

Računalni programi i programske usluge za otkrivanje plagiranja u znanosti i obrazovanju

Stella Lampret¹, Vanja Pupovac¹, Mladen Petrovečki^{1,2}

¹Katedra za medicinsku informatiku Medicinskog fakulteta Sveučilišta u Rijeci

²Klinički zavod za laboratorijsku dijagnostiku, Klinička bolnica Dubrava, Zagreb

SAŽETAK Razvoj računalnih tehnologija omogućuje jednostavan pristup informacijama, ali i njihovu zlouporabu. Problem se posebno ističe u djelatnostima poput znanosti, informatike ili književnosti. Plagiranjem se smatra preuzimanje tuđih ideja ili teksta bez navođenja izvora, i prikazivanje preuzetoga vlastitim. Kao odgovor na plagiranje, 80-ih godina prošlog stoljeća počinju se razvijati računalni programi za otkrivanje plagiranja u tekstovima. Danas postoje programi koji otkrivaju akademsko i znanstveno plagiranje. *Turnitin* i *SafeAssign* su najučinkovitiji kod otkrivanja akademskog, a *CrossCheck* i *eTBLAST* kod znanstvenog plagiranja. Programi se sastoje od algoritma za otkrivanje sličnosti, pripadajuće baze podataka i mrežnih stranica s kojima uspoređuju promatrani rad. Unatoč brojnim prednostima, računalni programi za otkrivanje plagiranja još uvijek imaju značajne nedostatke koje treba nadvladati.

KLJUČNE RIJEČI algoritmi; plagiranje; programska potpora; znanstveno nepoštenje

Napredak informacijsko-komunikacijske tehnologije (IKT) danas nudi brojne mogućnosti. Informacije nikad nisu bile dostupnije, ali osim što se time značajno povećava njihova korisnost, pojednostavljuje se i mogućnost njihove zlouporabe. Naredbe preslikaj (kopiraj, engl. *copy*) i zalijepi (engl. *paste*) omogućuju brzo i jednostavno preuzimanje velike količine teksta, što olakšava plagiranje. Plagiranje podrazumijeva preuzimanje tuđih ideja ili teksta bez navođenja izvora, predstavljajući ono što je neovlašteno preuzeto kao vlastito intelektualno vlasništvo. Osim plagiranja, često spominjano i jednako štetno jest samoplagiranje, a podrazumijeva preuzimanje već objavljenih vlastitih ideja ili preuzimanje vlastitog objavljenog teksta, predstavljajući ga kao novog i izvornog.¹ U djelatnostima poput znanosti, informatike ili književnosti, gdje su ideja i tekst ključni rezultati rada i pokazatelji napretka pojedinca ili skupine, plagiranje i samoplagiranje posebno su štetni, i izjednačeni su s krađom intelektualnog vlasništva i varanjem. Tema ovoga rada je prikazati postojeće računalne programe i programske usluge za otkrivanje plagiranja koje se koriste u znanosti i obrazovanju.

dence based medicine), vjerodostojnost znanstvenog rada je temeljni uvjet, zbog toga jer se taj sustav zasniva na prikupljanju i obradi rezultata znanstvenih istraživanja. Njihovim objedinjenjem oblikuje se razvoj medicinske prakse u obliku preporuka i smjernica za liječenje temeljeno na znanstvenim dokazima. Poštenje je, stoga, u biomedicinskim istraživanjima posebno važno jer prijevare i lažni rezultati mogu izravno utjecati na zdravlje i živote ljudi. Da bismo pridonijeli održanju i povećanju kvalitete u obrazovanju i znanosti, potrebno je proučavati pojave različitih oblika znanstvenog nepoštenja, na njih upozoravati i provoditi preventivne akcije. Proučavanju i istraživanju nečestitosti nije jedini cilj razotkriti nepoštenje u interesu znanosti, već na temelju dobivenih rezultata oblikovati upozorenja i smjernice u obrazovanju studenata i znanstvenika. Akademsko i znanstvena zajednica neodvojive su, studenti tijekom akademske izobrazbe, stječu znanja i vještine, ali i prihvaćaju stavove i ponašanje učitelja. Znanstvenici trebaju prihvaćati i isticati najviše etičke norme jer se na rezultatima njihova rada temelji dobrobit i povjerenje šire zajednice.²

ZAŠTO JE POŠTENJE U ZNANOSTI BITNO?

