

Plagiranje kao povreda znanstvene i akademske čestitosti

Plagiarism as a violation of scientific and academic integrity

Ksenija Baždarić^{1*}, Vanja Pupovac¹, Lidija Bilić-Zulle^{1,2}, Mladen Petrovečki^{1,3}

¹Katedra za medicinsku informatiku
Medicinskog fakulteta Sveučilišta u Rijeci

²Zavod za laboratorijsku dijagnostiku
Kliničkog bolničkog centra Rijeka

³Klinički zavod za laboratorijsku dijagnostiku
Kliničke bolnice Dubrava u Zagrebu

Primljeno: 31. 3. 2009.

Prihvaćeno: 20. 4. 2009.

Adresa za dopisivanje:

* Ksenija Baždarić, dipl. psiholog, prof.
Katedra za medicinsku informatiku,
Medicinski fakultet Sveučilišta u Rijeci,
Braće Branchetta 20, 51000 Rijeka
e-mail: bksenija@medri.hr

<http://hrcak.srce.hr/medicina>

SAŽETAK. Plagiranje je neovlašteno preuzimanje tuđih ideja, postupaka ili teksta bez odgovarajuće naznake s nakanom da se prikažu kao vlastita. Preuzimanje već objavljenih vlastitih ideja ili preuzimanje vlastitog već objavljenog teksta prikazujući ga novim i originalnim naziva se samoplagiranjem i smatra se jednako lošim kao i plagiranje. Učestalost plagiranja je u porastu, razvoj informacijsko-komunikacijske tehnologije olakšava neovlašteno preuzimanje teksta, no istovremeno, zahvaljujući istoj tehnologiji, razvijaju se računalni programi i mrežne usluge za otkrivanje plagiranja.

Za provjeru tekstova i činjenice jesu li oni plagirani postoje različita programska rješenja. Većina se zasniva na konkordanciji, tj. usporedbi teksta pri čemu program iznalazi i označava podudarne dijelove teksta i izračunava njegov udio s obzirom na cijeli tekst. Neki od programa, osim što uspoređuju tekstove, pretražuju internet radi iznalaženja tekstova s podudarnim sadržajem. Svi programi mogu uspoređivati tekstove napisane na istom jeziku; međujezično pretraživanje nije moguće uz pomoć programske potpore. Primjeri programa su računalni programi (*WCOPYFIND*) i mrežne usluge (*eTBlast*, *CrossCheck*). Prednost mrežne usluge je mogućnost pronalaženja mogućeg teksta izvornika.

eTBlast je besplatna mrežna usluga za pronalaženje podudarnih i vrlo sličnih sažetaka znanstvenih radova (pretražuje i bazu podataka *Medline*) koja je poslužila kao temelj za izradbu baze podataka *Déjà vu*. Mrežna usluga *CrossCheck* dostupna je samo članovima (akademske ustanove i časopisi), a korištenjem računalnog algoritma za pronalaženje sličnosti *iThenticate* tvrtke *iParadigms* (Oakland, CA, SAD) provjerava podudarnost teksta ispitivanog članka s cjelovitim tekstovima sadržanim u bazi podataka *CrossCheck* ustrojenom suradnjom uredništva i izdavača znanstvenih časopisa koji objavljuju članke pridružuju bazi podataka, što omogućuje nesmetano pretraživanje sadržaja zaštićenog pretplatom.

Vrlo je važno prepoznati plagiranje i podučavati o njemu u akademskoj zajednici na svim nivoima obrazovanja. Urednici znanstvenih časopisa i znanstvenici trebaju se zajedno boriti protiv neetičnih istraživanja koja su suprotna znanstvenoj ideji i štetna za znanstvenu i opću zajednicu, kritički čitati i provjeravati znanstvenu publicistiku, prijavljivati plagiranje i ostale sumnjive postupke u istraživanjima uredništvima časopisa i nadležnim tijelima.

Ključne riječi: bibliografske baze podataka, dvostruke publikacije, plagiranje, povlačenje objavljenih publikacija, prijevara, znanstveno nepoštenje

ABSTRACT. Plagiarism is unauthorized appropriation of other people's ideas, processes or text without giving correct credit and with intention to present it as own property. Appropriation of own published ideas or text and passing it as original is denominated self-plagiarism and considered as bad as plagiarism. The frequency of plagiarism is increasing and development of information and communication technologies facilitates it, but simultaneously, thanks to the same technology, plagiarism detection software is developing.

There are different software solutions for checking plagiarism. Most of them are based on concordance, i.e., comparison of text where program tools isolate and mark correspondent parts of the text and calculate its rate regarding the whole text. Several programs, besides comparing the texts, also search the Internet aiming for text with corresponding content. All programs can compare text written in the same language but translanguar comparison with plagiarism detection software is not yet possible. The software is available through computer programs (*WCOPYFIND*) or Web Services (*eTBlast*, *CrossCheck*). Their advantage is in the possibility of finding the original source paper.

eTBlast is the free of charge web based service for searching corresponding and highly similar scientific paper abstracts (it searches also *Medline* database), which served as the ground for constructing *Déjà vu* database. Web based service *CrossCheck* is accessible only for members (academic institutions and journals) and by using computer similarity algorithm *iThenticate* of company *iParadigms* (Oakland, CA, USA), it checks accordance of the given text with the complete texts in the *CrossCheck* database. It is organized by collaboration of journal editorial boards and publishers who pass the published papers to the base and enable searching of content usually protected by subscription.

The importance of recognizing and teaching plagiarism in the academic community at all levels of education is enormous. Scientific journal editors and scientists should fight together against unethical researches which are opposite to the scientific idea and harmful for scientific community and society, critically read and examine scientific publications, report plagiarism and other suspicious research misconduct to journal editorial boards and institutional authorities.

Key words: Bibliographic Databases, Duplicate Publication, Fraud, Plagiarism, Retraction of Publication, Scientific misconduct

UVOD

Objavljivanje rezultata znanstvenog rada u obliku izvješća – znanstvenog članka predstavlja krunu znanstvenoistraživačkog rada svakom znanstveniku. Često dugotrajan i naporan rad tek će objavljivanjem biti pridodan ljudskom znanju te trajno podastrijet znanstvenoj zajednici na kritičku prosudbu. Javnost, znanstvena i opća, očekuje iskrenost, pravednost i poštenje jer su upravo to osobitosti koje znanosti daju smisao i u samoj su njezinoj srži^{1,2}.

Znanstveni radovi danas su većinom dostupni u elektroničkom obliku u bazama podataka koje omogućuju jednostavan pristup i razmjenu informacija te uporabu rezultata kao najvažniji oblik suradnje u znanosti. Elektroničke baze podataka koje citiraju znanstvene radove, npr. *Current Contents (CC)*, *Medline*, *Web of Science (WOS)*, omogućuju jednostavno i brzo pretraživanje pouzdanih znanstvenih izvora. Znanstveni sadržaji koji nisu dostupni u elektroničkim bazama podataka nevidljivi su i neprepoznatljivi znanstvenoj zajednici²⁻⁴. Broj novih biomedicinskih informacija u stalnom je porastu pa se tako godišnje u bazu podataka *Medline* unese gotovo pola milijuna novih radova od kojih većina ima sažetke⁵.

