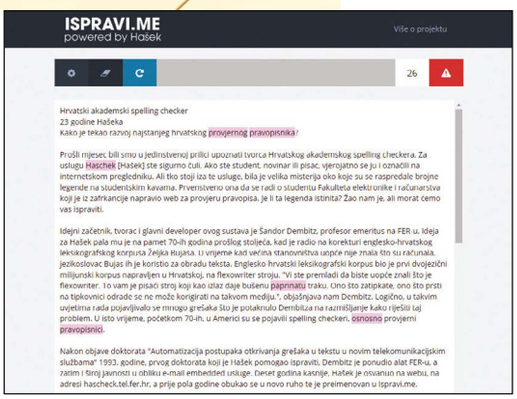
Microsoft Word jedini je program na ovoj listi koji se plaća, ali donekle opravdava tu cijenu. U pogledu prepoznavanja pravopisnih grešaka bolji je od ostalih alata za uređivanje teksta. Propustio je uočiti zarez viška te pogrešnu uporabu riječi "ukoliko" i "slijedeća", ali je zato ukazao na greške u korištenju č i ć, ije i je, tipfelere te ponavljanje riječi. Microsoftov alat za uređivanje teksta također vješto izbjegava krivo označavanje (stranih) riječi. Pogriješio je tek šest puta, i to u pogledu riječi "spell-checker" (spomenuta tri puta), "It Wasn't Me", "jamajkanskog" i "in flagranti". Kao i drugi programi, Microsoft Word vam ostavlja opcije da označenu riječ ispravite, zanemarite ili dodate u rječnik. Ono što bi vam dodatno moglo biti korisno je rječnik sinonima. Na provjeru gramatike nemojte računati. Microsoft Word košta od 69 eura godišnje (pretplatni model) do 109,99 dolara za stalnu licencu. Košta, ali se po pitanju provjere pravopisa (i mnogih drugih značajki) isplati više nego alternativni alati za uređivanje teksta. Ovaj program je također dostupan za mobilne uređaje.

LibreOffice Write

Najpopularniji open-source program za uređivanje teksta, normalno, ima predinstaliranu opciju za provjeru gramatike (koja ne radi kako treba) i pravopisa. Ali bosanskog jezika, ne hrvatskog! "Rječnik je permisivan što znači da u pravilu dozvoljava i "istočnu" i "zapadnu varijantu" pojedinih riječi. Zbog toga je upotrebljiv i za hrvatski i srpski jezik pod uslovom da vam nije stalo do potpune usklađenosti s aktuelnom jezičkom normom ovih jezika", piše na službenoj stranici LibreOffice alata. Kako to u praksi funkcionira? Pa... Ne puno bolje od Google dokumenata. Uočava greške poput krive upotrebe č i ć te ije i je, ali propušta označiti greške kao što su zarez viška, ponavljanje riječi te pogrešna uporaba riječi "ukoliko" i "slijedeća". Isto tako, označava određene normalne riječi poput "Valentinovo" kao neispravne. Doduše, u nešto manjoj mjeri nego Google dokumenti. Zbog toga vam preporučujemo da ne uključujete automatsko ispravljanje riječi, inače dio LibreOfficeova Writera. Označenu riječ možete ispraviti, ignorirati ili dodati u rječnik. LibreOffice Writer program je potpuno besplatan (spelling checker se mora dodatno skinuti sa službene stranice), ali je trenutno dostupan samo za klasične, desktop platforme (Linux, macOS, Windows).



ŠANDOR DEMBITZ, umirovljeni profesor FER-a, kreator je sustava kojeg znamo kao Hašek, Hascheck ili, u zadnje vrijeme, Ispravi.me



HAŠEK Samo je u siječnju obradio više od 300 tisuća tekstova ukupnog opsega od 90 milijuna pojavnica

možebitno krive riječi (slijedeći umjesto sljedeći), ali i krivu upotrebu sasvim ispravne riječi. Dembitz navodi primjer Donalda Trumpa, čije prezime mnogi autori omaškom pišu fonetski (Tramp): "Postoji legitimna hrvatska riječ trampa (razmjena). Ako vam se u tekstu pojavi ta riječ pisana s početnim velikim slovom, je li legitimno upozoriti korisnika: 'Nisi li možda mislio na Donalda Trumpa?' U Hrvatskoj imate toponim Trampov breg, postoji zaselak Trampi... Samo, kolika je danas vjerojatnost da će se pisati o Trampovom bregu, a koliko o Donaldu Trumpu? Stoga, ako se pojavi i taj nesretni Trampov breg, opet će program vikati: 'Jesi li možda mislio na Donalda Trumpa?' To je jedna kontekstualnost u širem smislu, koja ilustrira da jezik svaki dan donosi nešto novo. Stalno ga treba pratiti, stalno biti tu, jer živi jezik nije okamina."

