

Analiza softvera za otkrivanje plagiranja u znanosti i obrazovanju

Birkić, Tamara; Celjak, Draženko; Cundeković, Marko; Rako, Sabina

Other document types / Ostale vrste dokumenata

Publication year / Godina izdavanja: **2016**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:102:340099>

Rights / Prava: [Attribution-NonCommercial-NoDerivatives 4.0 International](#)/[Imenovanje-Nekomercijalno-Bez prerada 4.0 međunarodna](#)

Download date / Datum preuzimanja: **2024-07-27**



Repository / Repozitorij:

[Digital repository of the University Computing Centre \(SRCE\)](#)



SVEUČILIŠTE U ZAGREBU
SVEUČILIŠNI RAČUNSKI CENTAR



I Z V J E Š T A J

**Analiza softvera za otkrivanje
plagiranja u znanosti i obrazovanju**

Zagreb, rujan 2016.

Izvještaj pripremili: Tamara Birkić, Draženko Celjak, Marko Cundeković, Sabina Rako
Lektorica: Mia Kožul

U Zagrebu 9. rujna 2016.

Klasa: 650-03/16-421/018

Ur. broj: 3801-4-421-01-16-1

Voditeljica tima:

Pomoćnica ravnatelja za obrazovanje i
podršku korisnicima:

Tamara Birkić, prof.

Sandra Kučina Softić, dipl.ing., v.r.



Ovo djelo dano je na korištenje pod licencom *Creative Commons Imenovanje-Nekomercijalno-Bez prerada 4.0 međunarodna*. Licenca je dostupna na stranici:
<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.hr>.

SADRŽAJ

1. UVOD	5
2. SOFTVERI ZA OTKRIVANJE PLAGIJATA	7
3. KRITERIJI I ANALIZIRANI SOFTVERI	8
3.1. API i PIUG-IN.....	9
3.2. Smještaj servisa	10
3.3. Obuhvat baze i mogućnost uključivanja vlastitoga sadržaja.....	10
3.4. Podrška	12
3.5. Rasprostranjenost	12
3.6. Cijena korištenja softvera i licence.....	13
3.7. Autentikacija i uloge korisnika.....	15
3.8. Testiranje softvera i procjena kvalitete	16
4. ZBIRNI REZULTATI TESTIRANJA	19
5. ZAKLJUČAK	20
6. PRILOG – PREGLED API-JA	22
6.1. PlagScan.....	22
6.2. Turnitin	23
6.3. Unplag.....	23
6.3. Urkund.....	24
7. REFERENCE.....	25

1. UVOD

Plagiranje, odnosno etičnost u obrazovanju općenito, posljednjih je godina aktualna tema. Vijeće Europe je, uvidjevši trendove u obrazovanju, 2015. godine uspostavilo pan-europsku platformu za etičnost, transparentnost i integritet u obrazovanju (ETINED)¹ čiji je jedan od ciljeva i zaštita, razvoj i podupiranje akademske čestitosti, s posebnim naglaskom na borbu protiv plagiranja. Posebno je istaknuta važnost akademske čestitosti među obrazovnim institucijama u visokom obrazovanju zbog povećanja broja studenata, ali i rastuće konkurencije među sveučilištima.

Europska komisija je isto tako prepoznala važnost ove teme te je u razdoblju od 2010. do 2013. godine proveden projekt *Impact of Policies for Plagiarism in Higher Education Across Europe* (IPPHEAE) čiji je cilj bio istražiti politike i sustave osiguranja akademske čestitosti i sprečavanja plagiranja u sustavu visokog obrazovanja (Glendinning, 2015.). U okviru ovoga projekta nastao je i model zrelosti akademske čestitosti (AIMM – *Academic Integrity Maturity Model*).

U modelu AIMM prepoznati su ovi kriteriji (Glendinning, 2014.) temeljem kojih je izrađena evaluacija stanja u 19 država Europske unije²:

- transparentnost akademske čestitosti i osiguranje kvalitete (engl. *transparency*)
- postojanje pravične, efektivne i konzistentne politike upravljanja plagiranjem (engl. *policies*)
- standardizacija sankcija u slučaju nepoštivanja akademske čestitosti (engl. *sanctions*)
- uporaba digitalnih alata i repozitorija u otkrivanju plagiranja (engl. *software*)
- postojanje preventivnih strategija i mjera (engl. *prevention*)
- komunikacija o politikama i postupcima (engl. *communication*)
- poznavanje i razumijevanje akademske čestitosti (engl. *knowledge*)
- omogućavanje dodatne edukacije studenata i nastavnika (engl. *training*)
- istraživanja i inovacije u području akademske čestitosti (engl. *research*).

Iz navedenog popisa kriterija vidljivo je da je pitanje plagiranja kompleksno i zahtijeva promatranje iz više perspektiva (organizacijskih i tehničkih). Usklađivanjem i uključivanjem prepoznatih elemenata može se uspostaviti sveobuhvatan sustav sprječavanja plagiranja.

Postojanje softvera za otkrivanje plagiranja važan je element kada je riječ o sustavnom pristupu otkrivanju plagiranja. Softveri za otkrivanje plagiranja imaju niz prednosti kao što su mogućnost provjere velike količine radova iz repozitorija radova u kratkom vremenskom razdoblju, provjeru sličnosti te izradu izvještaja koji mogu poslužiti kao potvrda originalnosti rada.

¹ <http://www.coe.int/en/web/ethics-transparency-integrity-in-education>

² Hrvatska nije bila uključena u projekt IPPHEAE pa ne postoje podaci temeljem kojih bi bilo moguće odrediti stanje u sustavu visokog obrazovanja u Hrvatskoj u odnosu na ostale europske zemlje.

Na hrvatskim sveučilištima postoji nekoliko primjera sustavnoga pristupa u rješavanju pitanja plagiranja, međutim dojam je da nedostaje strateški i sustavan pristup ovom problemu.