Važnost biomedicinskih istraživanja i njihovog mogućeg izravnog utjecaja na živote i zdravlje ljudi naglašava potrebu za poštenjem znanstvenika u radu te potpuno poštivanje svih počela znanstvenoistraživačke čestitosti. Ukupna znanost je kumulativna i pripada cijelom čovječanstvu, svako je znanstveno otkriće nadopunjuje i gradi, znanost je “izvor stvarnog znanja (otkrića) za čovječanstvo” i “pridonosi općoj dobrobiti i sigurnosti u svagdanjem životu”⁶. Svako nepoštenje u znanosti, bez obzira na razloge zbog kojih je počinjeno, umanjuje je i obezvrjeđuje, priskrbljujući pritom autorima nezasluženu dobit⁷.

Znanstveno nepoštenje (engl. *scientific misconduct*) uključuje izmišljanje (engl. *fabrication*) podataka ili rezultata, prepravljivanje (od engl. *falsification*) rezultata i plagiranje (od engl. *plagiarism*) u predlaganju, provođenju ili razmatranju istraživanja ili u predočavanju rezultata istraživanja^{1,6-8}.

Izmišljanje podrazumijeva produkciju podataka ili rezultata istraživanja koje se ne provodi. Istraživač izmišlja, producira nepostojeće rezultate koje

kasnije obrađuje i prikazuje, odnosno objavljuje kao istinite^{6,7}.

Prepravljivanje uključuje naknadno mijenjanje ili izostavljanje postupaka ili rezultata kako bi se objavio željeni ili izbjegao neželjeni rezultat. Istraživanje se provodi, ali se neki rezultati mijenjaju ili zanemaruju kako bi se približili postavljenoj hipotezi. Kao i izmišljanje, predstavlja težak oblik znanstvenog nepoštenja⁶.

Plagiranje je preuzimanje tuđih ideja, postupaka, rezultata ili teksta bez navođenja izvora radi pri-

Plagiranje je preuzimanje tuđih ideja, postupaka, rezultata ili teksta bez navođenja izvora radi prikazivanja preuzetog kao vlastitog djela. Smatra se kako je povezano s četvrtinom prijavljenih prijevvara u znanosti. Postoje dokumentirani slučajevi plagiranja u znanosti od prije čak dvije stotine godina.

kazivanja preuzetog kao vlastitog djela^{6,9,10}. Plagiranje u znanosti vrlo je složen problem. Postoje dokumentirani slučajevi plagiranja u znanosti od prije čak dvije stotine godina. Pojave neovlaštenog preuzimanja vlasništva u medicini postaju ozbiljan problem, nepotrebno povećavaju znanstvenu publicistiku neoriginalnim rezultatima i tekstovima i u vremenu mjerenja znanstvene produkcije kao pokazatelja uspješnosti znanstvenika priskrbljuju značajnu nezasluženu dobit. Pojam plagiranja se u navedenom engleskom obliku *plagiarism* uvodi u tezaurus MeSH 1990. godine⁹. Pretraživanjem baze podataka *Medline* pretraživačem *PubMed* Nacionalne medicinske knjižnice u SAD-u doznajemo da je od 1990. godine do danas (brojeno 2. travnja 2009.) objavljeno 667 radova koji sadrže predmetnicu plagiranja.

Društvene norme i važnost djelatnosti od znanstvenika iziskuju potpunu predanost i poštenje u istraživanju i pisanju radova. Znanstvene se publikacije čuvaju u pismohranama pa, za razliku od lopova koji krade i nestaje, nepošteni znanstvenik zauvijek ostavlja za sobom trag i mogućnost provjere svoga rada. Kada se nepoštenog znanstvenika otkrije, istinitost svih njegovih znanstvenih radova i rezultata dovodi se u pitanje, a on sam može biti isključen iz znanstvene zajednice^{6,7,9}.

PLAGIRANJE

Točno definiranje pojma plagiranja vrlo je teško. Postoje mnogobrojne definicije, a danas je usvojeno stajalište kako je plagiranje neovlašteno preuzimanje autorskog vlasništva, tuđih ideja, postupaka ili riječi bez odgovarajuće naznake autora⁹. Osim preuzimanja ideja ili pisanih riječi ono uključuje i parafraziranje bez navođenja izvora te neovlašteno preuzimanje prezentacija, predavanja i slika iz znanstvenih radova ili udžbenika¹¹.

Plagirati se može na različite načine. Preuzimanje većine ili cijelog teksta bez odgovarajuće naznake najteži je oblik plagiranja i krađe intelektualnog vlasništva, što se ponekad opisuje kao napadno (engl. *blatant*) plagiranje. Jedan od primjera takva plagiranja opisan je nedavno u časopisu *Nature*. Francuski gerontolog Eric Le Bourg čitajući korejski časopis *Korean Journal of Biological Science* pronašao je svoj godinu dana ranije objavljen pregledni članak doslovno i u cijelosti prenesen, uključujući i slike¹². Poseban oblik plagiranja su i tzv. "kolažni radovi" (engl. *assembly-kit*), čiji su dijelovi sastavljeni od različitih znanstvenih radova¹³.

Plagirati se može i tekst sa stranoga jezika i to tako da se doslovno prevede cijeli rad, bez navođenja izvornih autora i bez njihova znanja i dopuštenja. Takav je oblik plagiranja vrlo teško dokazati jer je programskom potporom za otkrivanje plagiranja nemoguće usporediti dva ili više tekstova pisanih različitim jezicima^{5,13}. Takvi nalazi su uglavnom slučajni, upravo kao opisani slučaj u časopisu *Korean Journal of Biological Science*.

Ponekad se autori izvornih radova i navode, ali preuzeti sadržaj nije u navodnicima, pa se stječe dojam da je tekst parafraziran iako je doslovno preuzet, a to jednako tako nije dopušteno. Citiranje literature nije dopuštenje za preuzimanje teksta, treba citirati činjenice, ideje i rezultate drugih autora u kontekstu vlastitog istraživanja i vlastitim riječima. O plagiranju se radi i onda kada se sadržaj preuzme te se tek poneke riječi zamijene istoznačnicama, upravo zato da tekst ne bi bio identičan izvorniku¹³.

