

POJMOVNIK E-LEKSIKOGRAFIJE

1. Uvod

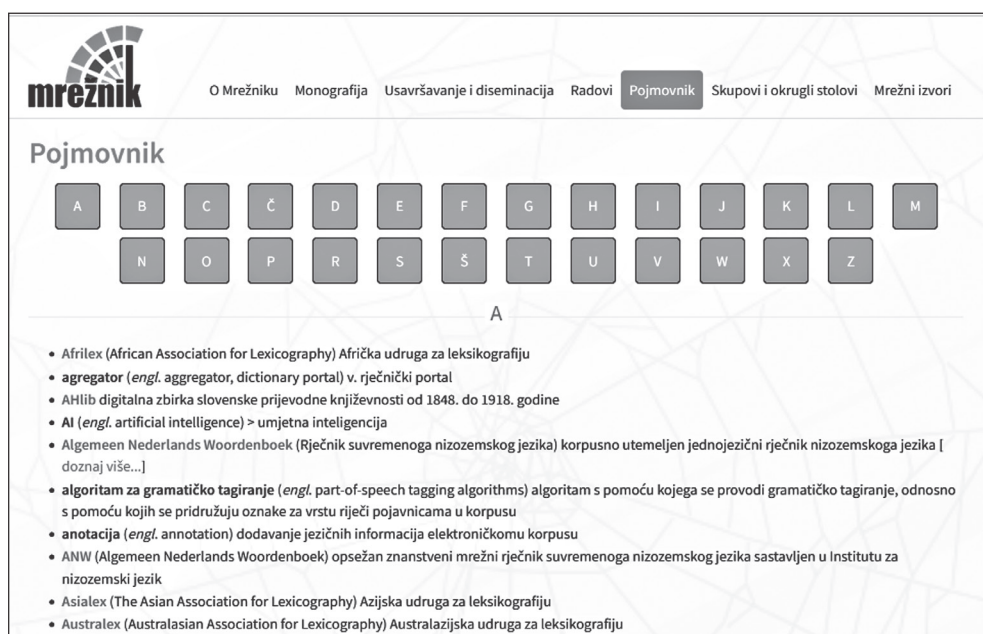
U početnome razdoblju provedbe projekta *Hrvatski mrežni rječnik – Mrežnik* pokazala se potreba da se neki nazivi koji se u hrvatskome jeziku upotrebljavaju u području e-leksikografije u prvome redu objasne te potom i urede u skladu s terminološkim pravilima, odnosno da se u nizu istoznačnih naziva jedan preporuči za uporabu odnosno uspostavi naziv kojega dotad u hrvatskome nije bilo. Kao i u mnogim novijim područjima, i u e-leksikografiji veoma se često, pa i u stručnoj literaturi, upotrebljavaju engleski nazivi. E-leksikografsko je nazivlje naglašeno interdisciplinarno te uključuje mnoge korpusnojezikoslovne nazive, ali i nazive nekih područja koja se ne povezuju s tradicionalnom leksikografijom, npr. igrifikacijsko nazivlje. Osim naziva *Pojmovnik* uključuje i imena relevantna za e-leksikografiju

2. *Pojmovnik e-leksikografije i Jena*

Pojmovnik e-leksikografije izrađen je u sklopu projekta *Hrvatski mrežni rječnik – Mrežnik* i u suradnji s projektom *Hrvatsko jezikoslovno nazivlje – Jena*.³⁶¹ Vrela za crpenje naziva radovi su iz područja e-leksikografije te *Jezikoslovni korpus* sastavljen u okviru projekta *Jena* (Marković, Mihaljević i Mihaljević 2020., Mihaljević i Marković 2023.). Suradnici na projektu bilježili su i nazive te imena koje su čuli na e-leksikografskim konferencijama te su i ti nazivi i imena zabilježeni u *Pojmovniku e-leksikografije*.

Dostupan je na stranicama projekta na adresi <http://ihjj.hr/mreznik/page/pojmovnik/6/>, vidi 1. sliku.

³⁶¹ Više o projektu *Hrvatsko jezikoslovno nazivlje – Jena* vidi na projektnim stranicama jena.jezik.hr i u poglavlju *Jezikoslovno nazivlje*.



1. slika: Pojmovnik e-leksikografije

Pojmovnik uključuje e-leksikografske nazive i imena institucija, tvrtka, inicijativa, programa i projekata relevantna za e-leksikografiju i iz njega se većina naziva (ne imena) preuzima u bazu *Jene*. U *Pojmovnik* nije uključeno tradicionalno leksikografsko nazivlje te u njemu nema naziva poput *natuknica*, *rječnički članak*, *rječnik* itd. Obrade naziva iz *Pojmovnika* prilagođuju se potrebama projekta *Jena*.³⁶² Na 2. i 3. slici prikazana je *natuknica regularni izraz* u *Pojmovniku e-leksikografije*. Nazivlje iz *Pojmovnika* preuzeto je u bazu *Jene* i njoj prilagođeno (vidi poglavlje *Jezikoslovno nazivlje*) te je iz baze *Jene* građa izvezena u bazu *Hrvatsko strukovno nazivlje* – *Struna*, vidi 3. sliku.