S druge strane, Dembitz je oprezan kod učenja vokativa imenica. Oni se u pravilu vrlo rijetko pojavljuju u tekstovima (iznimka su osobna imena), ali su zato česti zatipci. Kao primjere Dembitz ističe vokative imenica „struka" i „obuka" (struče i obuko). „Možete li se sjetiti ijednog primjera teksta u kojem se svrhovito koriste vokativi ovih imenica? Ali zato tekstovi vrve zatipcima poput 'obuko sam se', ili 'struče spreme', posebno u dokumentima mog resornog ministarstva."

Hašek se vrti na Linux Drive poslužitelju koji se sastoji od dva dijela: razvojnog, na kojem rade programeri te uslužnog koji poslužuje korisnike. U principu, učenje obavljeno tijekom dana na razvojnom ažurira se isti dan navečer na uslužnom serveru. Od programskih jezika najviše se koriste Perl i AWK.

Provjera hrvatskog pravopisa od Afganistana do Zambije

Hrvatski jezik je jezik s malim brojem korisnika u svijetu. Od 1000 jezika koji imaju pismo i prisutni su na internetu, samo 10 do 15% se može podičiti pravopisnim provjernikom. Osim toga, hrvatski je visokoflektivan jezik, s velikim brojem oblika koji se mogu napraviti iz pojedine promjenjive vrste riječi.

ŠTO KAD OBIČNI SPELLING CHECKERI NISU DOVOLJNI?

Što kad vam treba pomoć pri provjeri gramatičke ispravnosti teksta? Što kad tekst želite obogatiti živopisnim ili lokalnim izrazima? Na hrvatskom webu smo našli nekoliko stranica / servisa koji će vam pomoći u jezičnim dilemama i dodatno vas opismeniti.

Pravopis

http://pravopis.hr/

Ako ni nakon računalne provjere pravopisa niste sigurni kako se određena riječ piše, najbolje je da se uputite na stranicu Hrvatskog pravopisa, gdje ćete naći sva pravila. Ovdje se nalazi i pravopisni rječnik koji će vam pomoći u dvojbi oko načina pisanja problematične riječi. Pravopis je inače jedan od pet portala koji održava i uređuje Institut za hrvatski jezik i jezikoslovlje. Ostali su (također spomenuti) Struna, Hrvatski terminološki portal, Bolje je hrvatski! i Hrvatski na maturi.



HJP – Hrvatski jezični portal

http://hjp.znanje.hr/

Prva stanica za učenje novog vokabulara. Hrvatski jezični portal (HJP) je rječnička baza s više od 115 tisuća riječi (natuknica). Ako ste čuli za neku novu hrvatsku riječ, ali nemate pojma što ona znači, ovdje ćete gotovo sigurno pronaći odgovor. I ne samo to. Uz definiciju, dan je i primjer korištenja u govoru ili pisanju, izvedeni oblici (deklinacija, glagolska vremena), fraze (ako se mogu povezati s tom riječi) te etimologija i onomastika riječi.



Žargonaut

http://www.zargonaut.com/

Sve one riječi koje su previše kul za standardni rječnik ili su se u govoru pojavile tek nedavno možete pretražiti na Žargonautu. Čisto da se ne zbunite kad vam netko kaže da uzmete trešu za šetnju parkom, cimnut će vas za 5 minuta, da se ne kolindirate bez veze ili da u zrcalu pogledate svoj nosprodukt. Za svaku natuknicu naveden je primjer korištenja u rečenici, mjesto porijekla te vrsta riječi. Za pojedine žargone autori su također priložili informacije o tome koja ih dobna skupina koristi, sinonime te standardne riječi. Svatko može pridonijeti bazi ovog rječnika, ali se definicija riječi i opis naknadno provjerava i nadopunjuje od strane profesionalnih i iskusnih jezičara te prevoditelja, inače pokretača ove stranice.



Hrvatski terminološki portal

http://nazivlje.hr/

Tražite li definiciju neke stručne riječi ili njen hrvatski prijevod, najbolje bi bilo da se uputite na Hrvatski terminološki portal. Ovo je u biti nadogradnja na odličan portal Struna, iza kojeg stoje isti terminolozi. Osim Strune, omogućava pretraživanje baze strukovnih rječnika i glosara, terminoloških izvora Leksigrafskoga zavoda Miroslava Krleže te nazivlja iz hrvatskih norma.