Plagiranje jest krađa intelektualnog vlasništva, predstavlja kriminalan čin protiv znanstvenika i znanosti same, te je kao takav kažnjiv. Rezultati plagiranih znanstvenih radova neupotrebljivi su

jer ne donose nova znanja, ne pridonose znanosti, mogu dovesti u zabludu, a koriste isključivo stjecanju nezasluženi priznanja autorima rada^{6,8}. U svrhu promicanja znanstvene čestitosti i izobrazbe znanstvene zajednice znanstveni časopisi kao, primjerice, *Croatian Medical Journal* 2001. godine, uvode urednika za znanstvenoistraživačku čestitost¹⁴. Takvi bi urednici, kao "čuvari dobre znanosti", trebali osigurati objavljivanje samo izvornih radova u svojim časopisima, no ipak se, nažalost, većina znanstvenog nepoštenja još uvijek otkriva tek nakon objavljivanja¹⁵.

U slučajevima napadnog plagiranja, tj. neovlaštena preuzimanja većeg dijela ili čitavoga članka, ne postoji suglasnost urednika o tome kako postupiti, tj. kako kazniti autore članka. Ponajčešće je glavna mjera uredništva povlačenje članka iz časopisa (engl. *retraction*) nakon odluke nadležnog tijela koje se bavi znanstvenom čestitošću. Podatak o "retrakciji" rada upisuje se i u citatne baze podataka, pa tako, primjerice, pretraživanjem u *PubMedu* možemo pronaći članke uza čiji je naslov upisana oznaka o povlačenju, što znači da se taj rad smatra nepostojećim, te se ne smije rabiti u citiranju literature. Osim povlačenja članka i javne (pisane) isprike pojedini urednici zagovaraju i radikalna rješenja, kao što su prijavljivanje nadležnim institucijama (npr. odgovarajućem povjerenstvu nadležnog fakulteta ili ministarstva) te zabrana objavljivanja na neko vrijeme, posebice za autore u kojih se ustvrdi da su prijestup počinili opetovano ili s očitom namjerom da prevare¹⁶. Osim urednika časopisa postoje i državna tijela čiji je zadatak promicanje znanstvenoistraživačke čestitosti, sprječavanje znanstvenog nepoštenja, a imaju i ulogu savjetodavnog tijela u slučajevima znanstvenog i akademskog nepoštenja. U svijetu je najpoznatiji Ured za znanstvenoistraživačku čestitost, ORI (engl. *Office of Research Integrity*) pri Uredu za javno zdravstvo i znanost u SAD-u koji izrađuje preporuke za izobrazbu i daje smjernice u znanstvenoistraživačkoj čestitosti te vodi postupke u optužbama za prijestupe u biomedicinskim istraživanjima^{7,17}.

Po uzoru na ORI i srodne organizacije u svijetu u Hrvatskoj je 2006. godine osnovan Odbor za etiku u znanosti i visokom obrazovanju (OEZVO), savjetodavno i stručno tijelo Hrvatskoga sabora koje

djeluje pri Agenciji za znanost i visoko obrazovanje (www.azvo.hr). Odbor promiče najviše etičke vrijednosti u znanosti i čini ga devet cijenjenih stručnjaka iz područja znanosti i visokog obrazovanja koji na svojim sjednicama donose mišljenja i stavove o etičkoj prihvatljivosti slučajeva koje prijavljuju pojedinci, znanstvene i akademske ustanove. Odbor je zadužen i za oblikovanje smjernice znanstvenog djelovanja u obliku etičkoga kodeksa kojim se utvrđuju etička načela znanstvenog rada, objavljivanja rezultata i odnosa među znanstvenicima. Odbor planira osnivanje hrvatskoga Ureda za znanstvenoistraživačku čestitost^{14,18}. Odluke Odbora javne su i dostupne na mrežnim stranicama Agencije.

U izvješću iz 2005. godine ORI procjenjuje da je gotovo 25% od svih prijavljenih prijevara u znanosti povezano s plagiranjem¹. Učestalost znanstvenog nepoštenja u znanstvenim institucijama SAD-a ispitana je i objavljena u izvješću ORI-ja iz 2006. g. U istraživanju je sudjelovalo 2.216 znanstvenika iz 605 sveučilišta, instituta, bolnica i drugih zdravstvenih organizacija u SAD-u koji su ispunjavali upitnik o mogućim slučajevima znanstvenog nepoštenja. Odgovarajući na pitanja u upitniku, znanstvenici su trebali izvijestiti o pojavnosti mogućeg znanstvenog nepoštenja na svom radnom mjestu. Od ispitanih istraživača njih 164 (7%) izvijestilo je o 201 mogućem slučaju znanstvenog nepoštenja u posljednje tri godine (2002. – 2005.), od kojih je 36% povezano s plagiranjem. Više od polovice mogućih slučajeva znanstvenog nepoštenja (58%) prijavljeno je u ustanovama u kojima se sumnjivo istraživanje i provodilo. U izvješću se ističe da poznavanje pravila institucije o znanstvenoistraživačkoj čestitosti i znanje kome se može prijaviti znanstveno nepoštenje bitno utječu na prijavljivanje slučajeva mogućih znanstvenih nepoštenja^{17,19}.

Istraživanjem učestalosti i razloga povlačenja članaka iz baze podataka *Medline* ispitani su članci u vremenskom razdoblju od deset godina, od 1982. do 2002. U tom je razdoblju ukupno povučeno 395 članaka, od čega 107 (27%) zbog znanstvenog nepoštenja, a 244 (62%) zbog nenamjernih pogrešaka²⁰. Visoki udio povučениh članaka zbog nenamjernih pogrešaka upućuje na nužnu potrebu izobrazbe i promicanja istraživačke čestitosti unutar znanstvene zajednice.

SAMOPLAGIRANJE

Plagiranje vlastitih ideja ili preuzimanje već objavljenog teksta jednako je loše kao i plagiranje, ako se već objavljeni radovi ili njihovi dijelovi objavljuju ponovno i pri tome predstavljaju kao izvorni. Većini je znanstvenika pojam plagiranja u svojoj osnovi jasan i suglasni su kako je neispravan, međutim, vođeni zaključkom kako se "ne može krasti od sebe samoga", mnogima nije jasno zašto je samoplagiranje uopće štetno i zbog čega ga treba kažnjavati. U većini se slučajeva za samoplagiranjem i višestrukim objavljivanjem radova poseže radi povećanja vlastite znanstvene produkcije i to ponajčešće zbog akademskog napredovanja, pribavljanja potpore za prijavu ili pravdanja istraživačkih projekata i sl. Takva je djelatnost nevaljala, ni sa čim se ne može opravdati, a uz to ni ne pridonosi povećanju svjetskog znanstvenog opusa⁷. Učestalost samoplagiranja nije jednostavno ustvrditi, no primijećen je veći udio samoplagiranja u preglednim člancima u usporedbi s izvornim znanstvenim radovima¹⁶. Prema urednicima časopisa *Nature* učestalost samoplagiranja je u porastu i to vjerojatno stoga što se odobravanje financiranja i napredovanja u znanosti uvelike zasniva na samom broju znanstvenih radova²¹. Samoplagiranje se najčešće pronalazi kao preuzimanje prethodno objavljenog teksta bez navođenja izvora, objavljivanje sličnog teksta pod sličnim naslovom ili objavljivanje dvostrukih ili višestrukih publikacija (engl. *duplicate, multiple publications*), kada se isti (sličan) rad objavljuje u dva ili više časopisa. Granicu samoplagiranja teško je postaviti jer urednici nekih časopisa samoplagiranje smatraju lakšim oblikom znanstvenog nepoštenja te dopuštaju da čak do 30% teksta uvoda bude preuzeto iz već objavljenog znanstvenog rada istog autora²¹. Kako bi se izbjeglo nepotrebno prepisivanje tekstova, opise prethodnih istraživanja u uvodnim dijelovima znanstvenih radova treba pisati što je moguće kraće, s navođenjem izvora u kojem čitatelj može iznaći opširnije podatke. Uz uvodni dio rada, najčešće preuzimani dijelovi objavljenih tekstova su iz opisa postupaka i tvoriva (materijal i metode, ispitanici i postupci) koji se često ponavljaju. Pravilo je kako u opisu metodologije ne treba iznositi više od osnovnih informacija, već na odgovarajući način citirati ono što je već objavljeno. Preuzimanje dijelova rezul-