Ovim dokumentom Sveučilišni računski centar (Srce) želi potaknuti raspravu o pitanju plagiranja u sustavu visokog obrazovanja s naglaskom na softvere za otkrivanje plagiranja te prikazati i podijeliti trenutačna saznanja sa stručnom i širom javnosti. Saznanja Srca mogu biti preporuka ustanovama u sustavu visokog obrazovanja kod odabira ovakve vrste softvera, ali i potaknuti raspravu o potrebi za rješenjima na razini sveučilišta ili čak i na nacionalnoj razini.

Srce kao održavatelj nacionalnih e-infrastrukturnih sustava **Dabar**³ – Digitalni akademski arhiv i repozitorij, **Merlin**⁴ – sustav za e-učenje u visokom obrazovanju i **Hrčak**⁵ – Portal znanstvenih i stručnih časopisa RH rezultatom ove analize može dati preporuku ustanovama u visokom obrazovanju kako pomoću softvera otkriti i spriječiti plagiranje.

³ <https://dabar.srce.hr>

⁴ <http://www.srce.unizg.hr/usluge/sustavi-za-ucenje-na-daljinu/merlin>

⁵ <http://hrcak.srce.hr>

2. SOFTVERI ZA OTKRIVANJE PLAGIJATA

Rječnička definicija plagijata glasi „književno, znanstveno ili drugo djelo nastalo prepisivanjem u cjelini, u bitnim ili u prepoznatljivim dijelovima i prisvajanjem tuđega rada uložnog u to djelo“⁶. Kao što je i u uvodu navedeno, jedan od načina sprječavanja plagiranja je i uporaba softvera za otkrivanje plagiranja.

Zašto softveri za otkrivanje plagiranja? Što oni donose akademskoj zajednici?

Važnost softvera za plagiranje može se iščitati u mnogim pozitivnim primjerima uporabe takvih alata u svijetu. U Hrvatskoj postoji manji broj ustanova koje koriste ovu vrstu softvera.

Analizom radova softver daje korisniku informaciju o tome koliki je postotak sadržaja identičan, odnosno korišten iz drugih izvora. Međutim, važno je napomenuti isto tako da ne postoji definirana granica koja će pojedino djelo svrstati među plagijate pa je zato potrebno s razumijevanjem tumačiti rezultate dobivene softverskom analizom. Isto tako, rezultate softvera nije dobro interpretirati jednako u svim znanstvenim područjima. Mnogi uobičajeni tehnički izrazi ili matematičke formule koje ih opisuju ne mogu se smatrati plagijatom iako se pojavljuju u različitim radovima. Iz tog se razloga od nastavnika očekuje pregled rezultata analize i donošenje konačnoga zaključka o tome je li nešto plagijat.

Sveučilišta koja su implementirala softver za plagiranje primijetila su da njegova uporaba doprinosi i rastu svijesti među studentima o etičnosti te da studenti više obraćaju pažnju na to kako pravilno parafrazirati, referencirati i citirati nečiji rad (Stappenbelt i Rowles, 2009.).

Načini primjene softvera za plagiranje različiti su, od onog u kojem provjeru originalnosti rada provodi isključivo ovlaštena osoba visokoga učilišta pa do onog u kojem student samostalno provjerava rad prije službene predaje rada.

Važno je napomenuti da softveri za otkrivanje plagiranja imaju i određena ograničenja pa tako još uvijek mogu prepoznati plagijate samo u tekstualnim sadržajima, ali ne i u multimedijalnim sadržajima (slike, video itd.). Uglavnom softver uspoređuje nizove znakova iz rada predanoga na analizu s ostalim radovima koje ima pohranjene u svojoj bazi podataka pa je uspješnost softvera izravno ovisna o obuhvatu i kvaliteti baze koju taj softver koristi. Isto tako, za sada softveri ne otkrivaju plagijate ako se radi o prijevodima.

⁶ <http://hjp.znanje.hr/>

3. KRITERIJI I ANALIZIRANI SOFTVERI

Cilj analize koju je provelo Srce bio je upoznavanje s tehnologijama i mogućnostima softvera za otkrivanje plagijata te izrada preporuka za akademsku i istraživačku zajednicu u Hrvatskoj. Koraci su bili izrada kriterija, prepoznavanje najčešće korištenih softvera u akademskom okruženju i njihova analiza. Kontekst izbora softvera bio je usmjeren na uporabu na nacionalnoj razini i institucionalnoj razini (sveučilišta ili pojedine sastavnice) uzimajući u obzir relevantne izvore na hrvatskom jeziku (Hrčak, repozitoriji u Dabru) i relevantne informacijske sustave u kojima bi se softver mogao primijeniti (sustav za e-učenje Merlin).

Definirani kriteriji:

- postojanje programskih sučelja (API) i dodataka (PLUG-IN) zbog mogućnosti povezivanja sa sustavima Dabar i Merlin
- smještaj servisa (lokalno na poslužiteljima ustanove ili Srca i online na poslužiteljima održavatelja softvera)
- obuhvat baze i mogućnost uključivanja vlastitoga sadržaja
- podrška
- rasprostranjenost (uporaba softvera na svjetskoj ili europskoj razini te u susjednim državama – Slovenija, Bosna i Hercegovina, Srbija)
- cijena korištenja sustava i licence
- autentikacija putem korisničkoga računa iz sustava AAI@EduHr i načini autorizacije (postojanje više uloga u sustavu).

Nakon određivanja kriterija napravljen je pregled dostupnih softvera za plagiranje na tržištu te su odabrani oni koji su među najrasprostranjenijima u Europi. U ovoj analizi razmatrani su sljedeći softveri:

- PlagScan
- StrikePlagiarism
- Turnitin
- Unplag
- Urkund.