Znanstvena publikacija predstavlja vrhunac znanstvenikova rada, a nakon objave, rad postaje dio sveukupne ljudske znanosti i trajno je podložan provjeri i kritici. Razvojem medicine temeljene na dokazima (engl. *evidence*

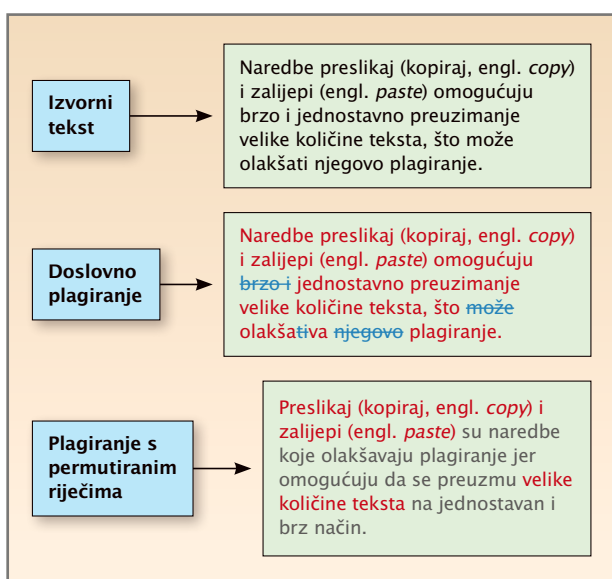
RAZVOJ RAČUNALNIH PROGRAMA ZA OTKRIVANJE PLAGIRANJA

Računalni programi za otkrivanje plagiranja počeli su se razvijati 80-ih godina prošlog stoljeća kao odgovor računalnih programera na neovlašteno preuzimanje izvornih programskih kodova. Tada najpoznatiji programi bili

su YAP (Michael Wise, *University of Sydney*, Australija), MOSS (Alex Aiken, *Berkeley*, California, SAD), SIM (Gitchell D i Tran N, *Wichita State University*, SAD) i *Plague System* (Michael Wise, *University of Sydney*, Australija).³⁻⁵ Predstavljali su dobru osnovu za razvijanje programa za otkrivanje plagiranja teksta. Programi za otkrivanje plagiranja tekstova koriste tehnike internet-skih pretraživača koji „dohvaćaju“ slične informacije i potom ih klasificiraju od identičnih ili najsličnijih do onih manje sličnih. Klasifikacija sličnosti u početku se obavljala tako da se pronalaze identični nizovi određenog broja riječi u slijedu (najčešće njih 6) pa se tekstovi s više takvih nizova smatraju sličnijima izvorniku od onih s manje pronađenih istih nizova. Nedostatak takvog po-

stupka je prepoznavanje isključivo doslovnog plagiranja ili tzv. plagiranja „od riječi do riječi“, ali ne i plagiranja s permutiranim riječima (slika 1). Kao moguće rješenje tog problema pojavila se tehnika digitalnog otiska (engl. *digital fingerprint*) koja pronalazi identične nizove određenog broja znakova (ne riječi!) u slijedu (odbacujući suvišne informacije poput razmaka) odabranih iz veličine teksta k , a izražene brojem znakova. Takva tehnika omogućuje otkrivanje permutiranih fraza, a ponekad i istoznačnica (sinonima) uporabljenih u svrhu prikrivanja plagiranja. Tehniku digitalnog otiska koristi većina današnjih računalnih rješenja za otkrivanje plagiranja. Njihova uspješnost ovisit će o duljini niza znakova (n) i veličini teksta (k) iz koje je odabran niz. Što je veći n , a manji k , algoritam je uspješniji.⁶ S obzirom da su izvorni kodovi današnjih programa zaštićeni, o veličini ključnih parametara, a time i o njihovoj uspješnosti, može se samo nagađati. Pregled značajki najkorištenijih programa prikazana je u tablici 1.

SLIKA 1. Doslovno plagiranje i plagiranje s permutiranim riječima.