tata, rasprave i zaključaka smatra se težim i najtežim prijestupom znanstvene čestitosti⁵ i teško može imati ma kakva opravdanja.

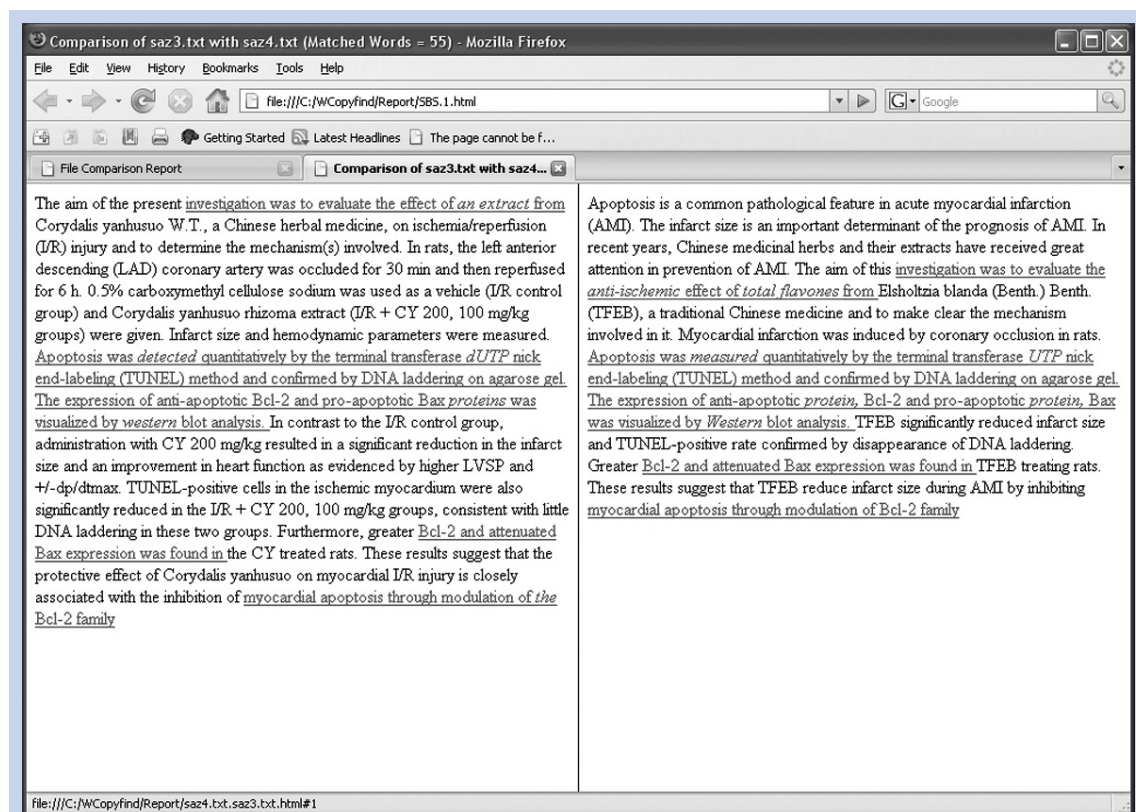
U višestruko objavljivanim publikacijama autori zbog različitih motiva predaju u isto vrijeme rukopis urednicima više časopisa. Upute autorima gotovo svih znanstvenih časopisa sadrže jasna upozorenja o zabrani takva postupanja²². Još 1969. godine urednik časopisa *New England Journal of Medicine*, Franz J. Ingelfinger, ustvrdio je kako njegov časopis neće objaviti rukopis koji je u isto vrijeme predan za objavljivanje u bilo koji drugi časopis, pa se pravilo o zabrani objavljivanja rada u više časopisa naziva Ingelfingerovim pravilom¹⁶. Vedran Katavić, urednik za znanstvenoistraživačku čestitost časopisa *Croatian Medical Journal* i predsjednik OEZVO, navodi u petogodišnjem izvješću urednika za znanstvenoistraživačku čestitost kako autori najčešće nisu svjesni da je višestruko objavljivanje neprihvatljiv čin, a razlog tomu nalazi u njihovoj nedovoljnoj educiranosti¹⁴.

Časopis *Otornolaringology – Head and Neck Surgery* objavio je istraživanje višestrukih publikacija

provedeno u vlastitom časopisu. Od gotovo 900 pregledanih članaka, 20% ih je višestruko objavljeno, od čega je većina autora imala dvostruko, a ne pravo višestruko objavljivanje članaka²³.

Nacionalna medicinska knjižnica (NLM) u SAD-u objavila je podatak da je u 2006. godini u bazi podataka *Medline* pronađeno 607 zapisa koji su prvim pregledom označeni kao dvostruko objavljeni radovi, što bi iznosilo oko 0,001% od ukupnog broja objavljenih radova (godišnje se objavi oko 16 milijuna radova). Godine 2008. Errami i sur. pregledali su radove koje je NLM pronašla u bazi podataka *Medline* te zaključili kako se za 171 rad (28%) može ustvrditi da se uistinu radi o slučajevima dvostrukoga publiciranja⁵.

Plagirani, samoplagirani radovi i dvostruko objavljeni radovi neprihvatljivi su u znanstvenoj zajednici. Takvi radovi nemaju smisla, a troše dragocjeno vrijeme urednika, recenzenata i čitatelja. Znanstvenici koji pribjegavaju takvoj praksi stječu nezasluženu dobit te odaju lažan dojam visoke znanstvene produktivnosti^{5,6}.