³⁶² Više o obradi jezikoslovnih naziva u projektu *Jena* vidi u Mihaljević, Hudeček, Jozić (ur.) 2023.

- **regex** v. **regularni izraz**
- **regularni izraz** (*engl.* regular expression) zadani niz znakova čija se kombinacija može upotrijebiti za pronalaženje određenih dijelova teksta, izraz koji se uspostavlja za pretraživanje korpusa s pomoću alata SketchEngine i NoSketchEngine za traženje ciljanih gramatičkih ili leksičkih uzoraka (opis regularnih izraza dostupan je na <https://sketchengine.co.uk/documentation/corpus-querying/>). Regularni izrazi iz SketchEngineova kalendara za 2018.:

all of these queries produce the same result, i.e. Daffodil including variants of various spelling

[word = "Daff?odi"]

[word = "Daf{1,2}odi"]

[word = "Daffodil|Dafodil"]

The queries will find the word Daffodil including cases when written incorrectly with one letter

1 question: Does the preceding character optional? i.e. whether Daffodil or daffodil

1 question: Is D repeated 1 and 2 repetitions of the preceding character? i.e. Daffodil or Dafodil

2. slika: Obrada natuknice *regularni izraz* u *Pojmovniku e-leksikografije*

© 2011 Institut za hrvatski jezik i jezikoslovje

Kopački rit, Osijek © Vlado Kos / Cropix

regularni izraz

definicija | zadani niz znakova čija se kombinacija može upotrijebiti za pronalaženje određenih dijelova teksta

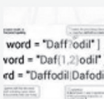
istovrijednice |

- engleski:** regular expression, regex
- njemački:** regulärer Ausdruck
- francuski:** expression régulière
- ruski:** регулярное выражение
- talijanski:** espressione regolare
- češki:** regulární výraz
- švedski:** reguljärt uttryck

poveznica | <https://sketchengine.co.uk/documentation/corpus-querying/>

napomena | Regularni izraz uspostavlja se za pretraživanje korpusa s pomoću alata SketchEngine i NoSketchEngine za traženje ciljanih gramatičkih ili leksičkih uzoraka.

privitak |



razredba |

- polje:** filologija
- grana:** opće jezikoslovje (lingvistika)
- projekt:** Hrvatsko jezikoslovno nazivlje (JENA)

3. slika: Obrada natuknice *regularni izraz* u *Struni* (<http://struna.ihjj.hr/naziv/regularni-izraz/51813/#naziv>)

Uz natuknicu u *Pojmovniku* donosi se engleska istovrijednica (u zagradama) te kratka definicija leksikonskoga tipa. Uz natuknicu se katkad nalazi i slika. Kao natuknice bilježe se svi sinonimni nazivi, a s pomoću uputnice *v.* (vidi) upućuje se na naziv pod kojim se nalazi obrada. Status preporučenoga naziva (naziva pod kojim se nalazi obrada), kao i u *Jeni*, naziv dobiva na temelju analize provedene s pomoću terminoloških načela (Hudeček i Mihaljević 2012., 2020., Mihaljvić, Hudeček, Jozić (ur.) 2023., vidi i poglavlje *Jezikoslovno nazivlje*).

U *Jeni* se donosi kraća rječnička definicija, navodi vrelo definicije, dodatni podatci daju se u napomeni. Natuknica je uvijek preporučeni naziv, a sinonimni se nazivi bilježe u rubrikama dopušteni naziv, nepreporučeni naziv, žargonski naziv, zastarjelica. Donosi se engleska, njemačka, francuska, talijanska, ruska, češka i švedska istovrijednica.

Specifičnost je e-leksikografskoga nazivlja da se u hrvatskome jeziku za određeni pojam veoma često upotrebljava (i) engleski naziv. Jedno je (i prvo) od terminoloških načela da domaći naziv ima prednost pred stranim. Na temelju toga načela prednost se u *Pojmovniku e-leksikografije* npr. daje hrvatskim nazivima *crpenje podataka* pred *data mining*, *gramatika skica* pred *Sketch Grammar*, *izbornik* pred *meni* i *menu*, *jezični izvori* pred *jezični resursi*, *označivanje pogrešaka* pred *error tagging*, *pametna leksikografija* pred *smart lexicography*, *pojavnica* pred *token*, *poveznica* pred *link*, *povezani podatci* pred *linked data*, *preklapanje n-grama* pred *n-gram overlap*, *pretraživač* pred *search engine*, *rastući rječnik* pred *growing dictionary*, *rječnički portal* pred *agregator*, *skice riječi* pred *Word Sketches*, *unutarrječničko povezivanje* pred *cross-referencing*. Prednost se daje i nazivu *sinonimni skup* pred prevedenicom *sinskup* (prema engleskome *synset*).

Međutim, neki su hrvatski nazivi nastali prilagodbom engleskih naziva te su u e-leksikografskome nazivlju posve prihvaćeni. To su npr. nazivi *parsiranje*, *parsirati*, *parser*, *tag*, *tager*, *tagiranje*, *tagirati*.