Nakon početne analize softver StrikePlagiarism nije dalje razmatran zbog nedostatka ključnih funkcionalnosti.⁷

⁷ Mogućnost pregleda jednoga dokumenta, teškoće s postavljanjem velikih datoteka i nepodržavanje formata PDF.

Pri analizi odabranih softvera kontaktirani su proizvođači softvera koji su kroz online prezentacije pružili tražene informacije o svakom softveru prema navedenim kriterijima. Također, praktično su testirane mogućnosti svakoga odabranog softvera.

Kriteriji za praktično testiranje softvera bili su:

- prepoznavanje citiranih dijelova
- podržani formati i ograničenja
- intuitivnost sučelja (subjektivna procjena i ocjena od 1 – 10, gdje je 1 najniža, a 10 najviša ocjena).

3.1. API I PLUG-IN

Cilj provjere po ovom kriteriju bio je utvrditi u kojoj se mjeri softver za otkrivanje plagiranja može povezati i integrirati s nacionalnim i drugim sustavima (npr. Dabar, Merlin) putem programskih sučelja (API) i dodataka (PLUG-IN) za Moodle i Drupal. Pri tome je važno postojanje podrške, mogućnost parametriziranja te dostupnost primjera i dokumentacije. Važan kriterij kod plug-ina za sustav Moodle bio je postojanje mogućnosti da se pojedinim ustanovama onemogućiti ili omogućiti uporaba softvera (sustav Merlin koriste sveučilišta i njihove sastavnice iz cijele Hrvatske).

PLAGSCAN	Dostupan je API za koji je dostupna dokumentacija na adresi http://www.plagscan.com/api-guide . Na istoj su adresi dostupni i primjeri za implementaciju u programskim jezicima Java, PHP i .NET. Dostupni su rezultati API-ja u obliku analiziranoga dokumenta s označenim dijelovima koje je potrebno provjeriti te statistikama (postotak podudaranja...). Postoji javno dostupan dodatak (<i>plug-in</i>) za instalaciju na sustav Moodle. Dodatak za Drupal ne postoji, ali prema potrebi voljni su ga izraditi.
TURNITIN	Nema API, ali je usklađen sa standardom Learning Tools Interoperability (LTI) koji omogućava povezivanje sa sustavima za e-učenje. Postoji javno dostupan dodatak (<i>plug-in</i>) za instalaciju na sustav Moodle pod nazivom Moodle Direct .
UNPLAG	Dostupan je API za koji se dostupna dokumentacija nalazi na adresi https://unplag.com/api/doc/ . Koristan je parametar <i>similarity sources</i> kojim se izostavljaju izvori s manje od 5% sličnosti. Postoji javno dostupan dodatak (<i>plug-in</i>) za instalaciju na sustav Moodle. Isproban na sustavu Merlin, ali instalacija nije bila uspješna.
URKUND	Dostupan je API, a dokumentacija je dostupna na zahtjev. Postoji javno dostupan dodatak (<i>plug-in</i>) za instalaciju na sustav Moodle. Dodatak je za potrebe testiranja uspješno instaliran na sustav Merlin. Jedino ovaj dodatak ima mogućnost ograničavanja uporabe po pojedinoj ustanovi na sustavu Merlin koji koristi više ustanova.

3.2. SMJEŠTAJ SERVISA

Ovim se kriterijem uspoređuju dostupne opcije za smještaj softvera i baze podataka: *online* na poslužiteljima održavatelja softvera i *lokalno* na poslužiteljima ustanove ili Srca. Prednost je online smještaja u tome da korisnik softvera ne mora planirati i održavati dodatne računalne resurse, a prednost lokalnoga smještaja je u tome da korisnik ima punu kontrolu nad podacima⁸.

PLAGSCAN	online Mogućnost instalacije na lokalni poslužitelj (paket <i>PlagScan in a Box</i>) koji se naplaćuje jednokratno 4000\$ po instalaciji te 99\$ mjesečno za održavanje. ⁹
TURNITIN	online Nema mogućnosti instalacije na lokalni poslužitelj.
UNPLAG	online Za dokumente pohranjene na razini ustanove i/ili suradničkih ustanova (<i>My Library</i>) može se instalirati lokalno (naplata 10.000 do 20.000 \$). ¹⁰
URKUND	online Nema mogućnosti instalacije na lokalni poslužitelj.

3.3. OBUHVAT BAZE I MOGUĆNOST UKLJUČIVANJA VLASTITOGA SADRŽAJA

Obuhvat baze odnosi se na izvore koje softver koristi kod otkrivanja plagijata. Za prepoznavanje plagijata nužno je da izvorni rad bude uključen u bazu koju softver koristi, stoga se može tvrditi da je uspješnost softvera izravno ovisna o obuhvatu i kvaliteti baze koju taj softver koristi. Od sadržaja dostupnih na internetu poželjno je da su u bazu uključeni ili se mogu uključiti relevantni hrvatski izvori radova (repozitoriji u Dabru, Hrčak). Od iznimne je važnosti da ustanova može u bazu softvera dodati i vlastitu bazu radova pri čemu je potrebno provjeriti uvjete pod kojima se dodaje vlastiti sadržaj.

PLAGSCAN	Četiri izvora: 1. internet (koriste Microsoft Bing tražilicu i odabrane akademske <i>web</i>-stranice) 2. vlastita baza (pretraživanje svih dokumenata ustanove i dokumenata korisnika) 3. publikacije i časopisi (imaju uključeno oko 21.000 znanstvenih časopisa) 4. PlagScan baza (radovi pohranjeni u bazi softvera PlagScan uz prethodno odobrenje autora rada ili onoga koji je rad predao na analizu).
-----------------	--

⁸ Vlasnici softvera politikom korištenja definiraju ovlasti nad radovima koji se predaju na provjeru.

⁹ Cijena dobivena na sastanku održanom u ožujku 2016., a informacija je dostupna i putem *web*-stranice na adresi <http://www.plagscan.com/in-a-box-local-server>.