Slika 1. Usporedba izvornog i plagiranog teksta uporabom programa WCopyfind
Figure 1. Comparison of original and plagiarized text using WCopyfind program

OTKRIVANJE PLAGIRANJA POMOĆU RAČUNALNE PROGRAMSKE POTPORE

Razvoj tehnologije i velika ulaganja u znanost pridonijeli su napretku znanosti, većem broju istraživanja i istraživača te, posljedično, većem broju znanstvenih publikacija. Ujedno je i, kao nikada do sada, pojednostavljena mogućnost neovlaštenog preuzimanja sadržaja. Dok je nekada tekst trebalo prepisivati od znaka do znaka, sada se odabirom jednostavnih naredbi preuzimanja (engl. *copy*) i postavljanja (engl. *paste*) mogu doslovce u sekundama preslikavati goleme količine teksta²⁴. No, sve se više razvijaju i računalni programi za otkrivanje plagiranog teksta koji omogućuju otkrivanje neovlaštenog preuzimanja autorskog vlasništva. Razvoj računalnih algoritama za pretraživanje teksta i elektronički indeksiranih znanstvenih radova omogućio je lakše pronalaznje neetičnih publikacija⁵.

Plagirani tekst može se provjeriti i bez ikakvog posebnog programa pomoću uobičajenih tražilica kao što su, primjerice, *Google* (www.google.hr) ili *Yahoo* (www.yahoo.com). Koristeći ključne riječi ili dijelove teksta rada za koji se sumnja da je plagiran moguće je pronaći izvornik. Rezultati takve pretrage su nepregledni, nepouzdana i najčešće ne obuhvaćaju relevantnu literaturu, pa je stoga razvijeno više vrsta programskih potpora namijenjenih provjeri plagiranih tekstova¹¹.

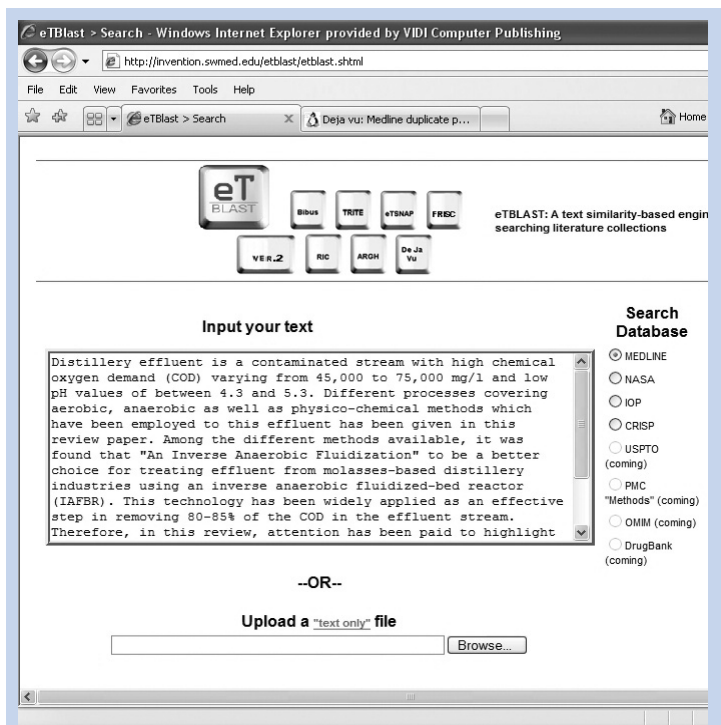
Računalni programi za provjeru plagiranih tekstova koriste algoritam za otkrivanje sličnosti i pronalaze istovjetnost u dijelovima ispitivanih tekstova. Programi mogu pretraživati i mrežne baze podataka i sličnost zadanog teksta usporediti s tekstovima u bazama podataka i tada se nazivaju mrežnim uslugama za provjeru plagiranja. Većina takvih mrežnih usluga prilagođena je tekstovima na engleskom jeziku.

Programi za otkrivanje plagiranja izračunavaju i iskazuju udio podudarnih dijelova tekstova. Uobičajeno je da prikaz rezultata usporedbe sadrži tekstove obaju članaka, ispitivani tekst u usporedbi s pronađenim podudarnim tekstom, a u oba se istaknuto prikazuju dijelovi teksta koji se podudaraju. Nažalost, unatoč brojnim prednostima računalnih programa i mrežnih usluga za otkrivanje plagiranja nije moguće otkriti tzv. pametno plagiranje, tj. parafraziranje teksta bez

navođenja izvora te doslovne prijevode znanstvenih radova s nekog drugog jezika⁵.

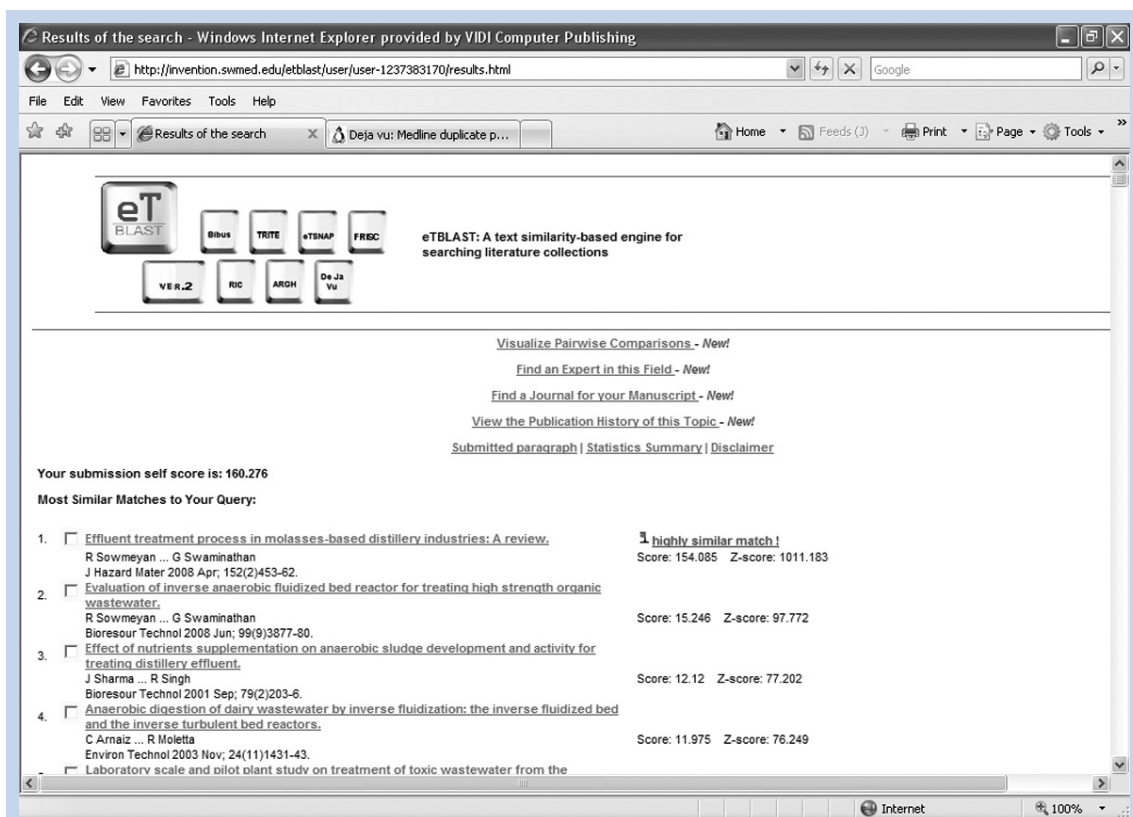
Računalni program *WCopyFind* Louisa Bloomfiel- da, profesora fizike s Virginijskog sveučilišta, može se preuzeti s mrežnog mjesta <http://plagiarism.phys.virginia.edu/home.html> i neograničeno rabiti bez novčane naknade (upravo to i jest bila temeljna Bloomfieldova ideja stvaranja programa – što veća uporaba u znanstvenoj i akademskoj zajednici). Program uspoređuje dva ili više zadanih predložaka, tražeći među njima podudarnost (slika 1). Podudarnim se, tj. plagiranim smatra istovjetni tekst duljine šest i više riječi u nizu. Kraći podudarni nizovi riječi mogu biti uobičajene je-