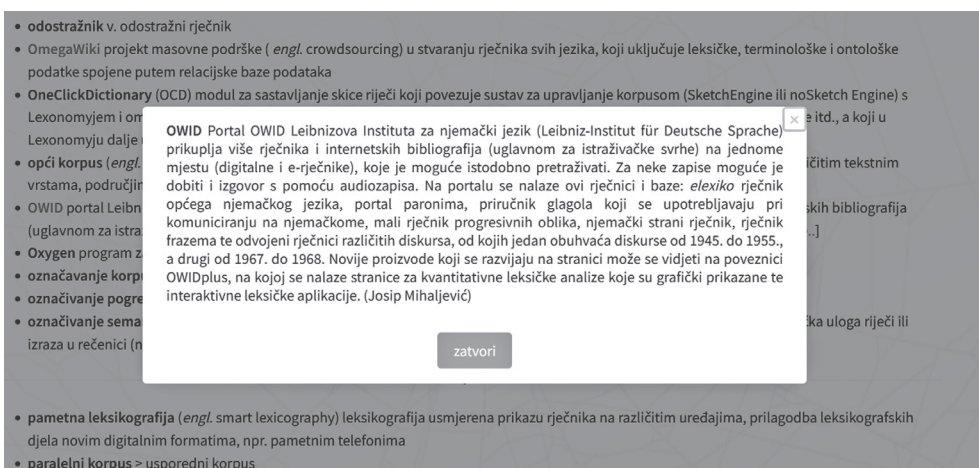
Prednost se daje i hrvatskomu nazivu pred internacionalizmom, npr. nazivu *čestoća* pred *frekvencija* i *frekventnost*.

Prednost se daje i nazivu usklađenu s normama standardnoga jezika pred onim koji to nije, npr. nazivu *govorni korpus* pred nazivom *govoreni korpus*, nazivu *konkordancija* pred nazivima *konkordanca* i *konkordansa*, nazivu *parsemsko stablo* pred nazivom *stablo parsema*.

Među jednako normativno prihvatljivim nazivima od kojih nijedan nije strani, prednost se daje bolje prihvaćenomu, npr. nazivu *odostražni rječnik* pred nazivom *odostražnik*.

Prednost se daje i manje metaforičnomu nazivu, npr. nazivu *crpenje podataka* pred nazivom *rudarenje podataka*.

Natuknica u *Pojmovniku* u pravilu je poveznica koja vodi na mrežne stranice projekta, institucije, portala, rječnika itd. Poveznice se nalaze i u tekstu članka. Također, neke se natuknice u *Pojmovniku*, i kad je riječ o nazivima i kad je riječ o imenima, proširuju opširnijim podacima dostupnim putem poveznice koja se nalazi na kraju primarnoga članka, vidi 4. sliku.



4. slika: Druga razina obrade natuknice OWID u *Pojmovniku e-leksikografije*

Zaključak

Pojmovnik e-leksikografije pokrenut je kako bi se uskladilo nazivlje kojim se u svojim radovima koriste suradnici na projektu *Mrežnik*. S vremenom je prerastao u pojmovnik u kojemu se može pronaći podatak o značenju kojega e-leksikografskoga naziva, ali i podatak o institucijama, tvrtkama, inicijativama, programima, projektima itd. povezanim s e-leksikografijom, ali i važnim e-leksikografskim konferencijama. U suradnji s projektom *Hrvatsko jezikoslovno nazivlje – Jena* to se nazivlje uređuje i dopunjuje.

Izvori i literatura

- Hudeček, Lana; Mihaljević, Milica. 2012. *Terminološki priručnik*. Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- Hudeček, Lana; Mihaljević, Milica. 2019. *Croatian Web Dictionary – Mrežnik vs. Croatian Linguistic Terminology – Jena. INFuture 2019: Knowledge in the Digital Age*. Ur. Bago, Petra; Hebrang Grgić, Ivana; Ivanjko, Tomislav; Juričić, Vedran; Miklošević, Željka; Stublić, Helena. Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences. Zagreb. 22–31.
- Hudeček, Lana; Mihaljević, Milica. 2023. Leksikološko, leksikografsko i e-leksikografsko nazivlje. *Hrvatsko jezikoslovno nazivlje*. Ur. Mihaljević, Milica; Hudeček, Lana; Jozić, Željko. Hrvatsko jezikoslovno nazivlje. Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- Marković, Mario; Mihaljević, Josip; Mihaljević, Milica. 2020. Kako pronaći jezikoslovni naziv. *Hrvatski jezik* 7/1. 18–22.

-
- Mihaljević, Josip; Marković, Mario. 2023. Jezikoslovni korpus. *Hrvatsko jezikoslovno nazivlje*. Ur. Mihaljević, Milica; Hudeček, Lana; Jozić, Željko. Hrvatsko jezikoslovno nazivlje. Institut za hrvatski jezik i jezikoslovlje. Zagreb.
- Mihaljević, Milica; Hudeček, Lana; Jozić, Željko. 2023. *Hrvatsko jezikoslovno nazivlje*. Institut za hrvatski jezik i jezikoslovlje. Zagreb.