¹⁰ Cijena dobivena na sastanku održanom u siječnju 2016.

	Iz web-sučelja moguće je odabrati pohranjivanje radova u bazu softvera. Ustanove mogu prijaviti vlastite baze i repozitorije, a u slučaju da je pristup prijavljenim sadržajima zaštićen, PlagScan im može pristupati i putem korisničkoga sučelja ili dostupnog API-ja. Postoji mogućnost uključivanja radova iz Hrčka i Dabra.
TURNITIN	Tri izvora: 1. internet (koriste komercijalnu tražilicu za akademske <i>web</i>-stranice) 2. znanstveni članci (u bazi obuhvaćen najveći broj časopisa od svih analiziranih softvera) 3. radovi postavljeni u njihovu bazu (korisnici nemaju mogućnost odabira žele li da se rad briše iz Turnitin baze). Pretraga obuhvaća preko 60 bilijuna <i>web</i>-stranica, 600 milijuna studentskih radova i 154 milijuna knjiga i časopisa. Radovi koji se provjeravaju trajno su dostupni u njihovoj bazi i ne mogu se brisati. Postoji mogućnost ograničenja na dokument uključivanjem opcije „<i>Nevidljiv drugima</i>“ (prilikom pretrage navodi postojanje sličnosti s dokumentom iz nekog izvora, ali se ne navodi ni naziv rada ni autor/i). U pretragu su uključeni radovi iz Hrčka, a na zahtjev i iz Dabra.
UNPLAG	Dva izvora: 1. internet (koriste Microsoft Bing i Yahoo tražilice te odabrane akademske <i>web</i>-stranice) 2. vlastita baza (softver može pretražiti radove dodane na razini ustanove i/ili suradničkih ustanova u bazu softvera Unplag (<i>My Library</i>)). Uz navedeno je omogućena i pojedinačna provjera (usporedba dva dokumenta). Prilikom dodavanja radova, postoji mogućnost podešavanja toga tko ima pristup dokumentu (globalno-svi, ustanova, nastavnik, student). Prilikom brisanja dokumenta iz baze, globalni se indeks također briše (ako je bio dostupan). Postoji mogućnost uključivanje radova iz Hrčka i Dabra (zatražena dozvola uključivanja repozitorija).
URKUND	Tri izvora: 1. internet (imaju vlastiti <i>crawler</i> i indeks, tj. ne oslanjaju se na tražilice) 2. objavljeni materijali u Urkund bazi (do 2014. ukupno 11,5 milijuna radova) 3. vlastita baza (pretraživanje svih dokumenata ustanove i dokumenata korisnika). Isključivanje radova iz Urkund baze moguće je na zahtjev. U pretragu su uključeni radovi iz Hrčka i Dabra.

3.4. PODRŠKA

Prilikom rada u sustavu institucijskim je korisnicima važna dostupnost stalne podrške. Osim s korisničke strane, važna je i podrška u razvoju i doradi sustava što osigurava stabilnost i dugoročnu održivost sustava. Podrška u ovom kontekstu ne uključuje podršku krajnjim korisnicima (studentima, nastavnicima).

PLAGSCAN	U standardnu podršku (uračunatu u cijenu softvera) uključena je: 1. on-line podrška (e-mail, webinar), telefon (9 – 17h) – očekivano vrijeme odgovora na upite unutar 24 sata 2. prema potrebi održavanje sastanaka na lokaciji klijenta Prilagođene dorade sustava se naplaćuju, osim u slučaju postojanja većeg broja zainteresiranih za istu doradu.
TURNITIN	U standardnu podršku (uračunatu u cijenu softvera) uključena je: 1. redovna podrška telefonom, e-mailom ili putem online obrasca
UNPLAG	U standardnu podršku (uračunatu u cijenu softvera) uključena je: 1. redovna podrška telefonom, e-mailom i preko kanala YouTube 2. ustanova dobiva osobu uz podršku koja je dostupna mobitelom (<i>key account manager</i>) Prilagođene dorade sustava se naplaćuju (cijena ovisi o opsegu posla).
URKUND	U standardnu podršku (uračunatu u cijenu softvera) uključena je: 1. redovna podrška telefonom, e-mailom ili putem Skype-a – očekivano vrijeme odgovora na upite je unutar 24 sata.

3.5. RASPROSTRANJENOST

Softveri za plagiranje koriste se u različitim dijelovima svijeta te je jedan od kriterija bio veličina korisničke zajednice. Poseban naglasak stavljen je na korisnike u Europi. Podaci u tablici dobiveni su u prvoj polovici 2016. godine.

PLAGSCAN	Prema podacima vlasnika softvera, u Hrvatskoj nema institucijskoga korisnika ovoga softvera. Najveći korisnici u Europi su sveučilišta u Njemačkoj (1200 institucija – 800 škola i 400 sveučilišta), Austriji, Ukrajini, Španjolskoj, Cipru i Švicarskoj. U susjednim zemljama ovaj softver dvije godine koristi Sveučilište u Mariboru.
TURNITIN	Prema podacima vlasnika softvera, u Hrvatskoj ovaj softver koristi Sveučilište u Rijeci, Sveučilište u Osijeku, Veleučilište VERN i Zagrebačka škola za menadžment te su započeti pregovori s nekoliciinom sastavnica Sveučilišta u Zagrebu. Više od 10.000 institucija cijeloga svijeta (135 država) koristi ovaj softver. Ovo je ujedno i najrasprostranjeniji softver.