Crvenom je označen doslovno plagiran tekst koji će programi za otkrivanje plagiranja prepoznati, a plavom su označene riječi koje su ispuštene prilikom doslovnog plagiranja. Sivom je obojen tekst koji je plagiran, ali koji zbog permutiranja plagiranog teksta jednostavni računalni programi neće prepoznati

OPIS RADA RAČUNALNIH PROGRAMA ZA OTKRIVANJE PLAGIRANJA

Rad programa može se pojednostavljeno prikazati u tri koraka (slika 2). U prvom koraku učitava se tekst sumnjiv na plagiranje i iz njega se oblikuje digitalni otisak.

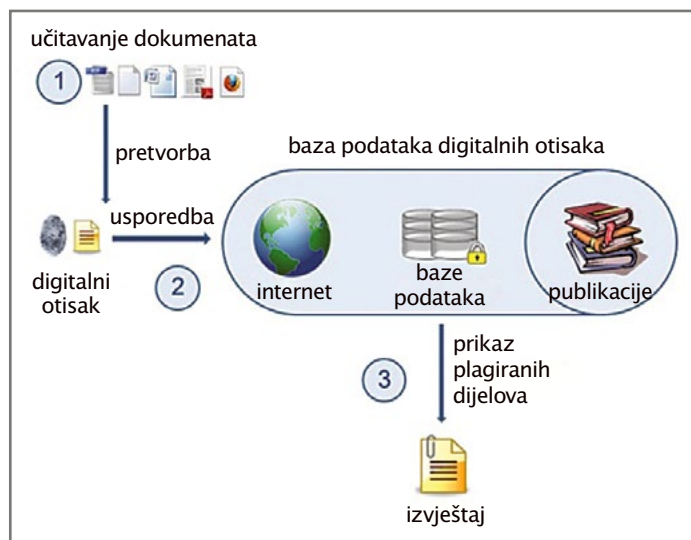
Digitalni otisak ispitivanog teksta uspoređuje se tada s digitalnim otiscima ostalih tekstova iz baza podataka ili s mrežnih stranica. Izbor baza podataka za usporedbu ovisi o programu koji se koristi, no otkrivanje plagiranja bit će uspješnije ukoliko su baze podataka velike i ukoliko imaju relevantne tekstove, npr. studentske radove ako je riječ o akademskom plagiranju. Za uspješno otkrivanje plagiranja temeljenog na izvorima s mrežnih stranica bitna je dohvatljivost informacija. Ona ovisi o tehnici za pretraživanje informacija, točnije o internetskoj tražilici koju koristi program. Internetski pretraživač *Google* pokazao se kao uspješan alat za otkrivanje plagijata.⁷ *Googleova* uspješnost potvrđena je istraživanjem Maurera i Zake, koji su izradili eksperimentalni program za otkrivanje plagiranja temeljen na *Googleovom* pretraživaču – *Benchmark Plagiarism Tool (BPT)*, i ispitali njegovu uspješnost u usporedbi s *TurnitInom* i *SafeAssignom*. *BPT* se pokazao uspješnijim u otkrivanju plagijata temeljenih na izvorima s mrežnih stranica.⁸

Treći, završni korak jest stvaranje izvješća koje najčešće sadrži:

- udio (postotak) sličnosti između dva teksta,
- prikaz ispitivanog teksta s istaknutim plagiranim dijelovima i
- mogućnost pronalaženja izvornih materijala putem mrežnih poveznica (tablica 1).

Nije rijetkost da je plagirani tekst nastao na temelju više različitih izvora te se u tom slučaju stvara i popis izvora iz kojih je tekst preuzet, najčešće označen različitim bojama s obzirom na različite izvornike (tzv. kolažno plagiranje, engl. *patchwork*). Opisana detaljna izvješća sadrže

SLIKA 2. Opis rada računalnih programa za plagiranje predstavljen u tri koraka



TABLICA 1. Pregled odabranih značajki najčešće korištenih programa za otkrivanje plagiranja (podaci preuzeti s mrežnih stranica proizvođača programa)