Računalna tehnologija olakšala je pristup informacijama u znanosti, ali i njihovo neovlašteno preuzimanje jednostavnim naredbama “kopiraj” i “zalijepi”. No, ista tehnologija omogućila je i nalaženje plagiranih sadržaja pomoću programske potpore za uspoređivanje teksta te mrežnih usluga koje pretražuju internet i bibliografske baze podataka u potrazi za podudarnim tekstovima.



Slika 2. Sučelje mrežne usluge *eTBLAST* pri unošenju sažetka za provjeru izvornosti teksta

Figure 2. Interface of the Web Service *eTBLAST* during abstract uploading



Slika 3. Rezultati provjere dobiveni mrežnom uslugom *eTBLAST*
 Figure 3. Results of the *eTBLAST* Web Service check

zične fraze i ne smatraju se plagiranjem. Iz ukupnog broja riječi u tekstu i broja prepisanih riječi izračunava se udio plagiranog teksta. Rezultat računalne programom *WCOPYFIND* je izvješće koje uključuje broj prepisanih riječi, udio prepisanog teksta i prikaz tekstova s istaknutim dijelovima za koje je utvrđena podudarnost. Program *WCOPYFIND* je vrlo jednostavan za korištenje te ima mogućnost računalne svih tekstova napisanih latiničnim pismom, bez obzira na jezik, koji ipak mora u oba teksta biti jednak²⁵.

Najpoznatije mrežne usluge za otkrivanje plagiranja su *eTBLAST* (<http://invention.swmed.edu/eTBLAST/eTBLAST.shtml>) s bazom podataka *DeJa Vu* te bazom podataka *CrossCheck* koja koristi program *iThenticate* (<http://www.crossref.org/crosscheck.html>) tvrtke *iParadigms* (Oakland, CA, SAD).

eTBLAST je besplatna mrežna usluga za pretraživanje sažetaka znanstvenih radova u više baza podataka između kojih i baza *Medline* koja je najzanimljivija za biomedicinu. *eTBLAST* je prvobitno zamišljen kao alternativni način pretraživanja

baze podataka *Medline* pomoću čitavih rečenica, dakle, prirodnim jezikom, na engleskom jeziku, a ne isključivo prema ključnim riječima, no s vremenom se pokazao kao uspješan alat za otkrivanje plagiranih sažetaka²⁶. Koristeći algoritam za sličnost teksta, *eTBLAST* izračunava sličnost između naslova i sažetaka znanstvenih radova te kao rezultat daje popis znanstvenih radova poredanih po stupnju sličnosti (slika 2)^{5,26,27}.

Postupak korištenja mrežne usluge *eTBLAST* vrlo je jednostavan: na mrežnu stranicu *eTBLAST* unosi se tekst sažetka kojem se želi provjeriti izvornost, te se kao rezultat dobiva popis sličnih radova. U pregledu rezultata može se odabrati nekoliko vrsti izvješća, a najjednostavnije i najzornije dobije se odabirom naslova pronađenog sažetka znanstvenog rada, u kojem su bojom istaknute riječi slične onima iz izvornika. Autori *eTBLAST* preporučuju pregled prvih pet rezultata te obvezno svih onih označenih kao "vrlo slični" (engl. *highly similar match*) (slika 3)²⁷. Prednosti mrežne usluge *eTBLAST* su jednostavnost i dostupnost, no rezultate treba vrlo oprezno tumačiti jer je moguće da

neki sažeci budu označeni kao slični, a cjeloviti tekst znanstvenih radova se uopće ne mora podudarati, pa se preporučuje dodatna provjera računalnim programom *WCOPYFIND* za izračun udjela podudarnosti u tekstovima i pomni pregled prepisanih dijelova teksta.

Déjà vu je besplatna mrežna baza podataka dvostrukih i vrlo sličnih bibliografskih zapisa *Medlinea*. Nastala je kao rezultat istraživanja u kojem je korištena *eBlast* mrežna usluga. Mrežnom uslugom *eBlast* provjereno je 74.780 sažetaka od kojih je gotovo 6.000 detaljno raščlanio etički tim baze podataka *Déjà vu* koji je pregledao i pročitao sve nalaze podudarnih radova te tumačio sličnosti prepoznate računalnim algoritmom. Samo računalnim programom provjereni radovi označeni su kao neprovjereni (engl. *unverified*), a detaljno raščlanjeni radovi svrstani su u jednu od sljedećih kategorija: duplikat (engl. *duplicate*), različit (engl. *distinct*), ispravljen (engl. *errata*), odobren (engl. *sanctioned*) i ažuriran (engl. *update*). Uvjeti koje rad mora zadovoljiti da bi bio označen kao mogući plagijat su dvostruko objavljeni članak i administratori baze podataka moraju tumačiti cjeloviti tekst obaju znanstvenih radova. Nakon što se posumnja da je rad mogući plagijat, šalju se upiti autoru za dodatne informacije o radu te se traži pokretanje istrage u časopisu objavljivanja. Pravo da rad proglaše plagijatom imaju isključivo etičko povjerenstvo ustanove u kojoj djeluje autor i urednički odbor časopisa u kojem je rad objavljen, dok etički tim baze *Déjà vu* izražava tek sumnju u poštovanje znanstvenoistraživačke čestitosti pri njegovoj izradbi. U bazi podataka *Déjà vu* trenutno postoje 73 kandidata za plagijat^{5,28}.