UNPLAG	Prema podacima vlasnika softvera, u Hrvatskoj nema institucijskoga korisnika ovoga softvera iako su započeti pregovori s nekolicinom sastavnica Sveučilišta u Zagrebu. Najveći korisnici u Europi su sveučilišta u Njemačkoj, Italiji, Španjolskoj i Velikoj Britaniji, a u svijetu imaju bazu od oko 150 korisnika.
URKUND	Prema podacima vlasnika softvera, u Hrvatskoj nema institucijskoga korisnika ovoga softvera. Najveći korisnici u Europi su sveučilišta u Italiji, Portugalu, Njemačkoj, Švedskoj, Norveškoj, Austriji, a u svijetu iz Sjeverne Amerike. Ovaj alat bilježi rast krajnjih korisnika za 81.000 u posljednjih 5 mjeseci, a ukupno broje nekoliko milijuna korisnika. U tijeku su pregovori za korištenje softvera na sveučilištima u Srbiji.

3.6. CIJENA KORIŠTENJA SOFTVERA I LICENCE

Za odluku o odabiru softvera važan je model po kojemu se formira cijena (po broju korisnika, broju stranica, broju dokumenata, broju znakova po stranici i sl.) te mogućnost korištenja i naplate po pojedinoj ustanovi. U nastavku su navedene usporedna tablica i tablica s cijenama.

PLAGSCAN	Dva su modela naplate: 1. po broju studenata (neograničen broj provjera) 2. po broju analiziranih stranica/riječi (jedna stranica sadrži 275 riječi). Valjanost licence: jednu godinu. Cijena je javno dostupna na <i>web</i>-stranicama i jasno je određena. Provjera znanstvenih časopisa je uključena u cijenu ako je visoko učilište koje izdaje časopis korisnik softvera.
TURNITIN	Dva su modela naplate: 1. po broju studenata (neograničen broj provjera) 2. po broju stranica/riječi. Valjanost licence: jednu godinu. Za provjeru znanstvenih časopisa koristi se poseban sustav pod nazivom iThenticate koji se zasebno plaća.
UNPLAG	Dva su modela naplate: 1. po broju stranica/riječi (stranica sadrži 275 riječi) 2. po broju studenata (neograničen broj provjera). Valjanost licence: jednu godinu. Za uključivanje provjera znanstvenih časopisa potrebno je dostaviti informaciju o približnoj količini provjera godišnje kako bi se utvrdilo mogu li se uključiti bez dodatnih troškova ili za dodatnu naknadu.
URKUND	Tri su modela naplate: 1. po broju stranica/riječi 2. po broju studenata (neograničen broj provjera) 3. po broju dokumenata (50% cijene se plaća unaprijed, a ostatak u intervalima ili na kraju godine).

	Valjanost licence: jednu godinu. Provjera znanstvenih časopisa uključena je u cijenu ako je visoko učilište koje izdaje časopis korisnik softvera. Za časopise koji žele zasebno koristiti softver cijena licence je 800 €, a cijena provjere po dokumentu 2 €.
--	---

Cijene po modelima naplate (*dobivene 20. travnja 2016.godine*). Zelenom bojom označene su najniže cijene.

Softver	Po broju studenata (godišnja licenca)			
	150.000	50.000	6.000	2.000
PLAGSCAN	0,85 €	0,92 €	1,00 €	1,00 €
TURNITIN*	0,80 €	0,80 €	1,50 €	1,50 €
UNPLAG	2,00 €	2,10 €	2,25 €	/
URKUND	0,49 €	0,68 €	1,00 €	1,20 €

* Ove cijene vrijede u slučaju pregovora s Ministarstvom znanosti, obrazovanja i sporta.

Softver	Po broju stranica (godišnja licenca)	
PLAGSCAN	100 milijuna riječi / 18.939 €	200 milijuna riječi / 36.582 €
TURNITIN	/	/
UNPLAG	35.000 stranica / 0,05 €	70.000 / 0,04 €
URKUND	/	/

Softver	Po broju dokumenata (godišnja licenca)	
	35.000	70.000
PLAGSCAN	/	/
TURNITIN	/	/
UNPLAG	/	/
URKUND	1 € /po dokumentu/do 400.000 znakova	0,75 € /po dokumentu/do 400.000 znakova

3.7. AUTENTIKACIJA I ULOGE KORISNIKA

Važno je da nastavnici i studenti prilikom korištenja softvera za otkrivanje plagijata mogu koristiti svoje postojeće elektroničke identitete u sustavu AAI@EduHr za prijavu u softver. Oslanjanjem na postojeću infrastrukturu AAI@EduHr izbjegava se potreba za kreiranjem i održavanjem potencijalno velikoga broja korisničkih računa u softveru.

Kod analize je također uzeta u obzir granulacija uloga koje postoje u softveru (administrator na razini sustava, administrator na razini ustanove, mentor, student...) te mogućnost podešavanja ograničenja (engl. *quota*) po ulogama i korisnicima.

PLAGSCAN	Omogućeno je korištenje AAI@EduHr elektroničkih identiteta (podrška za Shibboleth / SAML / Active Directory). Dodatno – mogućnost nastavnika da generira ključ koji šalje studentu putem kojega može unositi radove direktno na stranici pa u tom slučaju nije potrebna prijava u softver. Uloge: • administrator • podadministrator (administrator ustanove) • nastavnik • student. Administratori imaju mogućnost pojedinačnoga i masovnoga kreiranja korisničkih računa. Unutar softvera administrator može podesiti kvotu za određene korisnike – broj riječi koji taj korisnik može provjeriti. Koriste bodovanje gdje 1 PlagPoint iznosi 100 riječi.
TURNITIN	Omogućeno je korištenje AAI@EduHr elektroničkih identiteta (podrška za Shibboleth). Uloge: • administrator • nastavnik • student.
UNPLAG	Nemaju podršku za Shibboleth/SAML već je potrebno kreirati lokalne korisničke račune. Uloge: • administrator • nastavnik • student.
URKUND	Omogućeno je korištenje AAI@EduHr elektroničkih identiteta (podrška za Shibboleth). Uloge: • administrator • nastavnik • student.

3.8. TESTIRANJE SOFTVERA I PROCJENA KVALITETE

Za potrebe analize i usporedbe softvera provedeno je testiranje na uzorku od 10 radova. Radovi su izabrani iz sljedećih kategorija:

- radovi objavljeni u časopisu
- konferencijski rad
- disertacija
- diplomski rad
- seminarski rad.