Alati za otkrivanje plagiranja/ značajke							
Mrežna adresa	http://www.turnitin.com/en_us/home	http://www.crossref.org/crosscheck/index.html	http://etest.vbi.vt.edu/etblast3	http://safeassignment.com	http://plagiarism.bloomfieldmedia.com/z-wordpress/	http://www.checkforplagiarism.net/	http://www.compilatio.net/en/
Proizvođač	iParadigms	iParadigms	Harold Garner	SafeAssign Blackboard	Lou A. Bloomfield	CheckForPlagiarism.net	Compilatio SAS
Vrsta	program	mrežna usluga	mrežna usluga	program	program	mrežna usluga	mrežna usluga
Plaćanje	da	da + CrossRef članstvo	ne	da	ne	da	da
Namjena	akademsko obrazovne institucije (profesori)	znanstvene i izdavačke kuće	znanstvene institucije, privatna namjena	akademsko obrazovne institucije (studenti)	akademsko obrazovne institucije, privatna namjena	akademsko obrazovne i znanstvene institucije, privatna namjena	akademsko obrazovne institucije (studenti i profesori)
Izvori s kojima se uspoređuje zadani tekst	mrežne stranice, baza pregledanih radova, ProQuest baza	mrežne stranice, baza pregledanih radova, baze: CrossCheck, InfoTrac, OneFile, Emerald Journals, ABC-CLIO	mrežne stranice, baze: Medline, NIH Crisp, arXiv, NASA	mrežne stranice, baza pregledanih radova, baze: ProQuest ABI/Inform	unaprijed određena mrežna stranica, unaprijed zadan tekst na lokalnom računalu	mrežne stranice, baze: Lexus-Nexus, Blackwell-Synergy, Athens, Emerald, JSTOR, Harvard Journals, OptiQuest	mrežne stranice, korisnička baza pohranjenih radova
Pohrana pregledanog rada	da	da	ne	po izboru	ne	ne	po izboru
Stupanj razotkrivanja	doslovno prepisivanje, permutirane riječi	doslovno prepisivanje, permutirane riječi, sinonimi	doslovno prepisivanje, permutirane riječi	doslovno prepisivanje, permutirane riječi	doslovno prepisivanje	doslovno prepisivanje, permutirane riječi, sinonimi	doslovno prepisivanje, permutirane riječi
Izvrješće	postotak sličnosti, prikaz plagiranih dijelova, poveznica na originalan rad	postotak sličnosti, prikaz plagiranih dijelova, poveznica na originalan rad	postotak sličnosti, prikaz plagiranih dijelova, poveznica na originalan rad	postotak sličnosti, prikaz plagiranih dijelova, poveznica na originalan rad	prikaz plagiranih dijelova	postotak sličnosti, prikaz plagiranih dijelova, poveznica na originalan rad	postotak sličnosti, prikaz plagiranih dijelova, poveznica na originalan rad
Podržani oblici datoteka	doc(x), wpd, rtf, htm(l), txt, ps, eps, pdf	doc(x), wpd, rtf, htm(l), txt, ps, eps, pdf, zip	txt, html	doc(x), txt, rtf, pdf, html	doc(x),htm(l), pdf, txt	doc(x), wpd, rtf, htm(l), txt, ps, pdf	txt, pdf, rtf, doc(x), odt, xls, xlsx, ppt(x), htm(l), php, asp
Računalna platforma	MS Windows, Mac OS, Blackboard Moodle, Angel	internet	internet	MS Windows, Mac OS, Blackboard Moodle, Sakai	MS Windows	MS Windows, Mac OS	internet
Upute i često postavljena pitanja	da	da	da	da	da	da	da

programi i programske usluge *TurnitIn* (*iParadigme LLC, Berkeley, California, SAD*), *CrossCheck* (*CrossCheck powered by iThenticate, iParadigme LLC, Berkeley, California, SAD*), *Compilatio.net* (*Compilatio SAS, Saint-Félix, Francuska*), *SafeAssign* (*Blackboard Inc, Washington, SAD*).⁹⁻¹³