Hrvatski rječnik stopljenica

http://www.stilistika.org/hrvatski-rjecnik-stopljenica

Što su to pobogu stopljenice? Prema definiciji sa službene stranice rječnika, radi se o posebnoj "podvrsti posebnih rječnika koji se od općih rječnika razlikuju po tome što iz cjeline leksika nekoga jezika probiru ciljani dio, u ovome slučaju to su riječi nastale stapanjem neznačenjskih dijelova više ishodišnih riječi i pojmova koje one znače ". Primjerice, Banduševac je

(PRAVO)PISANJE NA STRANOM JEZIKU

U današnje vrijeme teško da možete izbjeći komunikaciju na stranom, najčešće engleskom jeziku. Pogotovo ako radite u IT industriji. Srećom, na internetu možete pronaći pregršt alata za provjeru pravopisa i gramatike na stranim jezicima. Ova dva su nas najviše oduševila.

Grammarly

https://app.grammarly.com/
Jedan od najboljih online alata za provjeru pravopisa i gramatike engleskog jezika. Osnovne funkcionalnosti Grammarlyja (ispravljanje kritičnih grešaka, dodaci za Chrome, Microsoft Office i Windowse) su potpuno besplatne. Prijedlozi za unaprjeđenje vokabulara, provjera originalnosti teksta (plagiranja), nezgrapno izražavanje te provjera od strane profesionalnog lektora se plaćaju (od 29,95 dolara mjesečno). Premium uslugu možete isprobati besplatno na tjedan dana. Jedino što nas smeta kod ovog alata je isključivi odabir između američko-engleskog te britansko-engleskog stila pisanja. Primjerice, ako odaberete američko-engleski način pisanja, softver će vas penalizirati za pisanje riječi "favourite" ili "centre". S druge strane, ako ste naklonjeniji britansko-engleskom stilu, morat ćete ispravljati riječi poput "traveled" ili "analyze" što zna biti naporno ako pišete veći članak.

LanguageTool

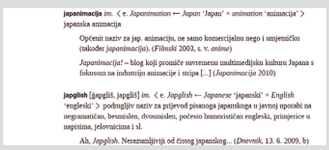
https://languagetool.org/
LanguageTool ne može se mjeriti s Grammarlyjem po mogućnostima i točnosti (u pogledu provjere teksta na engleskom), ali zato podržava provjeru teksta na više od 20 različitih jezika. LanguageTool je open-source alat, što znači da se po želji možete pridružiti, ali i maknuti s projekta. Trenutno je (aktivno) podržano desetak jezika. LanguageTool također ima dodatke za internetske preglednike Chrome i Firefox, programe Google dokumenti, LibreOffice i OpenOffice te Linux, MacOS i Windows platforme.

Razvoj spelling checkera za hrvatski jezik stoga je vrlo izazovan zadatak, objasnio nam je Dembitz. Unatoč tome, to je rad od kojeg nije odustao ni u mirovini. Kreator Hašeka ističe da usluga nikad nije bila sama sebi svrhom. Slobodna je i besplatna za korištenje svim malim korisnicima. Profesionalni korisnici moraju platiti uslugu. Hašek ima dvoznamenkasti broj profesionalnih korisnika, među koje se ubrajaju medijske kuće, teleoperateri i druge tvrtke. Hašek ima korisnike iz 159 IP-domena (zemalja), od Afganistana do Zambije, uključujući i Pitcairnovo otočje sa svega 261 dodijeljenom IP-adresom. Osim što služi kao alat za provjeru pravopisa, putem Hašeka se već deset godina skupljaju hrvatski n-grami. N-gramski sustavi su podloga za razvoj govornih tehnologija poput Siri ili Cortane. Pomoću njih sintetizator govora sposoban je razlikovati upotrebu riječi u kontekstu (luk kao povrće ili luk kao oruđe) te, shodno tome, pravilno ih izgovoriti. Hašek je do sad skupio oko tri-milijarde n-grama. Za usporedbu, najveći n-gramski sustavi, kao što su engleski i kineski, imaju 4 milijarde n-grama. „Dakle, imamo dobru podlogu za razvoj govornih tehnologija, ali nažalost, nema institucionalne potpore da se takvo nešto zaista i napravi, za dobrobit svih hrvatskih građana", s razočaranjem kaže Dembitz, dodajući: „Na vrata nam kucaju internetski kućni asistenti, sutra će se pojaviti i kućni roboti. Hoćemo li s njima u Hrvatskoj razgovarati na engleskome?".