Iz svake kategorije izabrana su dva rada pri čemu je jedan rad izrađen u razdoblju od 2005. do 2006., a drugi od 2015. do 2016. godine.

Količina prepoznatih kopiranih sadržaja nije značajno odstupala prilikom provjere istoga dokumenta u svim softverima te se može zaključiti da su promatrani softveri u tom kontekstu slični. Brzina obrade pojedinoga dokumenta ovisi o količini teksta koji se u njemu nalazi. Svi su softveri obradili dokumente u sličnom vremenskom razdoblju.

Rezultati provjere dostupni su u tablicama koje slijede:

a) PLAGSCAN

Prepoznavanje citiranih dijelova	- citirane dijelove prepoznaje po podudarnosti unutar navodnih znakova - mogućnost isključivanja prepoznavanja citata
Ograničenja	- preuzimanje rezultata analize u obliku formata pdf, docx - na preuzetom word-dokumentu kopirani dijelovi (potencijalni plagijati) označeni kao komentar
Intuitivnost sučelja	9
Napomene	- prilikom isključivanja određenih izvora koji su inicijalno prepoznati kao plagijat prema novi izračunati postotak - mogućnost podešavanja raspona postotka koji se smatra plagijatom i pripadajuće boje (npr. 20 – 30% zelena boja, 30 – 60% žuta boja itd.) - <i>whitelist</i> (mogućnost isključivanja određenih <i>web</i>-stranica u provjeru) - mogućnost isključivanja unaprijed definiranoga teksta u provjerama (npr. „Ime i prezime“)

b) TURNITIN

Prepoznavanje citiranih dijelova	- mogućnost isključivanja prepoznavanja citata i navedene literature/bibliografije (<i>smart filters</i>)
Ograničenja	- u pretrazi ne prepoznaje dijakritičke znakove (č,ć,š,đ,dž,ž) te riječi koje ih sadrže ne uvrštava u izvještaj - student može poslati na provjeru jedan dokument s računala, Dropboxa ili Google Drive-a
Intuitivnost sučelja	7,5
Napomene	- mogućnost dodavanja glasovnih komentara studentima, ocjenjivanje pomoću rubrika, naznačavanje pogrešaka metodom povuci i ispusti s već predefiniranim pogreškama (u obliku oznaka) koje se najčešće pojavljuju u radovima studenata - mogućnost dodavanja brze povratne informacije nastavnika prema studentu (nastavnici sami definiraju svoje standardne odgovore/komentare) - bilježi se statistika za sveučilište (broj studenata, nastavnika, izvještaja, postoci plagijata podijeljeni na grupe gdje se u svaku bilježi broj dokumenata, mogućnost izvoza statistike u Excel) - mogućnost podešavanja sadržaja izvještaja (dostupnost izvještaja studentu, određivanje minimalnoga praga u postocima ili riječima koji će sustav prepoznati kao plagijat).

c) UNPLAG

Prepoznavanje citiranih dijelova	- citirani dijelovi se prepoznaju prema stilovima citiranja (MLA, APA...) - dojam je da kod testiranja svaku zagradu prepoznaje kao citat - mogućnost uključivanja citiranih dijelova
Ograničenja	- nema posebnih zapažanja
Intuitivnost sučelja	7,5
Napomene	- otkriva plagijat i u slučaju kad su pojedina slova u kopiranom tekstu zamijenjena nekim alternativnim fontom (npr. ako se englesko a zamijeni ruskim)

d) URKUND

Prepoznavanje citiranih dijelova	- ne prepoznaje citirane dijelove, ali postoji mogućnost uključivanja prepoznavanja tekstova unutar navodnih znakova i zagrada
Ograničenja	- prilikom testiranja jedan pdf-dokument nije bilo moguće analizirati (tehnički problem) - mogućnost provjere dokumenta s minimalno 450 znakova
Intuitivnost sučelja	8
Napomene	- prepoznaje i označava dijelove dokumenata koji su preuzeti iz radova na srpskom jeziku - iz sučelja nije dostupna mogućnost odabira pojedinačnih izvora provjere (npr. internetom, drugim dokumentom ili bazom) - prilikom pretrage softver pronalazi više izvora s istim tekstom, navodi najzastupljeniji, a u dio <i>Sources not used</i> postavlja one izvore kod kojih je manji postotak sličnosti) - prilikom isključivanja određenih izvora, smanjeni postotak ne sprema, već postavlja početni koji je softver generirao.

4. ZBIRNI REZULTATI TESTIRANJA

U nastavku se nalazi skupna tablica za brzi pregled mogućnosti svakoga sustava prema odabranim kriterijima.

Legenda:

-  (nije ispunjen ili ispod prosjeka)
-   (zadovoljavajući kriterij)
-    (dostupne dodatne mogućnosti).

Kriteriji	PlagScan	Turnitin	Unplag	Urkund
API i PLUG-IN	  	 	 	  
PlagScan se ističe kvalitetnim i dobro dokumentiranim API-jem. Urkund dodatak za Moodle jedino ima mogućnost ograničavanja uporabe po pojedinoj ustanovi na sustavu Merlin koji koristi više ustanova.				
SMJEŠTAJ SERVISA	  	 	  	 
PlagScan i Unplag imaju dodatnu mogućnost instalacije na lokalni poslužitelj.				
OBUHVAT BAZE I MOGUĆNOST UKLJUČIVANJA VLASTITOGA SADRŽAJA	  	 	 	  
PlagScan i Urkund korisniku omogućavaju punu kontrolu nad radovima koji se provjeravaju. Prednost Turnitina je velik broj uključenih časopisa u bazu, a nedostatak je to što radovi trajno ostaju u njihovoj bazi i ne mogu se brisati.				
PODRŠKA	 	 	 	 
Svi softveri uključuju podršku korisnicima. Iz naših kontakata posebno ističemo dobra iskustva s predstavnicima softvera PlagScan i Unplag.				
RASPROSTRANJENOST	 	  	 	 
Turnitin koristi nekoliko visokih učilišta u Republici Hrvatskoj.				
CIJENA KORIŠTENJA SOFTVERA I LICENCE	 	 		  
Urkund ima najniže, a Unplag najviše cijene.				
AUTENTIKACIJA I ULOGE KORISNIKA	  	 		 
PlagScan omogućava administraciju na razini ustanove (administrator ustanove), a Unplag nema podršku za integraciju sa sustavom AAI@EduHr.				
TESTIRANJE SOFTVERA I PROCJENA KVALITETE	9/10	7,5/10	7,5/10	8/10