Administrativno osoblje baze podataka *Déjà vu* redovito ažurira i nadopunjuje bazu, te su sve promjene odmah vidljive. Korisnici mogu prijaviti sumnjivi rad, komentirati prijavljene radove ili dokazati neispravnost optužbi. Rad koji se nalazi u bazi podataka nije nužno izrađen suprotno načelima znanstvenoistraživačke čestitosti, nego je predložen znanstvenoj zajednici na procjenu. Članove znanstvene zajednice tako se potiču na raspravu i razmišljanje o odgovornom objavljivanju i pisanju znanstvenih radova²⁸.

Déjà vu se sastoji od dvije osnovne tablice: tablice unosa podataka (engl. *entries*) koja prikazuje slične radove i tablice znanstvenih radova (engl. *articles*)

koja detaljnije opisuje svaki rad. Tablica unosa sastoji se od bibliografskih zapisa (naziva i sažetka rada) koji su pri prvoj obradbi, provjerom *eBlastom*, automatski proglašeni sličnima ili dvostrukim publikacijama (slika 4). Uza svaki prijavljeni sažetak nalazi se i sažetak koji mu je sličan, te naznačeni datum i jezik objavljivanja, *Déjà vu* kategorija, a izračunat je i omjer sličnosti radova. Tablica znanstvenih radova predstavlja popis radova koji se nalaze u tablici unosa i sadrži naziv rada, autore, vrijeme i mjesto objavljivanja, identifikacijski broj u pretraživaču *PubMed* (*PubMed ID*), identifikacijski broj u bazi podataka *Déjà vu* (*Déjà vu ID*) i mrežnu poveznicu na baze podataka *Medline* i *Déjà vu*²⁸.

Pretraživanje baze podataka može se učiniti dvojako: jednostavnim pretraživanjem jedne od dviju tablica (unosa i znanstvenih radova) i statističkim pretraživanjem. Statističko pretraživanje omogućuje *Déjà vu* mrežna usluga, a temelji se na pet statističkih kriterija, i to prema ustanovi, državi, časopisu, jeziku rada i recenzentu. Statistika prema ustanovi, primjerice, daje prikaz svih radova čiji autori djeluju u određenoj ustanovi. Rezultat statističkog pretraživanja prikaz je radova prema kategorijama baze *Déjà vu*. Ako u statističko pretraživanje po državi unesemo Hrvatsku, kao rezultat ćemo dobiti 113 naslova, od toga 6 duplikata, 5 različitih, 2 odobrena, 1 ažuriran i 99 neprovjerenih²⁸.

CrossRef je udruga izdavača i uredništava znanstvenih časopisa koja je osigurala temelj za izradbu mreže stalnih poveznica na reference (citate) znanstvenih radova, disertacija, kongresnih priopćenja, tehničkih izvješća i ostalih znanstvenih publikacija. Osnovana je 2000. godine radi jednostavnog korištenja pouzdanih znanstvenih sadržaja u elektroničkom obliku. Udruga *CrossRef* danas broji 2.720 izdavača i udruženja (među kojima su i zvučna imena poput *American Society of Neuroradiology*, *BMJ Publishing Group*, *Elsevier*, *IEEE*, *Nature Publishing Group*, *Oxford University Press*, *Sage* i *Wiley Blackwell*) i svaki dan elektronički indeksira oko 16.000 različitih znanstvenih sadržaja²⁹.

Godine 2006. *CrossRef* prepoznaje problem otkrivanja plagiranja kao prioritet u znanstvenoj izdavačkoj zajednici te se u lipnju 2008. godine pokreće mrežna usluga za provjeru radova *CrossCheck*. Mrežna usluga *CrossCheck* razvijena je u suradnji s tvrtkom *iParadigme*, najpoznatijim proizvođa-

ID	Earlier Article	Later Article	Lag Mon	Language	Sim. Ratio	Share author	Classification	Modified date	Thumt
1	7587703 [Chai, B et al., 1995] [Medline]	9275340 [Chai, B et al., 1996] [Medline]	16	chi eng	0.91	Y	DUPLICATE	2007-02-14	
2	16475345 [He, Hong et al., 2005] [Medline]	16494333 [Liu, Junfeng et al., 2006] [Medline]	2	eng eng	0.59	Y	DUPLICATE	2007-02-14	
3	16734501 [Battershill, Anna J et al., 2006] [Medline]	16526829 [Battershill, Anna J et al., 2006] [Medline]		eng eng	0.96	Y	DUPLICATE	2007-02-14	
4	8199193 [Birchmeier, W et al., 1994] [Medline]	7587628 [Birchmeier, W et al., 1995] [Medline]		eng eng	0.79	Y	DUPLICATE	2007-02-14	
5	8980928 [Mahmood, I, 1996] [Medline]	8893036 [Mahmood, I et al., 1996] [Medline]		eng eng	0.91	Y	DUPLICATE	2007-02-14	
6	15497674 [O'Hagan, Derek T et al., 2004] [Medline]	16480788 [O'Hagan, Derek T et al., 2006] [Medline]	19	eng eng	0.90	Y	DUPLICATE	2007-02-14	
7	16868161 [Kim, Hyun S et al., ...]	16967220 [Kim, Hyun S et al., ...]		eng eng	0.68	Y	DUPLICATE	2007-02-14	

Slika 4. Tablica unosa vrlo sličnih sažetaka baze podataka Déjà vu
Figure 4. Déjà vu database table of highly similar abstracts entries

čem računalnih rješenja za otkrivanje plagiranja. Usluga sadrži *iParadigmeov* računalni algoritam provjere sličnosti teksta *iThenticate* i bazu podataka *CrossCheck*. Algoritmom *iThenticate* provjerava se podudaranje teksta ispitivanog članka s cjelovitim tekstovima znanstvenih radova koje sadrži baza podataka *CrossCheck*. Za sudjelovanje u bazi radova i za korištenje usluge *CrossCheck* potrebno je članstvo u udruzi *CrossRef*. Sudjelujući u izradbi baze podataka udruženja i izdavači ustupaju na korištenje svoje cjelovite tekstove. Većina znanstvenih članaka zaštićena je sustavom pretplate i nemoguće ih je besplatno pregledavati ili koristiti za otkrivanje plagiranja ostalim računalnim programima, stoga je pomoću usluge *CrossCheck* algoritma traženja sličnosti *iThenticate* omogućen privilegiran pristup relevantnoj bazi radova i to isključivo u svrhu otkrivanja plagiranja. Onemogućen pristup cjelovitim znanstvenim radovima najčešći je nedostatak ostalih mrežnih usluga za akademsko otkrivanje plagiranja³⁰. Sudjelovanje časopisa u udruzi *CrossRef* i nadogradnja baze podataka *CrossCheck* donose izrav-

nu dobrobit znanstvenoj zajednici jer povećavaju zajedništvo izdavača i časopisa, a time smanjuju vjerojatnost pojavljivanja plagiranih radova. Snaga baze *CrossCheck* mjeri se brojem njezinih članica i u stalnom je porastu (40 članova), a među njima je većina časopisa uglednih izdavača.