AKADEMSKI I ZNANSTVENI PROGRAMI ZA OTKRIVANJE PLAGIRANJA

Sadržaj baza podataka s digitalnim otiscima tekstova s kojima se uspoređuje ispitivani tekst određuje vrstu i specifičnu uporabu pojedinog programa za otkrivanje plagiranja. Programi koji sadrže baze podataka studentskih eseja i seminarskih radova otkrivaju akademsko, a oni koji sadrže baze podataka znanstvenih radova, znanstveno plagiranje. *TurnitIn* i *SafeAssign* su najčešće korišteni programi za otkrivanje akademskog plagiranja zato što imaju najveću bazu podataka radova s kojom uspoređuju tekst sumnjiv na plagiranje.⁸

Svi radovi koji se pregledaju pomoću *TurnitIna* automatski se pohranjuju u *TurnitIn* bazu podataka te postaju materijal za daljnju usporedbu. *SafeAssign* se također često koristi i ostavlja korisniku na izbor želi li pohraniti svoj rad u bazu podataka ili ne. *SafeAssign* i *TurnitIn* su podržani od sustava za obrazovanje na daljinu poput *Angel Blackboard* (*Blackboard Inc, Washington, SAD*), *Moodle* (*Moodle Pty Ltd, Perth, Western Australija*) i *WebCT* (*Murray Goldberg, University of British Columbia, Vancouver, Kanada*) što ih čini upotrebljivima za kontrolu studentskih radova u okruženju učenja na daljinu.^{9,12,13}

Za otkrivanje znanstvenog plagiranja najučinkovitija je mrežna usluga *CrossCheck*.¹⁴ Problem otkrivanja znanstvenog plagiranja je u ograničenoj dostupnosti znanstvenih radova koji su zaštićeni sustavom pretplate pa ih uobičajeni programi za otkrivanje plagiranja ne mogu pretraživati. *CrossCheck* sadrži bazu podataka znanstvenih radova koju je oblikovala udruga izdavača i urednika znanstvenih časopisa *CrossRef*. Članovi *CrossRef* udruge su ugledni znanstveni izdavači poput *American Society of Neuroradiology, BMJ Publishing Group, Elsevier, IEEE, Nature Publishing Group, Oxford University Press, Sage, Wiley Blackwell* i drugi koji ustupaju radove bazi podataka *CrossChecka*, te imaju mogućnost korištenja *CrossCheck* usluge za otkrivanje plagiranja u svojim izdanjima.^{1,10,15}

Besplatna mrežna usluga za otkrivanje znanstvenog plagiranja je *eTBLAST* (*Harold Garner, Virginia Bioinformatics Institute, SAD*). Razlog tome je besplatan znanstveni sadržaj koji se pretražuje – sažeci znanstvenih radova sadržani u sekundarnim bibliografskim bazama podataka, kao npr., *Medline, NIH Crisp, Institute of Physics* i *NASA*. Autori *eTBLAST* usluge izradili su *Déja vu* bazu podataka dvostrukih (jednakih) i sličnih sažetaka u koju pohranjuju rezultate pretrage *eTBLAST-om* koji upućuju na moguće plagiranje. Administratori *Déja vu* baze podataka provjeravaju sve zaprimljene radove, te odvajaju plagijate od onih koji to nisu, ali su prepoznati uslugom *eTBLAST*. (npr. pisma uredništvu,

uvodnike, kongresna priopćenja i sl.).^{16,17} Nažalost, broj pregledanih radova (172) u odnosu na broj zaprimljenih (79.383) još uvijek je jako malen.

NEDOSTACI PROGRAMA ZA OTKRIVANJE PLAGIRANJA

Ako je program za otkrivanje plagiranja pronašao veliki udio sličnosti između dva teksta, ne mora se nužno raditi o plagiranju, kao što i računalni program neće nužno otkriti plagiranje ako je ono prisutno. Te pretpostavke mora imati na umu svatko tko rabi programe za otkrivanje plagiranja. Da bi pronađeno plagiranje bilo nedvojbeno potvrđeno, nakon računalne obrade teksta, uvijek mora uslijediti i čovjekova provjera zbog nesavršenosti računalnih programa.