5. ZAKLJUČAK

Konačan odabir softvera za plagiranje u velikoj mjeri ovisi o potrebama pojedine institucije. U ovom smo dokumentu nastojali dati pregled činjenica koje smatramo značajnima, ali i prikazali rezultate i opažanja prilikom praktičnoga testiranja mogućnosti pojedinoga alata.

Mogućnost integracije s postojećim informacijskim sustavima preko API-ja ili dodataka (*plug-in*) važna je značajka softvera za plagiranje. U ovoj analizi naglasak je stavljen na mogućnost integracije sa sustavima Dabar, Hrčak i Merlin (Moodle) koji se u velikoj mjeri koriste u akademskoj zajednici i koje održava Srce. Svi analizirani softveri za plagiranje podržavaju neki oblik integracije (API, LTI ili Moodle Plug-in). Detaljnijim uvidom u dostupnu dokumentaciju uočene su manje razlike u softverima – dojam je da sustav PlagScan ima dostupnu detaljniju dokumentaciju koja može ubrzati i olakšati integraciju te da pruža veće mogućnosti upravljanja parametrima upita. Dodatno, prilikom razmatranja mogućnosti integracije sa sustavom Merlin (Moodle) razmatran je scenarij u kojem bi uporaba sustava za plagiranje bila omogućena pojedinačnim ustanovama, korisnicima sustava Merlin, a ne na čitavom sustavu. Zadovoljavajuće rješenje za ovaj scenarij ponudio je jedino softver Urkund.

Ova analiza uglavnom se temelji na dostupnoj dokumentaciji i prezentacijama predstavnika, a za sigurnu potvrdu funkcionalnosti API-ja i dodatka za Moodle potrebno je napraviti instalaciju i testiranje na konkretnom sustavu, ali i utvrditi odgovarajuće scenarije uporabe.

Svakako jedno od važnih pitanja jest zadržavanje kontrole nad dokumentima koji se šalju na provjeru putem softvera za otkrivanje plagiranja. Svi razmatrani softveri nude mogućnost slanja dokumenata online uporabom njihove infrastrukture, a PlagScan i Unplug uz to nude i mogućnost instalacije na lokalne poslužitelje, ali uz nadoplatu. Prilikom provjere radova svi sustavi pretražuju resurse na internetu, zatim pronalaze dokumente iz baza kao što su Hrčak i Dabar. Predstavnici svih sustava otvoreni su za prijedloge i zahtjeve za uključivanje novih izvora koje će indeksirati. Svi alati imaju mogućnost dodavanja/uključivanja vlastitoga sadržaja (radova s pojedine ustanove) u njihovu bazu, s time da je Turnitin naglasio da se iz njihove baze studentski radovi ne mogu obrisati.

Gledajući rasprostranjenost uporabe softvera u trenutku izrade ove analize, od analiziranih softvera jedino se Turnitin koristi na nekim ustanovama u sustavu visokog obrazovanja u Hrvatskoj.

Cijene uporabe sustava se razlikuju i ovise o broju studenata koji će koristiti softver, pri čemu vrijedi pravilo da se za veći broj studenata cijena smanjuje. Neki vlasnici softvera otvoreni su za pregovore oko cijene. Praksa korištenja licence je godinu dana, počevši od željenoga datuma (nije nužno da to bude 1.1. ili početak akademske godine).

Zbog korištenja SAML/AAI@EduHr autentikacije u Republici Hrvatskoj na svim ustanovama u sustavu visokog obrazovanja, korisna nam je bila informacija o tome podržavaju li softveri takav oblik prijave te u kojoj su mjeri definirane uloge (administrator,

nastavnik i student) i ovlasti. Sustav PlagScan ima razrađene opcije administratorskoga sučelja s mogućnošću podešavanja niza parametara iz sučelja, dok Unplag još uvijek nema podršku za autentikaciju putem elektroničkih identiteta u sustavu AAI@EduHr.

Potrebno je napomenuti da softveri ne daju izvješće o tome je li neki dokument plagijat već pronalaze sličnosti s drugim dokumentima. Iz tog se razloga od nastavnika ili odgovorne osobe očekuje pregled rezultata analize s razumijevanjem.

Provedeno praktično testiranje softvera omogućilo nam je bolji uvid u mogućnosti alata, ali je važno napomenuti da se softveri redovito nadograđuju te je opažanja navedena u ovom dokumentu potrebno promatrati u kontekstu vremena u kojem je provedeno testiranje. U ovoj analizi razmatrana su četiri komercijalna softvera, ali treba uzeti u obzir da se zbog brzog razvoja tehnologije pojavljuju novi komercijalni softveri i softveri otvorenoga koda.

Komunikacija sa svim predstavnicima tvrtki protekla je odlično, osim s predstavnikom Turnitina koji je tek nakon dužega pregovaranja omogućio testiranje.

Nakon provedene analize kao ozbiljne kandidate za primjenu u sustavu znanosti i visokog obrazovanja preporučujemo PlagScan ili Urkund. Za donošenje konačne odluke o odabiru softvera potrebno je provesti dodatno testiranje u koje treba uključiti korisničku zajednicu (nastavnici, studenti) te predstavnike ustanova. Istovremeno je potrebno donijeti politike kojima bi se osigurali organizacijski preduvjeti za primjenu softvera za plagiranje.