Njegovi studenti su već napravili prototip sintetizatora govora, HascheckVoice kao i sustav za pretvorbu kontinuiranoga govora u tekst ograničen na 15 tisuća riječi s najčešćom uporabom u hrvatskome, ali dodaje da "za napraviti nešto ozbiljnije treba sustavnije ulaganje." I ovaj se razvoj, koji pokazuje da se može kada se hoće, kao i razvoj i održavanje temeljne usluge, financiraju od prihoda klijenata.

Unatoč svemu, Dembitz je ponosan na ostvareno, posebno na široku bazu korisnika koja je s godinama formirana te se nada da će se usluga i dalje masovno koristiti: "Značajno je da korisnici u ovoj zemlji imaju pouzdan alat da mogu brzo i jednostavno svoj pisani uradak napisati pristojno. Ne piše se radi sebe, piše se radi drugog. A urednost pisanja je iskaz poštovanja prema onome kome se obraćaš pisanim putem."

podrugljiv naziv za Bandićeve fontane pokraj NSK nastao stapanjem riječi Bandić i Manduševac.


japanimacija im. < e. *Japanimation* ← *Japan* 'Japan' × *animation* 'animacija' > japanska animacija

Općenit naziv za jap. animaciju, ne samo komercijalnu nego i umjetničku (također *japanimacija*). (*Filmski* 2003, s. v. *anime*)

Japanimacija! – blog koji promiče suvremenu multimedijsku kulturu Japana s fokusom na industriju animacije i stripa […] (*Japanimacija* 2010)

jeglish [gagliš, jegliš] im. < e. *Jeglish* ← *Japanese* 'japanski' × *English* 'engleski' > podrugljiv naziv za prijevod pisanoga japanskoga u javnoj uporabi na negramatičan, besmislen, dvosmislen, počesto humoristično engleski, primjerice u napisima, jelovnicima i sl.

Ah, *Jeglish*. Nerazumljiviji od čistog japanskog... (*Dnevnik*, 13. 6. 2009, b)

Bolje je hrvatski!

http://bolje.hr/
Što bismo mi bez gaming računala, smartphoneova i ostalih gadgeta? Vjerojatno bismo se koristili računalima za igranje videoigara, pametnim telefonima i ostalim (pametnim) spravicama. Ako ste ojađeni poplavom anglizama u hrvatskom jeziku, pogotovo u ovom tehnološkom području te tražite alternativu (ili ju možda želite predložiti) na hrvatskom jeziku, nećete naći stranicu bolju od ove. Nažalost, na stranici je prevedeno tek nešto više od stotinjak pojmova.



Hrvatski na maturi

http://matura.ihjj.hr/
Vrlo koristan portal za maturante i sve one koji će to postati. Hrvatski na maturi pokriva sva pravopisna, morfološka, sintaktička i leksička pravila koje će vam trebati na ispitu zrelosti, ali i kasnije u životu.



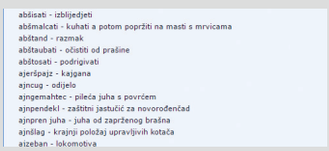
Portal hrvatske rječničke baštine

http://crodip.ffzg.hr/default.aspx
Portal za hardcore jezičare na kojem možete pretraživati stare hrvatske rječnike (i riječi). Uz pomoć ovih rječnika bez problema ćete razumjeti Juditu.



Purgerski rječnik

http://blog.dnevnik.hr/purgerskirjecnik/
Entuzijast Ivan Brnčić godinama je zapisivao stare purgerske izraze, a plod tog rada možete pronaći na ovoj stranici. Purgerski rječnik sadrži više od 700 pojmova.


abšisati - izblijedjeti
abšmalcati - kuhati a potom popržiti na masti s mrvicama
abštand - razmak
abštaubati - očistiti od prašine
abštosati - podrigivati
ajeršpajz - kajgana
ajncug - odijelo
ajngemahtec - pileća juha s povrćem
ajnpendekl - zaštitni jastučić za novorođenčad
ajnpren juha - juha od zaprženog brašna
ajnšlag - krajnji položaj upravljivih kotača
aizeban - lokomotiva

Dalmatinski rječnik

http://www.gimnazija-jbarakovica-zd.skole.hr/dalmatinski%20rjecnik
Na internetu ćete naći nekoliko dalmatinskih rječnika, općenitih ili karakterističnih za određeno područje (Omiš, Brela), a nama se najviše svidio glosar Gimnazije Jurja Barakovića iz Zadra.