Niti jedan od programa za otkrivanje plagiranja ne može otkriti plagiranje ideja, odnosno složeno parafraziranje. Moguće je otkriti tek promjenu redoslijeda riječi u rečenici, ubačen pokoji sinonim, ali ne i puno više od toga.^{6,8} Poseban je problem plagiranje iz nedigitaliziranih izvora tekstualnih sadržaja jer su nedohvatljivi računalnim programima za otkrivanje plagiranja te će takvo plagiranje iz nedigitaliziranih izvora ostati neotkriveno ili se otkriti slučajno. Jedno od mogućih rješenja za opisane probleme jest algoritam koji otkriva plagiranje temeljem sličnosti i razlika u stilu pisanja. Takvi algoritmi uspoređuju stil pisanja unutar teksta sumnjivog na plagiranje. Uoči li se razlika u stilu između dva dijela teksta, može se pretpostaviti da je riječ o plagiranju. Pri tome za otkrivanje plagiranja nije potrebno pronalaziti moguće izvore iz kojih se tekst preuzimao, već je dovoljan sami tekst koji ispitujemo na plagiranje, pa je to intrinzično otkrivanje plagiranja.^{18,19} Ovakvi algoritmi već postoje, ali su još uvijek niske učinkovitosti i njihovo se usavršavanje tek očekuje. Sljedeći nedostatak je nemogućnost pretraživanja priloga u tekstu, slika i tablica koje većina programa zaobilazi prilikom raščlambe, a bitni su jer mogu sadržavati plagirane dijelove ili biti prepisani i preslikani u cijelosti.⁶ U znanstvenim radovima tablice prikazuju rezultate istraživanja te njihovo neovlašteno preuzimanje povlači teške optužbe za znanstveno nepoštenje.

I na kraju, ostaje neriješen problem plagiranja na različitim jezicima. Još uvijek ne postoje učinkoviti računalni programi koji bi prepoznali plagiranje u prijevodima s jezika na jezik. To znači da će plagirani rad, koji sadrži tekst izvornika preveden na drugi jezik pri čemu se predstavlja kao izvornik, vjerojatno ostati neotkriven. Kako bi doskočili i tom obliku plagiranja, programi za otkrivanje plagiranja bi u svoj algoritam trebali ugraditi i programe za prevođenje. *TurnitIn* je krajem siječnja 2012. godine pokrenuo testnu verziju programa koji otkriva plagiranje višejezičnom usporedbom tekstova.⁹

ZAKLJUČAK

Poštenje u znanosti izuzetno je bitno zato jer prijevara i lažni rezultati izravno mogu utjecati na ljudski život. Računalnim programima za otkrivanje plagiranja

znanstvena zajednica si pokušava olakšati borbu protiv jednog od oblika znanstvenog nepoštenja. Otkrivanje plagiranja u znanosti i obrazovanju danas je gotovo nezamislivo bez računalnih programa. Razlog tome je prvenstveno velika količina informacija u znanosti, koje, s jedne strane, moraju biti provjerene na plagiranje, a, s druge se, koriste kao materija za usporedbu teksta sumnjivog na plagiranje. Današnji računalni programi vrlo brzo mogu usporediti sve dostupne relevantne tekstove i dosta učinkovito izdvojiti tekst sumnjiv na plagiranje, jasno prikazujući plagirane dijelove i izvore iz kojih su plagirani. O važnosti i utjecaju plagiranih dijelova uvijek mora prosuditi čovjek, stručnjak koji je upoznat s

pregledanom materijom. Da bi prevenirali plagiranje i bili sigurni u izvornost teksta potrebno je koristiti računalni program radi sveobuhvatnosti mogućih izvora plagiranja i potom samostalno procijeniti ispravnost rezultata zbog postojećih nedostataka programa. Razvatak i napredak računalnih programa ukazuje na brz pronalazak rješenja za sve dosadašnje nedostatke stoga je moguće očekivati da će se usavršiti do tog stupnja da budu dostatni u borbi protiv plagiranja.