6. PRILOG – PREGLED API-JA

6.1. PLAGSCAN

<https://www.plagscan.com/api-guide>

Dokumentacija: Online, primjeri za Java, PHP, .NET.

Sigurnosni zahtjevi:

- SSL
- Ograničenje IP adresa (pojedinačno, raspon IP adresa)

Rezultati u obliku:

- Samo statistika (razina plagiranja, broj riječi)
- Tekst dokumenta i linkovi za rezultate
- XML s linkovima na pronađene izvore
- Docx-dokument s oznakama plagiranja,
- HTML-dokument s oznakama plagiranja,
- HTML-izvješće,
- PDF verzija HTML-dokumenta i izvješća

Parametri rezultata:

- PlagScan ID analiziranoga dokumenta, korisnikov ID, broj riječi, datum, status analize (pauzirano, u tijeku, završeno, čekanje u redu)
- Razina plagiranja, ime datoteke i naziv
- Prikaz prvih 85 znakova sadržaja

Konfiguracijski parametri:

- Jezik (engleski, njemački, španjolski)
- Određivanje "žute" i "crvene" granice plagiranja (u postocima)
- Email izvješće (nikad, uvijek, samo ako je u "crvenom" području, tj. visoka razina plagiranja)
- Kreiranje Docx-dokumenta (generiraj i pošalji e-mailom, samo generiraj, ne generiraj)
- Automatsko pokretanja analize (da / ne)
- Analiziraj s izvorima sa interneta (da / ne)
- Analiziraj s ostalim izvorima (ne / mojim dokumentima / mojom institucijom / cijelom bazom podataka)
- Osjetljivost analize (niska, srednja, visoka)
- Izbriši rezultate nakon (tjedna, 4 tjedna, 6 mjeseci, nikad)

6.2. TURNITIN

Podrška za Learning Tools Interoperability (LTI):

[https://guides.turnitin.com/03_Integrations/Learning_Tools_Interoperability_\(LTI\)](https://guides.turnitin.com/03_Integrations/Learning_Tools_Interoperability_(LTI))

Napomena: Koristi LTI (*Learning Tools Interoperability*) umjesto standardnoga API-ja

Dokumentacija: Online

6.3. UNPLAG

<https://unplag.com/api/doc/>

Dokumentacija: Online PDF (s registriranim korisničkim računom)

Sigurnosni zahtjevi:

- SSL
- Ograničenje IP adresa (pojedinačno, lista IP adresa)

Rezultati u obliku:

- JSON, XML, MsgPack

Parametri:

- Pohrana dokumenta
 - Format datoteke
 - Datoteka
 - Ime datoteke
- Brisanje dokumenata
 - Id datoteke
- Pokretanje analize
 - Id datoteke za provjeru
 - Id-evi datoteka za provjeru Doc-vs-Doc
 - Izostavljanje citata (da / ne)
 - Izostavljanje referenci (da / ne)
- Dohvaćanje informacija o dokumentu (Id, Broj riječi, Ime, Format, Broj stranica)
- Dohvaćanje informacija o analizi (Id, Cijena, Vrsta, s koliko je dokumenata uspoređeno, Datum, Status, Progres)
- Dohvaćanje informacija o rezultatima (Datum, Sličnost, Broj izvora, Broj citata, Broj referenci, Link na URL)
- Izrada Izvješća PDF-a
 - Id datoteke
 - Jezik
- Dohvaćanje linka za online izvješće
- Uključivanje / Isključivanje Citata i Referenci
- Dohvaćanje trenutnoga stanja (Progres)

Podrška za *Learning tools Interoperability* (LTI):

<https://www.imsglobal.org/activity/learning-tools-interoperability>

6.3. URKUND

Dokumentacija: Na zahtjev

Sigurnosni zahtjevi:

- SSL

Rezultati u obliku:

- JSON, XML

Parametri:

- Pohrana (Id, Datum i vrijeme, Ime datoteke)
- Status (Pohranjen, Odbijen, Odobren)
- DocumentInfo (podaci o dokumentu)
 - Id dokumenta
 - Datum
 - Link za preuzimanje dokumenta
- ReportInfo (podaci o rezultatima ako je provedena analiza)
 - Id izvješća
 - Link za preuzimanje izvješća
 - Razina plagiranja
 - Broj podudaranja teksta
 - Broj izvora podudaranja
 - Upozorenja o greškama u dokumentu
- ReceiverInfo
 - E-mail adresa
 - Ime i prezime

7. REFERENCE

Glendinning, I. (2015.). *Promoting Maturity in Policies for Plagiarism across Europe and beyond*. Prezentirano na: 7th Prague Forum: Towards a Pan-European Platform on Ethics, Transparency and Integrity in Education, Prag.

Glendinning, I. (2014.). *Assessing maturity of institutional policies for underpinning academic Integrity*. International Integrity and Plagiarism Conference. Održano od 16. do 18. lipnja 2014. u The Sage Gateshead, UK. Dostupno na: <https://core.ac.uk/download/files/169/30620175.pdf> (2.7.2016.)

Pan-europska platforma za etičnost, transparentnost i integritet u obrazovanju (ETINED) Dostupna na: <http://www.coe.int/en/web/ethics-transparency-integrity-in-education>.

Stappenbelt, B. i Rowles C. (2009.). *The effectiveness of plagiarism detection software as a learning tool in academic writing education*. Prezentirano na: 4th Asia Pacific Conference on Educational Integrity (4APCEI), Wollongong. Dostupno na: <http://ro.uow.edu.au/apcei/09/papers/29/> (26.7.2016.)



(Analiza softvera za otkrivanje plagiranja u znanosti i obrazovanju.docx)