ZAKLJUČAK

Napredak znanstvene zajednice temelji se na poštenju i pravednosti znanstvenika pri izradbi i objavljivanju istraživanja. Plagiranje i samoplagiranje štetni su postupci neovlaštenog preuzimanja autorskog vlasništva koji autoru daju nezasluzenu dobit, umanjuju znanstvenu kreativnost i ne pridonose znanosti i društvu. Učestalost plagiranja i samoplagiranja teško je točno utvrditi, a istraživanja pokazuju kako su ti oblici znanstvenog nepoštenja u porastu.

Znanstvenici i urednici znanstvenih časopisa napredak informacijsko-komunikacijske tehnologije mogu iskoristiti u svrhu međusobnog povezivanja i stvaranja okruženja za raspravu te podizanje svijesti o etičkoj provedbi istraživanja. Računalni

programi i mrežne usluge za otkrivanje plagiranja te indeksirane baze podataka znanstvenih radova velika su pomoć u postupku dokazivanja plagiranja, no one su tek postupci kojima treba znati rukovati i čije rezultate trebaju pomno tumačiti isključivo relevantni stručnjaci. Učestalo korištenje tehnologije pri postupku otkrivanja plagiranja jedan je od koraka promicanja znanstvenoistraživačke čestitosti.

Otkrivanje plagiranja važno je za znanstvenu i akademsku zajednicu koliko i izobrazba njezinih članova o znanstvenoistraživačkoj čestitosti. Znanstvenik treba prijaviti etičkom povjerenstvu ustanove ili nadležnom državnom tijelu sumnjiv ili neetičan postupak u istraživanju, uključujući plagiranje, a urednicima znanstvenih časopisa na raspolaganju su računalne baze podataka, mrežne usluge i računalni programi za otkrivanje plagiranja. U sprječavanju plagiranja ključan je nadzor znanstvene publicistike, ali i pravodobna izobrazba znanstvenika.

ZAHVALA

Rad je dio znanstvenog projekta "Učestalost i stavovi o neovlaštenom preuzimanju autorskog vlasništva" br. 062-0000000-3552 Ministarstva znanosti, obrazovanja i športa Republike Hrvatske.

LITERATURA

- Benos DJ, Fabres J, Farmer J, Gutierrez JP, Henessy K, Kosek D et al. Ethics and scientific publication. *Adv Physiol Educ* 2005;29:59-74.
- Marušić M. Znanost i znanstveno istraživanje. u: Marušić M (ed.) Uvod u znanstveni rad u medicini. Zagreb: Medicinska naklada, 2008;1-25.
- Frković V, Skender T, Dojčinović B, Bilić-Zulle L. Publishing scientific papers based on Master's and Ph.D. theses from a small scientific community: case study of Croatian medical schools. *Croat Med J* 2003;44:107-11.
- Petrovečki M, Scheetz MD. Croatian Medical Journal introduces culture, control, and the study of research integrity. *Croat Med J* 2001;42:7-13.
- Errami M, Hicks JM, Fisher W, Trusty D, Wren JD, Long TC et al. Déjà vu – a study of duplicate citations in Medline. *Bioinformatics* 2008;24:243-9.
- Katavić V. Odgovorna provedba istraživanja. u: Marušić M (ed.) Uvod u znanstveni rad u medicini. Zagreb: Medicinska naklada, 2008;233-43.
- Bilić-Zulle L. Znanstvena čestitost – temelj postojanja i razvoja znanosti. *Biochemia Medica* 2007;17:143-50.
- Mojon-Azzi SM, Mojon DS. Scientific misconduct: from salami slicing to data fabrication. *Ophthalmic Res* 2004;36:1-3.
- Bilić-Zulle L, Frković V, Turk T, Ažman J, Petrovečki M. Prevalence of plagiarism among medical students. *Croat Med J* 2005;45:126-31.
- Bilić-Zulle L, Ažman J, Frković V, Petrovečki M. Is there an effective approach to deterring students from plagiarizing? *Sci Eng Ethics* 2008;14:139-47.
- COPE annual seminar 2007. London, UK: Committee on Publication Ethics; 2007. Available at http://publicationethics.org/files/u2/2007_report.pdf. Accessed March 18, 2009.
- Butler D. Entire-paper plagiarism caught by software. *Nature* 2008;715:7214-5.
- Harris RA. Educating yourself about plagiarism. In: Harris RA (ed.) *The plagiarism handbook: strategies for preventing, detecting, and dealing with plagiarism*. Los Angeles: Pyczak Publishing, 2001;1-23.
- Katavić V. Five year report of Croatian Medical Journal's research integrity editor – policy, policing, or policing policy. *Croat Med J* 2006;47:220-7.
- Marušić A. Author misconduct: editors as educators of research integrity. *Med Educ* 2005;39:7-11.
- Neill US. Publish or perish, but at what cost? *J Clin Invest* 2008;118:2368.
- Titus SL, Wells JA, Rhoades LJ. Repairing research integrity. *Nature* 2008;453:980-2.
- Odbor za etiku u znanosti i visokom obrazovanju. Available at <http://www.azvo.hr/Default.aspx?sec=139>. Accessed March 17, 2009.
- Wells JA. Final report: observing and reporting suspected misconduct in biomedical research: Rockville, SAD: The Office of Research Integrity; 2008. Available at http://ori.dhhs.gov/research/intra/documents/gallup_finalreport.pdf. Accessed March 17, 2009.
- Nath SB, Marcus SC, Druss BG. Retraction in the research literature: misconduct or mistakes? *Med J Aust* 2006;185:152-4.
- Giles J. Special report: taking on the cheats. *Nature* 2005;435:258-9.
- Corson SL, Decherney AH. Duplicate editorial on duplicate publication. *Fertil Steril* 2005;83:855-6.
- Bailey BJ. Duplicate publication in the field of otorynolaryngology - head and neck surgery. *Otorynolaryngology – head and neck surgery* 2002;126:211-6.
- Harper MG. High tech cheating. *Nurse Educ Today* 2006;26:672-9.
- Bloomfield LA. WCopyfind 2.6 Instructions. Available at <http://plagiarism.phys.virginia.edu/WCopyfind%202.6.html>. Accessed March 17, 2009.
- Lewis J, Ossowski S, Hicks J, Errami M, Garner HR. Text similarity: an alternative way to search MEDLINE. *Bioinformatics* 2006;22:2298-304.
- eBLAST: A text similarity-based engine for searching literature collections. Available at <http://invention.swmed.edu/etblast/index.shtml>. Accessed March 17, 2009.
- Deja Vu: a database of highly similar and duplicate citations. Available at <http://spore.swmed.edu/dejavu/>. Accessed March 17, 2009.
- CrossRef. Available at <http://www.crossref.org/index.html>. Accessed March 17, 2009.
- CrossCheck. Available at <http://www.crossref.org/cross-check.html>. Accessed March 17, 2009.