Zahvaljujemo doc. dr. Lidiji Bilić-Zulle na komentarima nakon kritičkoga čitanja članka

Using plagiarism detection software and software support in science and education

SUMMARY *The development of computer technologies allows not only easy access to information, but also its abuse. The problem is particularly evident in areas such as science, computer science and literature. Plagiarism is defined as appropriation of other's ideas or text without acknowledging the source and claiming to be one's own. During the 1980s, the development of plagiarism detection software came as an answer to the plagiarism issues. Today, there are two major groups of plagiarism detection software: one for detection of academic and the other for detection of scientific plagiarism. TurnItIn and SafeAssign are representatives of the first group, and eTBLAST and CrossCheck are representatives of the latter. Plagiarism detection software consists of an algorithm that detects similarities and searches related databases and web pages for sources of potential plagiarism. Although there are numerous advantages to plagiarism detection software, there are still shortcomings that have to be overcome.*

KEY WORDS *algorithms; plagiarism; scientific misconduct; software*

LITERATURA

1. Baždarić K, Pupovac V, Bilić-Zulle L, Petrovečki M. Plagiranje kao povreda znanstvene i akademske čestitosti. *Medicina*. 2009;45(2):108-17.
2. Bilić-Zulle L. Znanstvena čestitost – temelj postojanja i razvoja znanosti. *Biochem Med*. 2007;17(2):143-150.
3. Wise MJ. Detection of similarities in student programs: YAP'ing may be preferable to Plague'ing. Association for Computing Machinery, Special Interest Group on Computer Science Education Bulletin 1992;24(1):268-71.
4. Gitchell D, Tran N. Sim: a utility for detecting similarity in computer programs. Thirtieth SIGCSE Technical Symposium on Computer Science Education. New Orleans, LA. 1999.
5. Bowyer KW, Hall LOH. Experience using MOSS to detect cheating on programming assignments. Proceedings of IEEE Computer Society Conference on Frontiers in Education. San Juan, Puerto Rico, 10-13 studeni 1999.
6. Introna LD, Hayes N. On sociomaterial imbrica-
tions: What plagiarism detection systems reveal and why it matters. *Information and Organization* 2011; 21(2):107-22.
7. Farrokhi F, Khoddam M, Zamani N. Plagiarism in Submissions to an English-language medical journal: do authors have problems only with language? The Proceedings of the 4th Regional Conference on Medical Journals in the Eastern Mediterranean Region. Manama, Bahrain, 5-7 studeni 2008.
8. Maurer H, Zaka B. Plagiarism – A Problem And How To Fight It. U: Montgomerie C, Seale J, ur. Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications. Vancouver: Association for the Advancement of Computing in Education; 2007, str. 4451-8.
9. Turnitin. Dostupno s: http://www.turnitin.com/en_us/home Pristupljeno 22. ožujka 2012.
10. CrossCheck. Dostupno s: <http://www.crossref.org/crosscheck/index.html> Pristupljeno 22. ožujka 2012.
11. Compilatio.net. Dostupno s: <http://www.compilatio.net/en/> Pristupljeno 22. ožujka 2012.
12. Safeassign. Dostupno s: <http://www.safeassign.com/> Pristupljeno 22. ožujka 2012.
13. Barrie JM, Presti DE. How original – catching the cheats. *The Biochemical Society* 2008;30(6):16-9.
14. Baždarić K, Bilić-Zulle L, Brumini G, Petrovečki M. Prevalence of plagiarism in recent submissions to the Croatian Medical Journal. *Sci Eng Ethics*. 2011 Dec 30. (article in press)
15. CrossRef. Dostupno s: <http://www.crossref.org/> Pristupljeno 22. ožujka 2012.
16. eTBLAST. Dostupno s: <http://etest.vbi.vt.edu/etblast3> Pristupljeno 22. ožujka 2012.
17. Deja vu: a database of Highly Similar Citations. Dostupno s: <http://spore.vbi.vt.edu/dejavu/> Pristupljeno 22. ožujka 2012.
18. zu Eissen SM, Stein B. Intrinsic plagiarism detection. *Advances in Information Retrieval*. 28th European Conference on IR Research, ECIR 2006. London, UK, 10-12 travnja 2006.
19. Stein B, Lipka N, et al. Intrinsic plagiarism analysis. *Language Resources & Evaluation* 2011;45(1):63-82.

ADRESA ZA DOPISIVANJE

Stella Lampret, prof. informatike i pedagogije
Katedra za medicinsku informatiku, Medicinski fakultet Sveučilišta u Rijeci
Braće Branchetta 20, 51000 Rijeka
E-mail: lstella@medri.hr
Telefon: +385 51 651 255